Education Endowment Foundation

Realistic Maths Education

Evaluation Report

September 2022

Sean Demack, Dr Martin Culliney, Professor Mark Boylan and Claire Wolstenholme

Sheffield Hallam University | Sheffield Institute of Education

The Education Endowment Foundation (EEF) is an independent grant-making charity dedicated to breaking the link between family income and educational achievement, ensuring that children from all backgrounds can fulfil their potential and make the most of their talents.

The EEF aims to raise the attainment of children facing disadvantage by:

- identifying promising educational innovations that address the needs of disadvantaged children in primary and secondary schools in England;
- evaluating these innovations to extend and secure the evidence on what works and can be made to work at scale; and
- encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus - Private Equity Foundation) and received a founding £125m grant from the Department for Education. Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.

For more information about the EEF or this report please contact:

Jonathan Kay
Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP

0207 802 1653

jonathan.kay@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents

About the evaluator

The Sheffield Institute of Education is recognised for excellence and innovation in research, teaching, and learning. It has been developing and delivering educational innovations and evaluations locally, nationally, and internationally for over 25 years as one of the country's most successful providers of initial teacher education in one of the country's largest universities.

Corresponding author: Sean Demack (Principal Investigator) s.demack@shu.ac.uk

Acknowledgements

Linda Bray, Nick Speed, Caroline Cripps, and Freya Wallace all provided valuable administrative assistance.

Executive summary

The project

Realistic Mathematics Education (RME) is an approach to teaching mathematics that builds on students' informal sense-making in response to imaginable and meaningful contexts. Contexts are chosen for their potential to elicit flexible mathematical models and teachers employ interactive teaching techniques to enable students to develop their understanding and ability to work with mathematics in novel situations. RME training supports teachers in the development of a new classroom culture that enables extended discussion of multiple contexts and students' representations of them, the generation of multiple strategies and models, and sharing, explaining, and discussion of strategies.

The Manchester Metropolitan University (MMU) RME intervention was targeted at Year 7 and Year 8 pupils (11 to 13) and their mathematics teachers who participated in RME professional development over two years and delivered the intervention in class. The programme was originally designed to consist of eight days of continuing professional development (CPD) and ten modules of curriculum materials each designed to provide a minimum of two weeks teaching materials. Two days of CPD (one poorly attended, one cancelled) and two modules for Year 8 pupils were disrupted by the Covid-19 pandemic. Additional online CPD and alternative materials were provided, extending into the first term of Year 9.

RME was evaluated through a two-arm, three-level efficacy trial with pupils clustered into classes and classes clustered into schools. A total of 119 schools took part in the study; 60 were randomised to the RME intervention and 59 to the control group. The study aimed to assess the impact of the RME approach on mathematics attainment using the GL Progress Test in Mathematics (PTM13). The implementation and process evaluation (IPE) comprised interviews with teachers from intervention and control schools, CPD visits, a series of teacher surveys, and the analysis of management and delivery information and data collected by the delivery team. The trial recruited from January to September 2018 and delivery began in October 2018 when pupils were in Year 7, following them through to Year 8. Delivery was extended into the first term of Year 9 due to Covid-19 in order to offset delays to outcome testing, which was originally scheduled for the summer term of 2020 but was delayed until spring 2021 and again until summer 2021.

Table 1: Key conclusions

Key conclusions	
1.	Pupils in RME schools made, on average, no additional progress in mathematics compared to similar pupils in the control group as measured by GL PTM13. This result has very low security rating.
2.	There was no evidence to suggest that the programme had a differential impact on the GL PTM13 results of pupils eligible for free school means.
3.	Movement of ITT pupils between Year 7 and Year 8 classes was observed, meaning that only 43.5% of ITT pupils in the RME intervention schools received both years of the programme. This finding seems to present a challenge for the future design of RCTs for evaluating programmes in secondary schools that focus on change at a teacher or class level over two years.
4.	Some teachers who engaged in RME professional development were enthusiastic about RME as an approach to teaching mathematics and the professional development they experienced. There was some evidence that engagement changed self-reported teacher beliefs and practices when compared to teachers in the control condition.
5.	There were indications that the model for professional development of two teachers attending external professional development workshops over two years was challenging to implement in some schools due to ongoing pressures.

EEF security rating

These findings have a very low security rating. This was an efficacy trial, which tested whether the intervention worked under developer-led conditions in a number of schools. Attrition was high: 48% of the pupils who started the trial were not included in the final analysis. In addition to the partial closure of schools and other Covid-19 related disruptions, outcome testing was postponed from summer 2020 to summer 2021. This Covid19 context provides some explanation for the high attrition. In addition to delayed outcome testing and attrition, and prior to the arrival of Covid19, notable pupil movement between maths classes during Year 7 and Year 8 was observed. Specifically, only 43% of the ITT sample for the RME intervention group were located in a maths class taught by an RME teacher in both Year 7 and Year 8.

Additional findings

Pupils in RME schools made, on average, no additional progress compared to those in the control group equivalent across any of the six PTM13 subscales. This is our best estimate of impact, which has a very low security rating. As with any study, there is always some uncertainty around the result: the possible impact of this programme ranges from small negative effects of one month less progress and positive effects of up to two months of additional progress. The trial was disrupted due to the Covid-19 pandemic, which affected training day attendance and delivery.

Prior to randomisation, schools provided details on at least two Year 7 maths teachers who would participate in RME if their school was selected, along with class lists for all Year 7 maths classes and the teachers attached to these classes. The Intention-To-Treat (ITT) pupil sample was defined by identifying the pupils in classes taught by the teachers nominated by their schools to take part in RME. This resulted in an ITT sample of 7,956 pupils across the 119 schools.

While the study found no evidence that teacher-level fidelity to RME (in terms of teacher attendance on training days, self-reported use of materials, and time spent teaching RME) resulted in an impact on pupil-level attainment, there seemed to be a correlation between the number of days of training attended and pupils' attainment. ITT pupils in RME intervention schools that sent two or more teachers to at least four RME training days had higher maths attainment compared with ITT pupils in RME intervention schools that sent two or more teachers to training on fewer than four occasions.

The RME CPD approach is built around MMU professional development leaders (PDLs) modelling the use of teaching materials and resources alongside commentary on their approach. During these periods of modelling, participant teachers took the role of students. Participants particularly praised the modelling of lessons and the opportunity to 'act out' the teaching and learning themselves.

Cost

The average cost per pupil per year over three years is estimated at £32.22.

Impact[1]

Table 2: Summary of impact on primary outcome(s)

Outcome/ Group	Effect size (95% confidence interval)	Estimated months' progress	EEF security rating	No. of pupils	P Value	EEF cost rating
GL PTM13 (ITT)	+0.04 (-0.06; +0.15)	0		4,159	0.438	
GL PTM13 (FSM ITT)	-0.03 (-0.17; +0.11)	0		1,152	0.711	

[1] The impact evaluation of RME was done using statistical data from ONS. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

Introduction

Background

Rationale[2]

Based initially on the ideas of Hans Freudenthal, Realistic Mathematics Education (RME) is a pedagogical theory developed in the Netherlands over the last 40 years. RME uses realistic contexts and a notion of progressive formalisation to help the mathematical development of pupils. It is internationally recognised, and materials based on RME are used in many countries (De Lange, 1996; Van den Heuvel-Panhuizen and Drijvers, 2014).

The approach was taken up in England in 2004 by Manchester Metropolitan University (MMU) initially with Key Stage 3 (KS3) pupils (from 2004 to 2007 as 'Mathematics in Context') and later with KS4 pupils (2007 to 2010), particularly those studying towards foundation tier GCSE maths. The latter development involved, with Mathematics Education and Industry, the team producing a set of textbooks and related materials under the title *Making Sense of Maths*, published by Hodder.

In the following sections, ways that RME pedagogy is distinctive are described; details and examples of these features of the MMU approach to realistic maths education are available at **https://rme.org.uk/**

Use of context

Rather than being used primarily to illustrate the applicability and relevance of maths (as is customarily the case in English maths classrooms), context is used as a sustained underpinning for developing mathematical understanding (Fosnot and Dolk, 2002; Treffers, 1987). Contexts are carefully chosen for their potential for model building and teachers encourage students to develop and refine their informal mathematical strategies and models. Contexts can be taken from the real world, from fiction, or from an area of maths that students are already familiar with, but they need to be imaginable for students, and so able to support engagement with purposeful mathematical activity (Gravemeijer, Van den Heuvel-Panhuizen and Streefland, 1990).

Use of models

Models are used to bridge the gap between informal understanding and formal systems in ways that allow:

- the formal and informal to 'stay connected' in the minds of students; and
- mathematical activity at differing levels of abstraction—this enables learners who find more formal concepts difficult to engage with to make progress and develop strategies for solving problems.

The meaning of 'model' extends to 'materials, visual sketches, paradigmatic situations, schemes, diagrams, and even symbols' (Van den Heuvel-Panhuizen, 2003).

Mathematising

RME promotes two ways of 'mathematising': (i) solving the contextual problem under consideration ('horizontal mathematisation') and (ii) working within the mathematical structure itself by reorganising, finding shortcuts, and recognising the wider applicability of their methods: this is called 'vertical mathematisation' (Treffers, 1987).

Multiple strategies

Both formal and informal strategies are valued and, consequently, multiple strategies feature in RME lessons. The aim is for students to become more mathematically efficient and sophisticated over time; this efficiency and sophistication is

---

[2] The Realistic Maths Education (RME) study was proposed by the Manchester Metropolitan University (MMU) RME team and then details of implementation were further developed in collaboration with the Sheffield Hallam University (SHU) evaluation team and the Education Endowment Foundation (EEF) grant and evaluation management team. This included an intervention design and analysis workshop (Humphrey et al., 2016) that took place in February 2018. The main study materials design and training elements were proposed by the RME team. The independent randomisation, impact, and implementation and process evaluation data collection and analysis of the RME study are the responsibility of the independent evaluation team, led by Sean Demack (principal investigator) and Mark Boylan (co-investigator). The background rationale and description of RME draws on previous publications by the MMU team and collaborators (these include Barmby et al., 2011; Dickinson et al., 2010; Hough et al., 2017).

not taken as a starting point but as 'on the horizon' (Fosnot and Dolk, 2002), and informal methods are still valued (Webb et al., 2008).

Redefining progress

In RME, 'progress' is defined in terms of the progressive formalisation of models (Van den Heuvel-Panhuizen, 2003), particularly the progression from 'model of' to 'model for' (Streefland, 1985). Initially, the model is very closely related to the specific context being considered but eventually becomes a model which can be applied in numerous mathematical situations (Van den Heuvel-Panhuizen, 2003).

As a result of the above distinctive features, classroom cultures where teachers follow RME pedagogy differ from those based on more prevalent forms of maths teaching. RME classroom culture is underpinned by different socio-mathematical norms (Yackel and Cobb, 1996) that involve more active mathematising by pupils and discussion and sharing of strategies. Specifically, an RME classroom culture might be marked by:

- time spent talking about context, which might appear to be non-mathematical;
- time spent generating/discussing various representations of a context;
- systematic provision of spaces in which students are invited to talk about their strategies for solving even apparently straightforward questions;
- teacher questions that are open rather than closed; and
- student willingness to initiate/share/discuss/question/explain strategies.

Learners will potentially be more engaged as maths is more meaningful and some alienating features of regular teaching diminish (see Nardi and Steward, 2003). However, affect in maths classrooms is complex. It has been noted that RME may be resisted by students who have become used to particular maths classroom cultures: while students may not like their usual classroom experience, they are familiar with it and may have developed strategies that enable them to participate on some level (Hough, Solomon, Gough and Dickinson, 2017).

Evidence of potential impact

RME is employed by over 80% of schools in the Netherlands, which is considered one of the highest-achieving countries in the world in maths (TIMSS 1999, 2007, 2010; PISA 2000, 2006, 2009, 2015). In PISA 2015, the Netherlands was ranked 11th in maths (compared to the U.K. in 27th place). However, in science the Netherlands was ranked 17th (behind the U.K.) and in reading it was ranked 15th. Thus, this indicates that national mathematical success in the Netherlands may be due to the specifics of maths education and not the education system in general. An international comparison of numeracy levels among 16- to 18-year-olds by the OECD showed the Netherlands in second position and the U.K. in 17th (Department for Business, Innovation and Skills, 2013).

An evaluation of the application of RME in the U.S.—Mathematics in Context—led to a large number of detailed monographs and other outputs.[3] Comparison of intervention classes, with high fidelity of implementation, with classes that followed conventional teaching identified an association between RME and higher achievement (Romberg and Shafer, 2005).

MMU's previous implementation of RME in England has also been evaluated through comparative analysis of the performance of 50 Year 7 pupils who had experienced RME and 50 comparison group pupils. The pupils experiencing RME did so over a year through the use of Mathematics In Context textbooks during an initial pilot in 2004/2005 (Dickinson and Eade, 2005). Details of the extent to which RME methods and materials were used by their teachers are not provided in the report. Analysis focused on a test of students' capacity to solve nine problems and explain their approaches. This found that the intervention group were both more likely to solve the problem and to be able to explain their methods (Searle and Barmby, 2012). In 2014 to 2016, RME was further extended for use with post-16 pupils taking GCSE resit examinations with one class in each of three colleges participating (Hough, Solomon, Gough and Dickinson, 2017).

An independent evaluation of the post-16 GCSE project involving delivery of a number module (12 hours) and an algebra module (nine hours) found an initial positive effect on intervention students for the number module, but not of the algebra module, in comparison with students not experiencing RME. However, the positive effect was not sustained in a delayed

[3] http://micimpact.wceruw.org/

post-test when students were retested three months later (Boylan and Jay, 2017). Thus, overall, the evidence for an effect from this study was mixed.

Practice context

RME pedagogy and materials differ from those commonly used in England in their emphasis on sustained use of realisable contexts and model building to visualise mathematical processes so that learners make sense of what they are doing. RME also differs from regular teaching in that formal maths develops out of students' informal mathematising, rather than the primary lesson objective being acquisition of a formal process. The process of formalisation in RME seeks to maintain a link back to the original context and model that students worked with (Van den Heuvel-Panhuizen, 2003; Hough, Solomon, Gough and Dickinson, 2017).

Maths education traditions in England have had a well-documented impact on classroom cultures and on student experiences and expectations. Many young people see maths as a question of learning rules that lead to answers based on received wisdom and the authority of the teacher (De Corte, Op 't Eynde, and Verschaffel, 2002). It can often be seen as irrelevant to everyday life and as meaningless and abstract (Boaler, 2002). A premise for RME approaches is that patterns of classroom interaction that are fostered by a traditional transmissionist approach to teaching maths can lead students to have lower expectations of themselves as well as of maths. It is widely acknowledged that many students become disaffected with school maths (Swan, 2006; Nardi and Steward, 2003; Lewis 2011). More meaningful maths may lead to enhanced student engagement and interest.

RME arguably aligns with more recent developments such as the new emphasis at GCSE on solving non-routine problems and, consequently, the implications of this for early secondary maths. Current interest in East Asian maths education under the title of 'teaching for mastery' (Boylan et al., 2018) has led to an increased interest in the use of models and representations, mathematical meaning, and developing the capability of pupils to apply maths to non-familiar contexts (Drury, 2014). Current encouragement to spend more time on mathematical topics to deepen understanding (NCETM, 2014) directs attention to pedagogies that slow down the process of formalisation. RME potentially addresses this.

The interest in East Asian maths education has also led some schools to explore alternatives to grouping pupils by prior attainment (setting or streaming) as a way of addressing the 'long tail' of low achievement in maths that is more prevalent in students who are socioeconomically disadvantaged (Jerrim and Choi, 2014). It is believed that the RME approach may be particularly well suited to all types of attainment teaching contexts (Hough, Solomon, Gough and Dickinson, 2017).

Several RME-type practices are advised in recent guidance for KS2 and KS3 maths teaching (EEF, 2017; see also Hodgen, Foster, Marks and Brown, 2018). The following evidence-based recommendations appear closely aligned with RME: the use of representations, focusing on problem-solving strategies, developing pupils' independence, and using tasks and resources to challenge and support pupils' maths.

Although there are some indicators that RME pedagogy may improve outcomes, there is an absence of robust evidence. The previous studies in England have been evaluated using a variety of mixed-method designs (Searle and Barmby, 2012; Boylan and Jay, 2017). However, RME in England had not been subject to a randomised controlled trial. By undertaking a trial of RME, evidence might also be gained on other related pedagogical approaches of current interest in practice and policy in England. The intervention was designed for lower secondary pupils, so building on and extending MMU's previously designed curriculum materials and drawing on their experience of working with teachers and departments in secondary schools.

Integrated evaluation design

RME is a whole-class teaching approach requiring professional development for teachers. Therefore, it was evaluated with a two-arm, three-level efficacy trial with pupils clustered into classes and classes clustered into schools. The IPE comprised CPD observations, interviews with teachers from intervention and control schools, and a series of teacher surveys to assess implementation in practice and changes in teacher beliefs.

In each intervention school, at least two teachers participated in professional development and then taught curriculum materials to Year 7 pupils starting from October 2018. Ideally, the same two teachers would then teach the same pupils in Year 8, until the end of 2019/2020, using materials designed for Year 8 pupils, and continue with CPD. Anticipating that some of the original nominated teachers would be unavailable (due to staff turnover, absence, or other staffing requirements), schools nominated replacement teachers to teach the same pupils who experienced RME in Year 7, and to attend the training programme, to provide continuity.

Recipients were mainly pupils in classes taught by the teachers nominated to participate in RME training prior to randomisation. The primary Intention to Treat (ITT) impact analysis comprised pupils located in a Year 7 maths class that were taught by any of the RME teachers nominated prior to randomisation. However, it was expected that due to movement between classes over the two years, some pupils would have a partial experience of RME. Thus, all (then Year 9) pupils in the intervention schools were tested in summer 2021 and this data was included within the impact evaluation analyses; the original design was for testing of these pupils when in Year 8 in summer 2020 but this was delayed due to the Covid-19 pandemic.

The Covid-19 pandemic created issues for the delivery and evaluation of RME. The intervention was delivered as intended up to February 2020. However, as schools were closed from March 2020 until the end of the academic year,[4] RME delivery (to May/June 2020) was disrupted and a final training day in May 2020 was cancelled. In response, MMU delivered two sets of training sessions online to support any RME teaching that might be taking place either online or in classrooms where a small minority of students were still being taught during the summer 2020 school closures. However, based on interview and survey data, the classroom engagement with RME materials was close to nil during this time and online RME teaching very patchy. Almost all schools had neither begun nor delivered the final (tenth) module (on functions/algebra), and the preceding (ninth) module (on data) had also not been completed by a number of schools. Additionally, MMU offered additional optional online training to support an additional ‘connections’ module (see Modifications section).

Intervention

In this section, the design of the intended intervention is outlined. Covid-19 caused considerable disruption including, for an extended period, closure of schools for most pupils from March 2020. Changes to the original design, including adaptations to offer remote professional development, are described below under ‘modifications’.

1. Brief name
Realistic Maths Education.

2. Why
The intervention aimed to improve the quality of maths teaching in lower secondary school. It provided professional development that aimed to prepare teachers to confidently use RME materials and be able to adapt other existing maths materials to extend the approach to their overall maths teaching.

3. What (materials)
The intervention curriculum materials were organised in five modules per year, the modules being in both Year 7 and Year 8, and sequenced in the following order in Year 7. In Year 8 the sequence changed so that proportional reasoning and geometry changed places:

- number
- geometry
- proportional reasoning
- data
- algebra.

Modules were loosely divided into lessons which comprised sets of activities. As well as providing printed copies of materials at CPD events, participants were also able to access material on the project website.

[4] Schools were closed to most pupils from March 2020 but for children of key workers and other children, schools remained open. For simplicity, when we refer to school closures we mean partial closure—where most pupils did not experience any in-person schooling.

Teachers were provided with teacher guides and supporting materials (PowerPoints, worksheets) which should have enabled them to teach five two-week blocks of RME in each of Years 7 and 8 as a minimum. Ways this was integrated or undertaken alongside other teaching are indicated in implementation and process evaluation reporting below.

4. What (procedures)

1. Professional development

Teachers involved in the intervention received specialist training on how to use RME materials and develop their practice over the two years of the trial. Professional development focused on both effective use of the curriculum materials and on developing the underlying RME pedagogical principles. Professional development combined off-site CPD days and in-school tasks referred to as 'gap tasks' (detailed in Table 3, below). CPD days were based around curriculum materials with participants first experiencing materials from a learner's perspective and then examining them from a teaching perspective. These two aspects were used to develop RME teaching approaches such as the use of context. The intervention can be considered a 'curriculum professional learning' innovation (Boylan and Demack, 2018) in that the CPD itself was intended to lead to pedagogical change, and so to pupil impact, but that also the use of the curriculum materials has a potential professional learning effect as well as directly impacting on pupil learning. Professional development took place in regional clusters (see Trial Design below).

2. RME teaching

Curriculum materials were organised into materials for Year 7 and then Year 8.

Each lesson or set of activities (outlined above) may have lasted more than one teaching lesson depending on lesson timing. Each module consisted of approximately eight lessons' worth of materials. As the approach promotes flexibility, a lesson might last one hour but could be extended to two hours or more. One lesson should involve a mean of approximately three activities. Overall, the module design involved giving teachers more material than they might need for two weeks, and teachers were encouraged to take more time if they felt it appropriate. For this reason, some teachers did not complete all activities in a module if they were sticking to a two-week block only or if they had limited extra time.

5. Who (provider)

RME curriculum materials were based on materials previously developed by MMU, which updated and refined the materials for current teaching contexts.

6. How

Professional development components were:

- an eight-day course—teachers attended up to eight training days between October 2018 and April 2020; the course involved demonstration lessons, training in RME principles (such as the role of context, mathematical 'landscapes', and pedagogy), reflections on their teaching of previous modules, and discussion of teachers' experience in 'gap tasks';
- personal and collaborative development—between sessions, teachers worked with the RME materials and, following guidance given during sessions, worked collaboratively to evaluate their practice and undertake gap tasks;
- the use of RME materials and scheme of work guidance for Year 7 and Year 8 (see below, Curriculum Materials); and
- access to online resources—teachers had access to a dedicated website to support them in gap tasks and access materials.[5]

A key component of the professional development was the gap tasks, which encouraged teachers to reflect on their RME practice in school with their paired teacher (see implementation and process evaluation reporting for detail about how this was enacted in practice). These were undertaken between CPD sessions and focused on developing diagnostic insights and pedagogic practices which supported an RME classroom culture. The aims and purposes of gap tasks are summarised below.

---

[5] All Year 7 and Year 8 materials are available at https://rme.org.uk/

Table 3: RME CPD gap tasks

Gap task aim/purpose	Specific issues addressed
Encourage focus on teachers' pedagogic practice through 'noticing'	Time spent using activities and resources 'Wait time' and related issues Response to student strategies and contributions Orchestrating and responding to disagreement/mathematical argument Physical positioning in the classroom
Develop diagnostic skills in analysing student work:	Understanding student approaches and strategies Moving away from deficit views of student work Understanding how materials do or do not support development
Practice adapting existing materials towards an RME approach	Use of local contexts/interests Thoughtful selection of materials and activities Reflection on how materials progress through ideas Anticipating how students respond to tasks

The table above reports the intended gap tasks. With regard to the last row—'adapting existing materials towards an RME approach'—implementation of the gap tasks in practice varied from this. In general, the idea of adapting existing materials towards an RME approach was not emphasised. An important focus was thinking about the teaching and planning by annotating module teaching guides and developing skills in analysing student work and generally reflecting on their teaching and pedagogic practice. In addition, teachers created 'postcards' to put up in their classrooms to remind them of particular strategies.

Gap tasks frequently involved observation but it was anticipated that teachers may often not find time to observe each other. Gap tasks focused on specific aspects of RME pedagogy and encouraged teachers to think about this in their lesson planning and work on this aspect in their lessons, and in lesson observations with their paired teacher. This practice draws on and aimed to develop 'the discipline of noticing' (Mason, 2002) in terms of teachers' thinking about how students' progress is supported in a lesson and how their practice enables this. The gap tasks were designed to enable teachers to focus on one strategy at a time in order to build towards RME 'fluency'.

7. Where
CPD took place in six geographical hub areas and teachers then worked in their own schools to teach RME. In summer and autumn 2020, teachers participated in online CPD.

8. When and how much
CPD days took place between autumn 2018 and spring 2020, in October and Dec 2018, February, April, June, and October 2019, and February 2020. Face to face CPD planned for later than March 2020 did not take place due to Covid-19 restrictions.

Teaching took place over two school years, between autumn 2018 and spring 2020. The planned summer 2020 teaching was curtailed. As described above, each module consisted of approximately eight lessons' worth of materials and teachers were encouraged to spend at least two weeks of curriculum time per module. In Year 7, curriculum material could potentially have been taught as planned. During Year 7, a total of ten weeks of RME-specific material was available, so replacing approximately 25% of usual teaching material. In Year 8, due to Covid-19 restrictions the maximum amount of teaching time that available materials would cover would be 20% but in practice the timing of module use meant that the amount would not reasonably exceed 15%.

9. Tailoring
In addition to refinement of materials specifically for this intervention, the CPD framework included one 'catch up' CPD day in September 2019 with incoming Year 8 teachers to support continuity of RME experience for pupils whose Year 7 teachers did not continue with the programme.

10. Modifications
After the introductory CPD days, five schools were identified by MMU as ‘design schools’ and MMU undertook visits to these schools to support refinement of materials and development of CPD materials.

During the programme, from the first year, less emphasis was placed on joint completion of gap tasks and tasks were modified to allow individual engagement rather than collaboration between teachers in the same school. The rationale for this change was the challenge of undertaking tasks together due to time restrictions and teacher availability.

As a modification to the planned approach, a small number of schools decided to use curriculum materials across whole year groups and so used cascade models where teachers who participated in CPD days shared their learning and introduced materials to colleagues. MMU did not specifically encourage or discourage this happening. This increased variation in implementation is discussed in IPE reporting below. The reasons for such modifications by schools related to the approach to curriculum planning across KS3 and wanting to maintain a common experience for all pupils.

In March 2020, due to Covid-19, schools were closed to most pupils. However, before this happened, the public health crisis already impacted on some teachers’ participation in training day seven, delivered in February/March 2020.

School closure affected the use of the Year 8 ‘data handling’ module in some schools and the Year 8 algebra module in almost all schools. The final planned training day in May 2020 was cancelled. Alternative online training was offered to support teachers using RME materials in remote teaching. This was offered on a hub basis with two sessions taking place in May and June for each hub, with each session lasting two hours and using a video conferencing platform.

In recognition of the disruption to RME teaching in Year 8, the use of curriculum materials and training was extended into Year 9 with the provision of an optional ‘capstone’ module (‘making realistic connections’). This additional module was not part of the original planned RME curriculum. It was designed to consolidate and connect RME approaches to the topics covered in Year 7 and Year 8. The module had four parts, the first being available to teachers in July 2020 and parts two to four in October, November, and December 2020 respectively. Additional associated training was offered for teachers of Year 9, comprising three twilight 90- or 120-minute sessions which took place in October, November, and December, offered on a combined hub basis with teachers able to choose from three dates on each occasion.

Theory of change

Causal mechanisms

The intervention involved a number of potential change agents and related mechanisms or processes that may lead to change either directly or indirectly in pupil outcomes; these are described in Table 4. The latter two mechanisms are both posited as intermediate outcomes and causal processes.

Table 4: Theory of change mechanisms

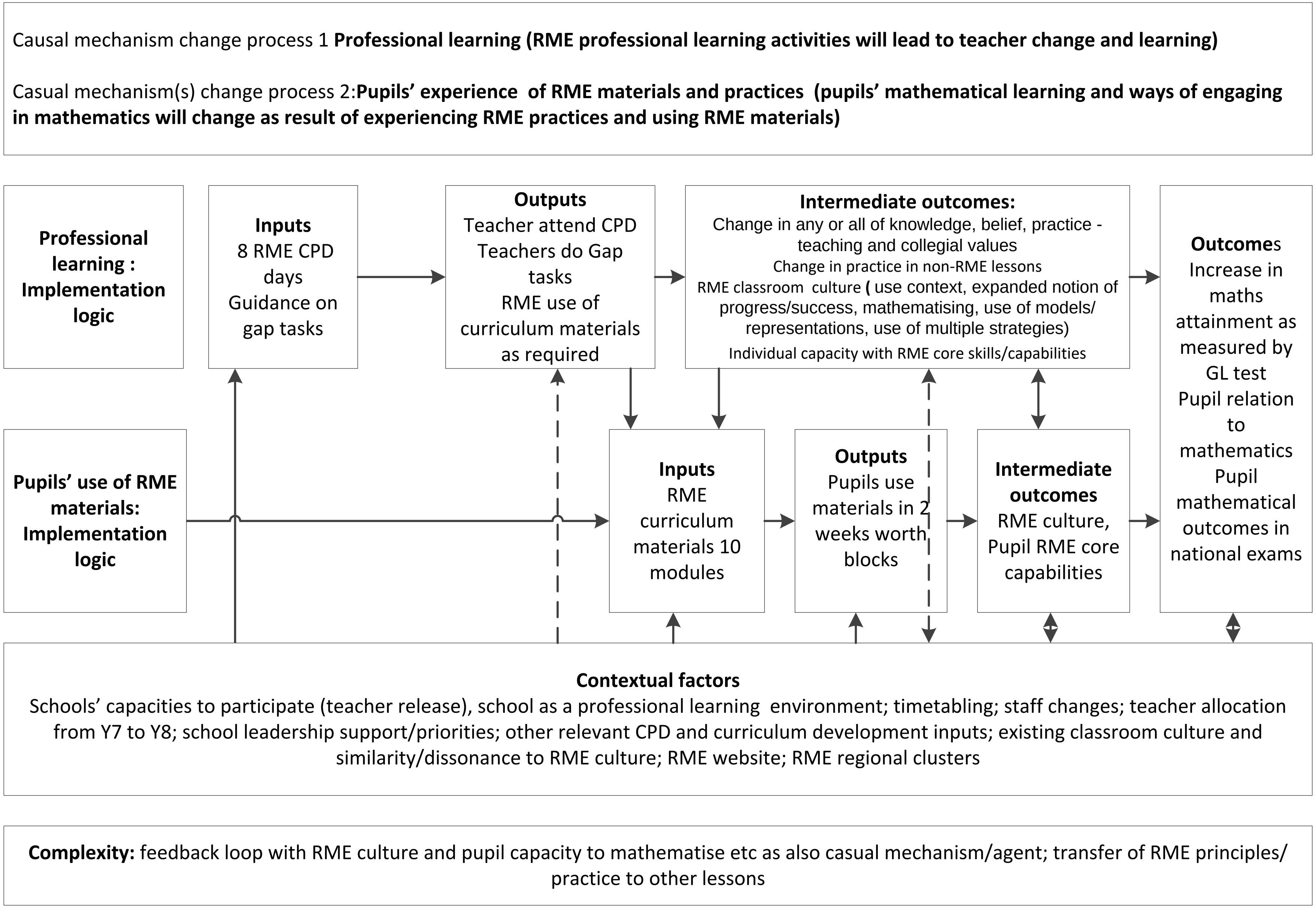
Causal agent	Mechanism/process	Outcome/desired change (intermediate or endpoint)
RME professional development events	Indirect: professional learning	Teacher: change in any or all of knowledge, belief, practice (teaching and collegial), values
In school PD tasks	Indirect: professional learning	Teacher: as above
Use of RME principles in other maths lessons	Indirect: professional learning	Change in any or all of knowledge, belief, practice—teaching and collegial values

Use of RME materials	Indirect 1: professional learning Indirect 2: pupil and class change Direct: pupil engagement with RME maths	Teacher: as above Class: foster RME culture—use of context, expanded notion of progress/success, mathematising, use of models/representations, use of multiple strategies, pupil listening and explaining Pupil: capacity to use context, expanded notion of progress/success, mathematising, use of models/representations, use of multiple strategies, capacity to explain and engage in teacher led discussion including listening to other pupils' explanations
Development of RME culture (use of context, expanded notion of progress/success, mathematising, use of models/representations, use of multiple strategies)	Direct: enhanced pupil mathematical activity	Pupil mathematical learning and capacity to use context, expanded notion of progress/success, mathematising, use of models/representations, use of multiple strategies

Logic model

Figure 1 provides a logic model that links main inputs and outputs. For simplicity RME culture is presented as an intermediate outcome only as is the use of RME principles in other lessons. Two potential causal processes are foregrounded in this logic model: RME professional development as an indirect change process on pupil outcomes and the use of RME materials as a direct process

Figure 1: Logic model



Evaluation objectives

Impact evaluation

Primary question

RQ1. Does the RME intervention improve pupil attainment in maths over two years as measured by the GL Assessment Progress Test in Mathematics (PTM13) for the intention to treat group compared to the control group, in general and specifically for FSM pupils?

Secondary questions

RQ2. What is the impact of RME on maths attainment for pupils known to have been taught by one of the nominated RME teachers throughout the trial period?

RQ3. What is the impact of RME on maths attainment for pupils who experience partial intervention effects due to pupil and/or teacher movement during the trial period?

RQ4. What is the effect on attainment of components of the GL PTM13 most closely related to the RME curriculum material content and on questions identified as related to problem solving?

RQ5. What is the relationship between maths attainment and fidelity of implementation?

RQ6. What is the impact of RME on maths attainment for pupils taught by teachers identified as implementing the RME evaluation with high fidelity?

Implementation and process evaluation questions

RQ7. What is the fidelity of participation in CPD and implementation of RME school-based gap tasks and use of materials by nominated teachers?

RQ8. What has influenced variation in implementation by nominated teachers?

RQ9. What are teachers' perceptions of differences between RME practices and those they used prior to the intervention and what are the differences in reports of practices between teachers participating in RME and those in control schools?

RQ10. How do teachers change as a result of participation in the intervention, and in particular in relation to beliefs about maths and teaching maths, and self-reported changes (if any) in maths pedagogical knowledge, pedagogical content knowledge, and maths content knowledge?

RQ11. What issues are important, if any, for the security of the trial outcome, replication, and scalability?

The trial protocol and statistical analysis plan are publicly available on the EEF website:
https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/realistic-maths-education

Ethics, trial registration, and data protection

Details of ethics, trial registration, and data protection are contained in the Memorandum of Understanding, parent/carer letter, and Fair Processing Notice. These are included in Appendix E.

Ethics

The RME evaluation was submitted to the SHU ethics committee and received approval in July 2018 (SHU Ethics Review ID: ER6803657).

In line with General Data Protection Regulation (GDPR) guidelines, schools were not obliged to seek parental consent unless they preferred to do so for their own purposes because this evaluation was conducted on the legal basis of a public task. However, parents/carers were informed that they could ask for their child's data to be excluded at any stage of the trial. We suggested wording for the parent/carer communication sent out by schools in September 2018. This related to participation in the trial, linking to NPD data, and participation in the GL PTM13 test in June 2020. Separate permission was sought from parents/carers for the IPE data collected.

Updated information letters for parents/carers were issued in autumn 2020 and spring 2021 to describe the changes to the trial due to Covid-19, including the opportunity to withdraw from rescheduled outcome testing. Changes to the trial timescale, data collection, testing period, and sharing the GL PTM13 pupil data with schools was submitted for ethical approval in December 2020, which was approved in January 2021 (SHU Ethics review ID ER6804043).

Registration

The trial was publicly registered (ISRCTN21086912) on 31/05/2019.[6]

Data protection

SHU acted as data controller for the duration of the project. The EEF will be data controller once the data has been submitted to its data archive. A Fair Processing Notice was sent to all participating schools—as per General Data Protection Regulation (GDPR) requirements that came into force from 25 May 2018—specifying the personal data to be processed (pupil names, pupil UPN, FSM status). For information on how SHU treats data from research participants, please see the University's **privacy notice**.

Pupil data supplied to the evaluation team by schools was linked with information about the pupils from the National Pupil Database (NPD) to access baseline test data (KS2 maths results) and identify FSM pupils for subgroup analysis. Further matching to NPD data may take place during subsequent research through the EEF data archive.

Special category data (specifically, FSM status, accessed from the NPD) was collected and processed for the purpose of scientific research as permitted under GDPR Article 9 (j). As the EEF was established with a remit to break the link between family background and educational attainment, all EEF projects conduct subgroup analysis on FSM pupils.

As an additional incentive due to the Covid-19 extension, schools were offered outcome test data for their pupils. A parent information sheet distributed before the test period stated that results would be available to schools. Two schools contacted the evaluators to request the data for their pupils. This was sent using a secure transfer service.

Project team

MMU RME team

Principal investigator: Yvette Solomon, Professor of Education

RME CPD lead: Sue Hough, Senior Lecturer in Mathematics Education

RME design lead: Steve Gough, Senior Lecturer in Mathematics Education

Project manager: Vinay Kathotia, Senior Research Associate

RME trainers: Jo Kennedy, Senior Lecturer in Mathematics Education; Fiona Haniak-Cockerham, Senior Lecturer in Mathematics Education; Franke Eade, former Principal Lecturer in Mathematics Education; Marisa Bartoli, Laurus Trust Director of Mathematics

Project administrator: Jo Dennis until November 2018

Research assistant and administrative support: Kate O'Brien

SHU Evaluation team

Principal investigator and impact evaluation lead: Sean Demack, Principal Research Fellow

Trial statistician and data manager: Dr Martin Culliney, Senior Research Fellow

Co-investigator and IPE lead: Professor Mark Boylan, Professor of Education

Project manager and IPE fieldworker: Claire Wolstenholme, Research Fellow

IPE fieldworker: Bernie Stiell, Senior Research Fellow

---

[6] https://doi.org/10.1186/ISRCTN21086912

Methods

Trial design

The impact of RME on pupil maths attainment was evaluated through a two-armed, three-level, multisite clustered RCT that accounts for clustering of the GL Progress Test in Mathematics 13 (PTM13) outcome at the school and classroom levels. The design was blocked by area to reflect the six geographical locations (or clusters) of RME training (Yorkshire and Northeast, North West, West Midlands, East Midlands, London, South) but, as an efficacy trial, no variation in the impact of RME was assumed across these training locations.

The 119 recruited schools that submitted the specified data were randomised to the intervention (60 schools) or control (59 schools) group in October 2018. To be included in the randomisation, recruited schools were required to supply (1)—before summer 2018—the names of at least two nominated Year 7 maths teachers who would participate in the RME programme if the school was randomised to the intervention group, (2) detail on the school's setting/streaming policy for KS3 maths, and (3)—in September 2018—pupil lists for all Year 7 maths classes along with the names of teachers allocated to teach these classes.

Following randomisation, the nominated teachers in each intervention school participated in RME CPD training and taught RME curriculum materials to Year 7 pupils starting from October 2018. Ideally, the same teachers would teach the same pupils in both Year 7 and Year 8, until the end of 2019/2020. In intervention schools, having teacher/class-pupil consistency in Year 7 and Year 8 was encouraged within the school MOUs but some teacher changes were still anticipated within a two-year programme. In the RME intervention schools, when the original nominated teachers were unavailable (due to staff turnover, absence, or other reasons), schools nominated replacement RME teachers to continue with the programme. In the control schools, a business-as-usual approach was adopted with no encouragement for teacher/class-pupil consistency through Year 7 and Year 8.[7]

Complete teacher/class-pupil class lists were collected for all Year 7 maths classes in all 119 schools before randomisation. The Intention to Treat (ITT) sample comprises pupils located in the 328 Year 7 maths classes taught by the teachers nominated prior to randomisation.

The extent to which pupils would be with the same teachers/classes from Year 7 into Year 8 was unknown at the start of the evaluation although as mentioned it was recommended that teachers would teach the same pupils across both years of the trial. To explore this within the RME delivery period, teacher/class-pupil lists were collected for all Year 8 maths classes during the first term of 2019/2020 and prior to the arrival Covid-19. Within the extended evaluation period teacher/class-pupil lists were also collected for Year 9 maths classes in the first term of 2020/2021 but participation in this final collection of class lists was very low. In addition to exploring Year 7 to Year 8 teacher/class-pupil consistency, data on maths attainment data for the complete year group at outcome and baseline enabled exploration of how full and partial classroom exposure to RME was associated with maths attainment.

[7] Business as usual teaching is reported in the implementation and process evaluation results in the Usual Practice section.

Table 5: Trial design

Trial design, including number of arms		Two-armed, multisite, clustered randomised controlled trial
Unit of randomisation		School
Stratification variable (s) (if applicable)		Geographical area (six areas) School phase (secondary/middle) Use of setting/streaming in Year 7 maths (yes/no)
Primary outcome	Variable	Maths attainment
	Measure (instrument, scale, source)	GL Progress Test in Mathematics (PTM13)
Secondary outcome(s)	Variable(s)	Subscales of PTM13 maths attainment
	Measure(s) (instrument, scale, source)	GL PTM excluding 'mental maths' test items Selection of 11 of the 23 GL PTM test items closely aligned to RME (see SAP Appendix IV) GL PTM subscale1: fluency in facts and procedures GL PTM subscale2: fluency in conceptual understanding GL PTM subscale3: problem solving GL PTM subscale4: mathematical reasoning
Baseline for primary outcome	Variable	Maths attainment
	Measure (instrument, scale, source)	KS2 maths score (NPD variable KS2MAT_SCORE, measured on 80-120 scale)
Baseline for secondary outcome(s)	Variable	Maths attainment
	Measure (instrument, scale, source)	KS2 maths score (NPD variable KS2MAT_SCORE, measured on 80-120 scale)

The control condition in this efficacy trial was business as usual. Control schools that participated in the data collection activities were offered an incentive payment of £1,000. Additionally, MMU and the EEF intended to host a one-day event in summer 2020 for staff from each control school after outcome testing but this did not happen according to the original schedule due to Covid-19 restrictions, being delayed until July 2022. Control schools also had full access to materials and guidance at the end of the trial, as they became publicly available after outcome testing in July 2021. A further £250 payment was offered to all schools (intervention and control) taking part in outcome testing in summer 2021 in return for continuing in the trial after Covid-19 closed schools for most pupils from March to September 2020 and from January to March 2021. All participating schools were offered access to results data from the outcome test, GL PTM13 (described below), so that they could use this to support pupil maths learning.

Participant selection

Recruitment for this trial was overseen by the MMU RME team. To minimise variation at the school level, the aim was to recruit 120 non-selective, established secondary schools for the evaluation. This criterion was relaxed towards the end of June to expand the offer to middle schools and newly opened schools to maximise the chance of reaching the target of 120 schools. Of the final 119 recruited schools, eight were middle schools with the oldest pupils being in Year 8 (four intervention and four control).

If randomly selected into the intervention group, schools were required to commit at least two teachers to the RME CPD training. The names of these nominated teachers were collected for all 119 recruited schools before randomisation. Additionally, teacher/class pupil lists were collected for all Year 7 maths classes for all 119 recruited schools. The pupil-level ITT sample was identified using this data. Specifically, the list of nominated teachers was used to identify ITT pupils who were in a Year 7 maths class taught by one or more of the nominated teachers (=1) or not (=0).

Initially, it was estimated that nominated teachers would teach a mean of three Year 7 maths classes of 25 pupils in the 119 schools (8,925 pupils overall; 4,500 in the 60 RME intervention schools and 4,425 in the 59 control schools). In the observed data, nominated teachers taught a mean of 2.8 Year 7 maths classes of 24 pupils in the 119 schools (7,956 pupils overall, 4,035 intervention and 3,921 control).

Outcome measures

Baseline measures

To maximise statistical precision, KS2 maths attainment (from tests taken in summer 2018) was used as a baseline covariate. This measure was obtained from the NPD (KS2MAT_SCORE) and used in the analysis of all outcome variables. The multilevel models included the baseline covariate at pupil, class, and school level. Pupil-level baseline scores were centred around the class mean, class baseline scores were centred around the school mean, and school baseline scores were centred to the grand mean. As noted by Demack (2019), one reason for centring the KS2 maths covariate at the three levels is to ensure that the pupil, class, and school baseline variables are uncorrelated, therefore avoiding problems of multicollinearity within the impact analyses.

Primary outcome

The primary outcome measure was maths attainment assessed through the GL Progress Test in Mathematics (PTM), administered in school with test materials provided by the evaluators. GL PTM is a commercial standardised assessment of pupil mathematical skills and knowledge[8] and was selected as the primary outcome for four reasons. First, it is well known and widely used by schools and teachers. Second, it is aligned to the maths national curriculum. Third, it assesses pupil maths attainment at the end of Year 8. Fourth, it assesses maths attainment in terms of curriculum areas and mathematical capabilities, which in turn facilitates the creation of secondary outcomes for the impact evaluation in order to address RQ4. During the evaluation, PTM curriculum areas and mathematical capabilities closely related to the RME curriculum were selected and published in the statistical analysis plan in spring 2019.

In response to the Covid-19 pandemic and school closures, the evaluation was extended into Year 9 and outcome testing was rescheduled from summer 2020 to spring 2021. Following the second period of partial school closures in spring term 2021, outcome testing was rescheduled again to summer 2021. This prompted consideration of which GL PTM test to use as the primary outcome.

Under normal circumstances, the most suitable GL PTM test for pupils in the spring term of Year 9 would be PTM14. According to GL, the PTM14 test is designed for pupils in spring/summer terms in Year 9 or autumn term in Y10.[9] The original primary outcome for the RME evaluation was GL PTM13, which is designed for pupils in the spring/summer terms of Year 8 or in the autumn term of Year 9. School closures during the spring/summer terms of Year 8 and ongoing need for pupils and teachers exposed to Covid-19 to stay away from schools since the start of Year 9 led to the decision to maintain GL PTM13 as the primary outcome rather than change to PTM14. There are three (related) reasons for this decision. First, the PTM14 test includes items that test mathematical knowledge and understanding for the curriculum following the Year 7 and Year 8 RME delivery period. Second, as most pupils experienced highly disrupted schooling between March and July 2020, the progression into the Year 9 maths curriculum is likely to have been delayed. Third, Year 9 pupils were at the very start of the recommended age range for PTM14 and likely to have had limited exposure to much of the curriculum it is designed to test even without taking account of the impact of Covid-19. An additional benefit for using PTM13 relevant to the Covid-19 context was that identifying gaps in mathematical learning in Year 8

[8] www.gl-assessment.co.uk/products/progress-test-in-maths-ptm/
[9] See: https://www.gl-assessment.co.uk/support/ptm-product-support/progress-test-in-maths-test-level-age-guide/

and earlier is likely to be of greater value to intervention and control schools as opposed to confirming gaps in the Year 9–10 curriculum not yet covered.

Following the second period of school closures in January and February 2021, the decision to maintain the PTM13 outcome was revisited. Testing was rescheduled to take place in May and June 2021, which pushed the testing period further into the PTM14 age band. However, the additional disruption to education in early 2021 raised questions over the validity of the GL PTM age bands within the pandemic context. The decision was to maintain the PTM13 primary outcome for similar reasons to those discussed above, yet we acknowledge that the suitability of GL PTM13 is open to question and have therefore scrutinised the marks distribution for ceiling effects.

Outcome test administration was also adapted in response to Covid-19. Due to restrictions on university travel and outside visitors, and to minimise disruption for schools, GL PTM was not independently invigilated as had been intended. Schools administered the assessment to Year 9 pupils under test conditions. Papers were sent to schools by the evaluation team by post. Upon completion, these were returned to the evaluators who then forwarded them to GL for marking. Results were supplied to the evaluators in a spreadsheet in late August 2021.

Secondary outcomes

A total of six secondary outcomes were examined (see Table 5, above). Each of these were subscales taken from the PTM13 primary outcome test. It should be noted that these subscales have not undergone the same psychometric evaluation that the overall GL PTM13 score has and are not psychometrically designed to stand alone. Therefore, analyses of secondary outcomes should be considered exploratory.

First, the GL PTM primary outcome includes a ‘fixed time’ mental maths section. This runs counter to the RME ethos, which encourages pupils to take time in order to develop their depth of mathematical understanding. Therefore, the first secondary outcome for the RME evaluation is the GL PTM outcome with all 20 of the mental maths items removed. Second, MMU selected 11 of the 23 GL PTM13 written test items that were closely related to the RME curriculum as published in the statistical analysis plan. The second secondary outcome is the test score based just upon these 11 GL PTM13 test items. The other four secondary outcomes are the GL specified subscales for PTM13. These are: fluency in facts and procedures, fluency in concepts, problem solving, and mathematical reasoning.

For all secondary outcomes, the same modelling stages for the primary analyses were undertaken using the same ITT sample.

Sample size

Minimum Detectable Effect Size (MDES) estimates were dependent on how the variance of the primary outcome variable was structured. The strength of clustering at the school and class levels determined the smallest effect size that the MSCRT design could detect as statistically significant ($p < 0.05$) with a statistical power of 0.80 for a set number of participants and clusters. At the randomisation and protocol stages, the strength of clustering of the primary outcome at the school and class levels was unclear. Demack (2019) highlighted how the strength of class-level clustering depends on the use of setting/streaming in Year 7 and Year 8 maths across the sample. From data collected prior to randomisation, 106 of the 119 schools in the trial (89%) reported setting or streaming in Year 7 and/or Year 8 maths. This led us to assume that half of the variation in the GL PTM outcome will be within schools, between classrooms and 15% to 20% will be between schools. The class ICC of 0.50 draws on recommendations from Demack (2019) for trials evaluating secondary maths programmes and is comparable to what was found by Boylan et al. (2015) with a GL PIM13 outcome in a Year 8 pupil sample.[10] The proportion of variance assumed to be at the pupil level (within classes between pupils) was therefore between 30% and 35%.

We estimated the correlation between covariates and the outcome to be 0.70 at all three levels ($R^2 = 0.49$). This was a conservative estimate drawing on Boylan et al. (2015) who found a much stronger correlation between KS2 and GL PIM (12, 13 and 14) at school, class, and pupil levels. As the class-level clustering is largely a result of school policies that

---

[10] For the Year 8 cohort, Boylan et al. (2015) reported a class level ICC of 0.47. GL Progress in Mathematics (PiM) was replaced by GL Progress Test in Mathematics (PTM) in 2015. For Year 8 measures (PiM13 and PTM13) a correlation of 0.83 between the two is reported (page 12 of GL Assessments PTM technical report).

sort pupils (at least in part) along attainment lines, it was anticipated that the class-level correlation would exceed 0.70. The school-level correlation was more unknown. As Boylan et al. (2015) included grammar and secondary modern schools along with comprehensives, their estimates for school-level explanatory power (R^2 = 0.88 for Year 8) may well be an artefact of the sample. The RME sample only included comprehensive middle and secondary schools and is therefore not directly comparable.

Table 11 (below, Impact Evaluation section) presents the MDES summary and sample sizes at protocol, randomisation, and analysis stages. Protocol estimates were based on predicted numbers of pupils per class (25) and classes per school (3). At the randomisation stage, data was collected directly from schools before randomisation. At both protocol and randomisation stages, the estimated MDES for the primary outcome was between 0.21 and 0.23. For the FSM subgroup, the MDES estimate was between 0.22 and 0.24. The indicative MDES estimates from the analysis stage suggest that the design had higher sensitivity with an estimate of 0.15 standard deviations overall and 0.16 to 0.17 standard deviations for the FSM subgroup. The gain in sensitivity came from covariate explanatory power at the class level (estimated in protocol and SAP as $R^2_{class} = 0.49$, observed as 0.90), which reflects how class-level clustering in the primary outcome was largely accounted for by prior attainment, in other words, a result of setting or streaming.

The impact of Covid-19 on trial sensitivity

After schools reopened in September 2020, a number of those in the sample did not continue their involvement with the trial. In the eight middle schools, this was because pupils were no longer at the school they attended in Year 7 and Year 8. Other schools withdrew due to the impact of Covid-19. Additional financial incentives were offered in acknowledgement of the additional time commitment from schools as discussed elsewhere in this report.

With specific reference to the impact evaluation design, Covid-19 resulted in particularly high attrition in outcome testing and a forced late change to the timeline that delayed the outcome testing by about a year.

After the eight middle schools withdrew in summer 2020, 111 secondary schools remained in the trial (56 intervention and 55 control) and GL PTM13 outcome data was collected summer 2021 in 79 of these schools (71%).

While RME delivery was completed in Year 8 (with the final months disrupted by Covid-19), the outcome testing did not take place until late Year 9; in other words, Covid-19 resulted in an enforced and unforeseen lag of approximately one year between the end of the RME programme and outcome testing. Even with the additional Year 9 'refresher' online sessions run by MMU, this lag may have resulted in some dissipation of the impact of RME on pupil attainment—perhaps even more so given other distractions and anxiety caused by Covid-19, although this is not certain.

Randomisation

Schools were randomised within each of the six geographical areas (listed above). As this was an efficacy trial, the design was not powered to detect variation in effect sizes across geographical areas. The fourth (multisite) level is accounted for in the design via the degrees of freedom for the impact evaluation power analyses. Geographical area was not included as a level in the model but is included in the analysis through dummy covariates. Randomisation took place at the school level to simplify recruitment and implementation and to minimise spillover risk. School-level randomisation does not eliminate the possibility of spillover, yet it provides stronger protection than within-school randomisation.

Prior to randomisation, schools that signed the MoU were required to submit two sets of data. First, between June and September 2018, schools were required to specify whether setting or streaming policies were used in Year 7 maths and the names of at least two nominated teachers to participate in the RME programme if their school was randomly selected to the intervention group. Second, in September 2018, schools were required to submit pupil lists for all Year 7 maths classes attached to named teachers. As part of the IPE, a baseline teacher survey was also undertaken in September 2018. Completion of this survey was voluntary and so not a requirement for randomisation, but the survey was closed prior to randomisation to ensure that teacher responses were not influenced by their school's allocation.[11]

[11] The RME delivery partners at MMU requested that the survey be reopened following this in order to maximise the baseline response. We obliged this request but, for the evaluation, only used data collected prior to randomisation.

A stratified randomisation approach was adopted that included controls for geographical hub region, whether a school was a middle or standard secondary school and whether a school currently had a policy of using within-school selection (setting/streaming) to sort pupils into maths classes in Year 7. Using these three controls best ensured that the RME intervention and control samples were balanced across the geographical hub regions, had similar numbers of middle and secondary schools, and schools that had within-school selection policies.

A total of 129 schools were recruited to the RME evaluation and signed a Memorandum of Understanding. Of these, 119 provided pupil data to the evaluation team and were entered into the randomisation, which took place on 5 October 2018. All schools were notified of their allocation immediately except for ten that were late in submitting the required data. Their allocation was withheld until this was received. Nine schools were informed of their allocation before the first week of RME CPD training and one was informed between the first and second training days.

Statistical analysis

Primary analysis

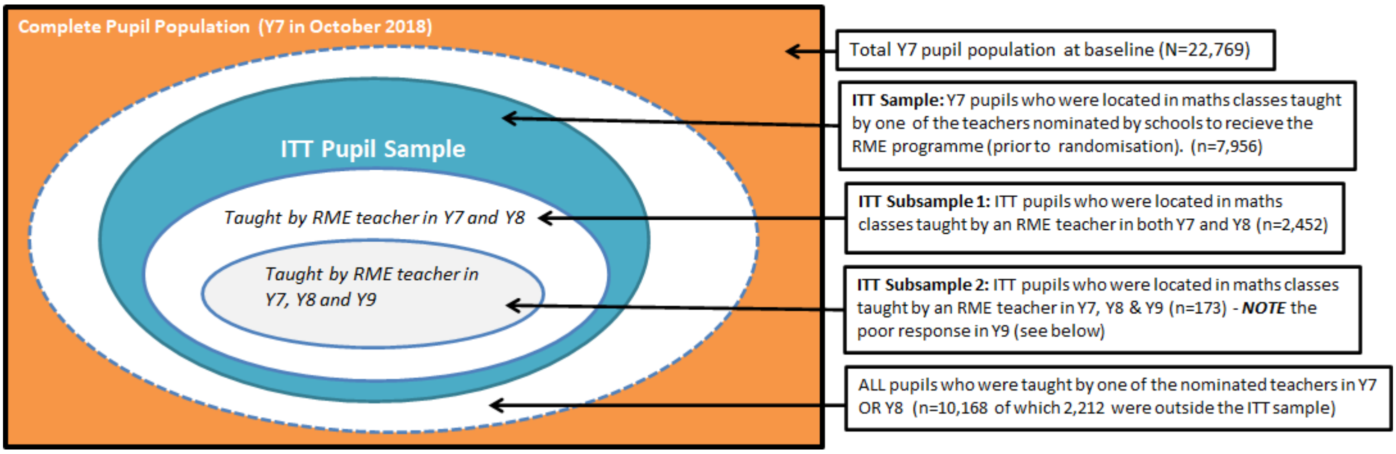
A multilevel approach was adopted with pupils clustered into classes and classes clustered into schools (three-level random intercepts multilevel models). Multilevel linear regression models were constructed for the primary outcome using Stata software version 15. KS2 maths attainment was used as the baseline covariate. The first model included the school-level group identifier (an outcome only model). The second model added KS2 maths attainment as a covariate at the pupil, class, and school level. The final model also included the three variables used within the stratified randomisation (geographical hub area, setting/streaming dummy, and middle/secondary dummy). This final model was used for the headline ITT impact analysis.

For each model, the coefficient of the school-level dummy variable was used to distinguish intervention pupils from control pupils. This coefficient was converted into Hedges' g effect size statistics with 95% confidence intervals as detailed below.

Secondary analysis

Follow-on exploratory analyses of the primary outcome drew on the teacher/class-pupil lists collected prior to randomisation in Year 7 in the first term of Year 8 prior to the arrival of Covid-19 and in Year 9 following the arrival of Covid-19. These lists were used to examine pupil movement between classes and, for the intervention ITT group, classroom exposure to RME in Year 7 and Year 8. This links to the Venn diagram shown as Figure 1 in the evaluation protocol and Figure 2 in the SAP (replicated as Figure 2 below).

Figure 2: Pupil samples and subsamples for the RME impact evaluation



Initially the Venn diagram related to just Year 7 and Year 8 (the RME delivery period) but this was expanded into Year 9 following the extension to the evaluation because of Covid-19. However, in spite of regular communication by MMU with over 30 of each of the control and intervention schools, there was a poor response to requests for the Year 9 teacher/class-pupil lists during the Covid-19 pandemic (only 23 schools provided data) and the Year 9 details are only used to complete the central ellipse in Figure 2. Given that no Year 9 data was collected for 96 of the 119 schools in the trial (eight of which are middle schools and therefore have no Year 9 pupils), this detail needs to be treated with caution.

Table 6 provides additional details on the samples and subsamples shown in Figure 2. At baseline, there was a total of 22,769 Year 7 pupils located in maths classes across the 119 schools (11,582 in intervention schools and 11,187 in control schools). Of these, 7,956 were located in a maths class taught by one of the nominated RME teachers (the ITT sample: 4,035 in intervention schools and 3,921 in control schools).

Table 6: ITT and other pupil samples (populating Figure 2)

		Intervention	Control	Total
Total	All Year 7 pupils	11,582	11,187	22,769
ITT	In Year 7 maths class taught by nominated RME teacher	4,035	3,921	7,956
ITT Sub 1	In Year 7 AND Year 8 maths taught by RME nominated teacher	1,754	698	2,452
ITT Sub 2	In Year 7, Year 8 and Year 9 maths taught by RME nominated teacher	120	53	173
Some (ITT)	In class taught by RME nominated teacher JUST in Year 7	1,711	2,113	3,824
Some (ITT)	In class taught by RME nominated teacher in Year 7 but missing Year 8 data	570	1,110	1,680
Some (outside ITT)	In class taught by RME nominated teacher JUST in Year 8	1,479	733	2,212
Any (in and outside ITT)	In a maths class in Year 7 OR Year 8 taught by an RME nominated teacher	5,514	4,654	10,168
None	No evidence of being in maths class taught by RME nominated teacher in Year 7 or Year 8	6,068	6,533	12,601

Teacher/class-pupil data for Year 8 maths classes was collected from 100 of the 119 schools in the first term of the 2019/2020 academic year, prior to the arrival of Covid-19. It should be noted that this data for maths classes was collected once in Year 7 (before randomisation) and again in Year 8 (first term) and therefore only shows movement between these two data collection points. However, the data provides two snapshots of pupil class compositions along with their assigned teacher(s) from which some detail on the movement of ITT sample pupils could be observed.

Of the 4,035 Year 7 pupils located in the ITT sample across the 60 RME intervention schools, only 1,754 (43.5%) were observed to have been in a maths class taught by an RME teacher (nominated or replacement) in Year 8 and therefore to have received both years of the RME programme (Year 7 and Year 8). There were 1,711 pupils (42.4%) observed to be in a Year 8 maths class taught by a non-RME teacher and for 570 pupils (14.1%), the Year 8 maths teacher data was missing.

Of the 3,921 Year 7 pupils located in the ITT sample across the 59 control schools, only 698 (17.8%) were observed to be in a Year 8 maths class taught by one of the teachers nominated at baseline in Year 7. A total of 2,113 pupils (53.9%) were in a Year 8 maths class taught by a different teacher. For 1,110 pupils (28.3%), Year 8 maths teacher data was missing.

In summary, we found that less than half of the ITT pupil sample in intervention schools received the second year of the two-year RME programme. This has implications for the validity of our ITT impact analysis for RME and potentially more widely for the design of future RCTs evaluating secondary maths programmes. Reasons for this observation relate to teacher or pupil movement between Year 7 and Year 8 maths classes. This might be because of the introduction or reconfiguring of sets/streams/classes. A comprehensive examination of this is beyond the scope of this evaluation report but it is worth highlighting that the Year 7 ITT sample was observed to disperse from 328 Year 7 maths classes at baseline into 631 Year 8 maths classes a year later.[12] As this observed issue of pupil movement is before the arrival of Covid-19, it seems that the teacher- and class-level focus of RME is methodologically problematic within the structural contexts of secondary maths for RCT design.

While the extent of teacher/pupil movement was unknown, the SAP anticipated this by specifying follow-on exploratory analyses that focused on pupil exposure to RME in the classroom. These analyses drew on the teacher/class-pupil data.

RQ2 for the evaluation focuses on the subsample of 1,754 ITT intervention pupils known to have had RME classroom exposure in both Year 7 and Year 8. For these analyses, two control groups were used: first, the full ITT control group sample (n = 3,921) and, second, the ITT control sample known to be in a class taught by a nominated teacher in Year 7 and Year 8 (n = 698). The reason for selecting the two control conditions was the slight distinction between Year 7 and Year 8 teacher/class-pupil consistency for intervention (nominated or replacement) and control (nominated only). As this was not clearly specified in the SAP we consider these to be the two most feasible control groups in these exploratory analyses.

RQ3 focused on the sample of pupils in intervention schools who were in an RME classroom in Year 7 or Year 8. This comprised the 4,035 ITT pupils along with 1,479 pupils outside the Year 7 ITT sample but listed as being taught by an RME teacher in Year 8 to form the sample of 5,514 pupils in the 60 intervention schools known to have had some classroom exposure to RME in Year 7 or Year 8. For these analyses, the full ITT control group sample (n = 3,921) was used as specified in the SAP.

In addition to these follow-on secondary analyses of the PTM13 primary outcome exploring classroom exposure to RME, exploratory analyses of six subscales of PTM13 were undertaken. Two of these were constructed by the evaluation team to adapt the GL PTM13 to be more closely reflective of the RME approach and four were subscales provided by the GL testing company.

---

[12] If the Year 7 sample is restricted to just schools where teacher/class-pupil data was collected in Year 8, the dispersion is wider: from 272 Year 7 into 631 Year 8 maths classes.

Analysis in the presence of non-compliance

The purpose of the Compliers Average Causal Effect (CACE) analyses was to estimate the impact of RME for pupils deemed to have 'complied' with the intervention during the delivery period (Year 7 and Year 8). In this instance 'compliance' related to both pupils and their maths teachers and to classroom RME exposure and teacher fidelity to the RME intervention. Specifically, to be considered 'compliant', a pupil must have been located *consistently* in a maths class taught by a nominated RME teacher who attended at least four of the eight RME training days, used RME in maths lessons for at least ten weeks (20% or more of total available pre-Covid-19 maths teaching time in Year 7 and Year 8), and covered at least six of the ten RME modules. These thresholds were revised downwards in response to Covid-19 disruption (see Table 7 below) While these thresholds are at two levels (pupil and teacher), the final compliance measure is a binary variable at the pupil level (1 = complied; 0 = did not comply). This binary measure was used to estimate the CACE for the RME intervention.

The SAP specified that a two-stage least squares (2SLS) instrumental variable regression approach would be adopted for the CACE analyses (Gerber and Green, 2012). A 2SLS approach assumes that the compliance variable is endogenous such that it will be correlated with the error-term of the final model of the ITT analysis for the primary outcome. An endogenous variable breaks assumptions of ordinary least squares regressions which is why the two stage instrumental variable approach is needed.

The analysis undertaken diverged slightly from what was specified in the SAP by including the Wu-Hausman test[13] to examine the assumption that compliance to RME was an endogenous variable using the Stata 'IVRegress' command. A non-significant Wu-Hausman test would support the assumption that compliance to RME was exogenous and therefore a 2SLS instrumental variable approach is not needed. Here, CACE would be estimated by replacing the group identifier in the final headline ITT analyses of PTM13 with the compliance binary variable using a standard random effects three-level multilevel regression model.

Missing data analysis

In the impact evaluation results section, the amount of missing baseline and outcome data is summarised and possible reasons for missing data are discussed. A multilevel mixed-effect logistic regression model was estimated to check for statistically significant predictors of missing data at the pupil level (KS2 maths attainment centred around class mean), class level (mean class-level KS2 maths attainment centred around the school mean), and school level (mean school-level KS2 attainment centred around the grand mean; use of setting/streaming and geography). Significant predictors and possible mechanisms for the missing data are discussed.

After examining whether it was reasonable to assume that the data was missing at random (MAR), the impact of missing data on the primary analysis was assessed using multiple imputation by chained equations, predicted by the percentage of FSM pupils, the percentage with English as first language, and the percentage with SEN, and statistically significant variables from the model above. A 'burn-in' of ten was used and 20 imputed datasets were created. The primary analysis was then redone using the imputed datasets and Rubin's rules were used to combine the multiple imputed estimates.

Subgroup analyses

Follow-on subgroup analyses focused on the impact of the RME programme on maths attainment for pupils ever classed as FSM.[14] The same three model stages used for the headline ITT analyses were used for these analyses.

Fidelity analyses

As specified in the SAP, to address RQ5 exploratory analyses were undertaken to examine the relationship between fidelity to RME and the GL PTM13 outcome. These fidelity analyses relate to the compliance analyses discussed above but only include RME intervention schools, classes, and ITT pupils. Rather than focusing on identifying the impact of RME compared with the control group, the fidelity analyses focused on the separate dimensions of teacher fidelity (attendance of RME training, RME curriculum time, and RME module use) and pupil-level classroom exposure to RME in both Year 7 and Year 8 and how these were associated with the PTM13 outcome.

---

[13] See **https://www.stata.com/manuals/rivregresspostestimation.pdf**
[14] The NPD EverFSM_6 variable was used.

At the pupil level, fidelity to RME focused on ensuring pupils had consistent exposure to RME teaching in the Year 7 and Year 8 classroom. Using the teacher/class-pupil data collected in Year 7 and Year 8 we were able to identify 43.5% of pupils in the ITT sample who were listed as being in a Year 7 and Year 8 maths class that was taught by one of the RME teachers (either nominated at baseline or a replacement).

The teacher fidelity variables began as scales derived from attendance lists provided by MMU (number of days a teacher attended RME training) or teacher surveys (number of weeks RME was used in the classroom in Year 7 and Year 8; number of RME modules) but lower and higher thresholds for each were agreed with MMU to represent lower and higher fidelity thresholds for teacher fidelity to RME relating to training, curriculum time, and module use. Maintaining the original scales is statistically attractive in terms of constructing a finely grained composite measure of teacher fidelity. However, bringing the three scales together would require a known and/or agreed approach, such as equal or variable weighting of dimensions, which was not feasible for this efficacy trial. The identification of lower and higher thresholds for each of the three teacher-level dimensions of fidelity was agreed and published in the first RME SAP. An updated SAP reported changes to two of these (curriculum time and module use) to account for the reduced classroom time following the partial closure of schools in the final months of the RME programme.

Teacher fidelity data is examined at the teacher level within the IPE section of this report. For these fidelity analyses we report on the impact of teacher fidelity at the pupil level on maths attainment. To bring this teacher-level detail into the impact analyses, it was attached to intervention school ITT pupils at the class level. First, unique teacher IDs were generated for teachers listed as attending RME training or participating in a teacher survey. These teacher names and IDs were then matched with the teacher/class-pupil-level data for the RME ITT intervention group sample. At baseline, the complete RME ITT pupil sample comprised of 4,035 pupils in 169 Year 7 maths classes across 60 RME intervention schools; 111 teacher IDs were attached to 147 of these classes in 56 intervention schools accounting for 3,168 ITT sample pupils (78.5%). The analyses that explore the association between teacher fidelity to RME and pupil maths attainment are therefore restricted to this pupil ITT subsample within RME intervention schools.

One school-level measure of fidelity to RME was requested by MMU—a variable that identified when schools sent two or more teachers to the RME training sessions. Unlike the pupil- and teacher-level fidelity variables, this school-level variable was not selected for inclusion in the compliance analyses discussed above. The variable is included within these fidelity analyses as a scale and then as thresholds by adopting the same two lower/higher attendance thresholds applied for the teacher-level variable.

Table 7 summarises the pupil-, teacher-, and school-level measures of fidelity to RME along with the lower and higher thresholds.

Table 7: Pupil-, teacher-, and school-level fidelity dimensions and thresholds

Component	Data and scale	Lower threshold	Higher threshold
Pupil level (consistent exposure to RME in the classroom in Year 7 and Year 8)	Pupil/teacher class lists collected prior to randomisation Year 8 (pre Covid-19) [1 or 2 instances]	2 instances	2 instances
Teacher level 1 (attendance of RME training)	MMU attendance list [0 to 9]* 7 face to face, 2 online	4+ days	8+ days
Teacher level 2 (RME curriculum time)	From teacher surveys Reported number of weeks that RME is used in the Year 7 and Year 8 classroom [0 to 60]	10+ weeks (reduced from 12+ weeks following Covid-19)	20+ weeks
Teacher level 3 (RME module coverage)	From teacher surveys Reported number of RME modules used in the Year 7 and Year 8 classroom [0 to 10]	6+ modules (reduced from 7+ modules following Covid-19)	All 10 modules
School level (2+ teachers attending same RME training session)	MMU attendance list [0 to 9]* 7 face to face, 2 online	4+ days	8+ days

*The original schedule was for eight face to face training days in each of the six geographical hubs but the final day was cancelled because of Covid-19. MMU developed and delivered two additional online sessions to help support teachers in this period. This resulted in increasing the training scale by one from 0–8 to 0–9. The lower and upper teacher training thresholds remained unaltered for the updated SAP.

A similar three-level multilevel regression approach used for the ITT analyses was adopted for these fidelity analyses, which are confined to the ITT sample of pupils in the 56 RME intervention schools where teacher fidelity data could be attached to the pupil via the maths class. The intervention/control group identifier is therefore redundant and not included in the models. The pupil-, teacher-, and school-level fidelity measures were included into the models separately and then simultaneously. All models also included KS2 maths attainment (at pupil, class, and school levels), geographical hub area, and reported setting or streaming policy.

Estimation of effect sizes

The impact of RME on pupil maths attainment (GL PTM13) was measured using the Hedges g effect size statistic. Hedges g standardises the difference between the attainment of pupils in RME schools and pupils in control schools into units of standard deviations. As specified in the EEF analysis guidance, the unconditional variance was used to obtain the effect size. Specifically, the variance in the GL PTM13 outcome that is clustered at school, class, and pupil levels was used, as set out in equation 1:

Equation 1 $$ES = \frac{(T-C)_{adjusted}}{\sqrt{\delta^2_{sch}+\delta^2_{class}+\delta^2_{pup}}}$$

Where:

- δ^2_{sch} is the school-level variance, δ^2_{class} is the class-level variance, and δ^2_{pup} is the pupil-level variance for the GL PTM13 outcome from the empty/null multilevel model.
- $(T-C)_{adjusted}$ is the mean difference between the attainment of pupils in RME schools and pupils in control schools in the original raw (GL PTM) units. This is obtained from the coefficient for the school-level ‘group’ variable from the final (headline) analyses.

Estimation of ICC

For both pre- and post-test, ICCs at the school and class levels were estimated using a null (empty) three-level multilevel variance components model. Variance decomposition for the three levels (school, class, and pupil) is presented below along with the ICC estimates.

Longitudinal analysis

None planned.

Implementation and process evaluation

Outline and aims

The implementation and process evaluation blended quantitative and qualitative methods (Humphrey et al., 2016) and was informed by both the rationale and description of the RME intervention (see above) and the theory of change logic model.

The aim of the IPE was to provide evidence that supports the interpretation of impact analysis findings, including providing insights into possibly causal explanations. The IPE focused on:

- fidelity (RQ7);
- implementation including variation in this (RQ7, RQ8); and
- teacher perceptions of differences between RME and business as usual (RQ9).

The theory of change model was interrogated by examining effects on teachers' and pupils' activity (RQ10). Finally, implications for scalability and replications were considered (RQ11).

Overview of methods, rationale, and analysis

The IPE was affected by the consequences of the Covid-19 pandemic on participating schools and restrictions on research activities. This section considers:

- overview of methods;
- data collection activities and analysis;
- changes to planned IPE activity; and
- samples.

Overview of methods

Table 8, below, provides a summary of IPE research methods.

Table 8: IPE methods overview

Research methods	Participants/ data sources	Data analysis methods	Research questions addressed	Implementation/logic model relevance
Data collected from schools; school census data.	Data collected from schools and DfE.	Descriptive statistics	RQ7, RQ8, RQ9, RQ10, RQ11	Contextual factors
Online teacher surveys.	Year 7 and Year 8 teachers in intervention and control schools, plus follow up survey. *Year 9 survey added due to Covid-19.*	Descriptive statistics	RQ 7, RQ9, RQ10, RQ11	Causal change 1, professional learning: implementation logic, inputs, outputs, intermediate outcomes.
Implementation survey and MMU attendance data.	Year 7 and Year 8 teachers in intervention schools, MMU.	Descriptive statistics	RQ7	Professional learning implementation logic; pupils' use of RME materials: outputs implementation logic, outputs.
Teacher interviews by telephone.	Year 7 and Year 8 teachers in intervention and control samples. *Year 9 teachers also interviewed as original*	Interview summary, analysed thematically	RQ7, RQ8, RQ9, RQ10, RQ11	Causal change 1, professional learning; causal change 2, pupil experience of RME materials and methods:

	Year 8 teacher interviews delayed due to Covid-19.			professional learning and implementation logic intermediate outcomes, pupil use implementation logic, intermediate outcomes. Contextual factors Complexity.
Visits to RME CPD events; attendance at online CPD	CPD events, observation notes.	Descriptive analysis	RQ8	Professional learning implementation logic: inputs.
Summary data on recruitment approaches; tri-annual updates covering: informal feedback, updates on any attrition, and reasons for withdrawal; project materials.	MMU.	Descriptive analysis	RQ8, RQ11	Contextual factors.

Detail of methods

1. School contextualising data

Data collected from schools prior to randomisation and publicly available data from the DfE schools information service was used to contextualise and interpret findings.

2. Teacher surveys (online)

Online surveys collected data on self-reported teacher practices and beliefs. Questions were devised with consideration of the features of RME pedagogy and the intervention theory of change, that is, related to multiple strategies, mathematising, redefining progress, representations, classroom talk, and use of context. Additionally, more generic beliefs about maths teaching or preference for practices were included as well as efficacy in relation to RME-related practices. Where appropriate, item design drew on existing questions, for example, the OECD TALIS survey (OECD, 2014) and other appropriate instruments (see RME evaluation protocol page 35 for details of the design process).

Teacher surveys were administered electronically using Qualtrics software and were fully GDPR compliant.

Survey 1 in September 2018 focused on teacher beliefs and practices in teaching maths and collected information on any prior use of *Making Sense of Maths* textbooks or of other RME materials.

Survey 2 (summer 2019) was tailored to the two trial conditions in the following ways:

- In RME schools, the teachers originally nominated (or replacements) completed a survey (30 minutes) on their maths teaching including, in relation to RME lessons, issues that might have affected implementation, attitudes to the value of project activities and materials in supporting effective teaching and professional development, and triangulation of data on implementation. Other teachers in the department were invited (but not required) to complete the survey.
- In control schools, the two teachers originally nominated to take part in the project were invited to complete a short survey (20-minutes) on their maths teaching and issues that might affect pupil performance in maths in their school. The aim was to assess the control condition, monitor programme differentiation, and check for spillover.

Survey 3 (summer 2020) was undertaken during partial school closures due to Covid-19 restrictions. The surveys for the two conditions were based on survey 2 but, for the intervention condition, adapted to account for interruption to CPD sessions and classroom teaching of RME. For both conditions, additional data was collected on activity during partial school closures. Ultimately it was decided to exclude the Year 8 survey data from the IPE analysis due to potentially confounding effects of the survey taking place during the 2020 Covid-19 lockdown.

Survey 4 (spring/summer 2021) saw teachers from intervention and control schools surveyed again. This was in addition to the surveys originally planned. This survey focused on maths teaching in Year 9, including during the second period of partial school closure in January to March 2021. This survey was timed to coincide with the period of outcome testing.

Survey data was downloaded and matched using SPSS and Excel software. Descriptive statistics were generated from survey data to compare intervention and control teachers and to identify change in RME intervention teacher responses over time. Relationships to data on fidelity and variation in implementation were analysed.

3. Implementation data

Implementation data was collected in two ways:

(1). Implementation data was collected directly from RME teachers on the use of materials and engagement in other intervention activities. Information was collected through a short survey (5 to 10 minutes to complete) undertaken online via participants' handheld devices with a paper back-up option at or after each of the seven CPD days that took place in person. These short surveys recorded, for example:

a) how far they progressed through a module;

b) to what extent did they use the approaches in the teacher guide; and

c) did they do the gap task.

Surveys were timed to coincide with CPD days as these coincided with when participants would have completed the modules if compliant with the RME protocols cycle.

(2). Implementation data was collected from MMU related to RME CPD attendance.

4. Teacher interviews

These were conducted by telephone in spring 2019 and were due to take place again in spring 2020. Due to Covid-19, the latter were delayed until late autumn 2020 and, following the period of partial school closure, completed in spring 2021.

Intervention interviews

Interviews lasted between 30 and 40 minutes and covered RME implementation, teacher views and beliefs about maths teaching and RME, and pupil responses. In spring 2019, SHU sought to undertake interviews with 12 intervention schools and invited one of the participating teachers from each of these schools to undertake a telephone interview of 30 to 45 minutes (see below, subsection Samples for details of achieved samples and rationale and approach to sampling). During this period, five intervention schools were also selected as sites for case studies with an intention to visit each in spring 2019 and summer 2020 (see RME evaluation protocol for details of case study design and rationale).[15] However, only two schools were visited before Covid-19 restrictions were introduced in 2020. Subsequently, in-person fieldwork was suspended at SHU due to ethical and health and safety considerations. For the purposes of data analysis, the interviews undertaken during the case study visits were treated as part of the spring 2019 teacher interview cycle.

Control school interviews

These lasted around 20 minutes and focused on normal practice and the identification of any spillover issues and programme differentiation. Interviews in autumn 2020 also covered teaching practices that had taken place during partial school closures between March 2020 and September 2020 as well as in-school practices after the schools reopened from September 2020. Spring 2021 interviews covered the same periods as well as the second partial school closure period of January 2021 to March 2021. Fieldworkers completed summary notes (along with some verbatim quotes) recorded and managed in Excel in relation to research questions and sub-questions specific to the interview.

---

[15] https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/realistic-maths-education

5. Field visits to CPD days

Visits to CPD days were planned on four occasions to consider implementation of the professional development component and to understand more fully the intervention and teacher responses. Fieldworkers observed and took notes on CPD activity and on teacher reflections on implementation including engagement in gap tasks.

Covid-19 restrictions meant that two of the four CPD field observations took place online.

Each visit was summarised on a visit record and initial analytical notes were made in relation to observation and relevant RQs, as well other issues of note.

6. MMU data summaries and project materials

MMU provided the following data:

- summary data on recruitment approaches;
- tri-annual updates to include reflections on CPD events, a summary of online activity, informal feedback, updates on any attrition and reasons for withdrawal, and updates from design schools in relation to design research themes; and
- project materials including module materials, CPD materials, and gap task guidance.

Data from these summaries was analysed principally to address RQ11 but also to inform the design of instruments and IPE data collection and to refine the project description in the final report.

Samples

Table 9 gives details of the samples of teachers who participated in IPE data collection.

Table 9: Teachers participating in IPE data collection

	Intervention N	Control N
Survey 1: pre-randomisation (September 2018)	94 (120)	94 (118)
Survey 2: summer 2019	68 (139)	46 (130)
Survey 3: summer 2020	67 (139)	44 (130)
Survey 4: summer 2021	40 (75)	32 (103)
Teacher interviews spring 2019	7+2 (case study)	7
Teacher interviews: between autumn 2020 and spring 2021	12	7

Source: SHU administrative data. Number of teacher responses (number of teachers sent survey in brackets).

Sampling teachers for telephone interviews was done randomly in 2019 with stratification by hub, choosing from teachers originally nominated to participate in RME or a nominated substitute where necessary. If a school declined, a reserve school was identified that preserved the matrix criteria. If a teacher declined or was unable to participate then the other teacher was asked and if both teachers declined or were unable to participate then an alternative school was identified. In the event, recruitment continued until samples were achieved regardless of school or teacher characteristics. The five schools MMU designated as 'design schools' (see above) were excluded from the sample to reduce burden. However, due to administrative error one design school was included in one round of interviews.

One intervention school participated in both the spring 2019 interviews and 2020/2021 interviews and a teacher from one school interviewed originally as a planned case study in 2019 was re-interviewed in the 2020/2021 interview rounds.

In reporting interview data below, the timing of interviews is given as the school year—either 2019/2020 or 2020/2021. This is to reflect the suspension of interviews during the 2021 partial school closure.

Process for designing instruments

As noted above, the description of the intervention and programme theory of change was informed by an IDEA workshop (Humphrey et al., 2016). In addition, specific instruments were designed with input from the MMU RME team. This includes teacher surveys, interview schedules, a pupil survey, and observation rubrics.

The following process for instrument design was followed. The SHU evaluation team considered issues identified by MMU through updates on implementation. Drafts of instruments were circulated for a round of substantive comments, these were taken into account and, where appropriate, instruments were revised.

Covid-19 changes to implementation and process evaluation

Collection of implementation data. MMU provided additional data on participation in online training and in Year 9 CPD. The final teacher survey collected data on use of Year 9 materials.

Teacher interviews. The planned spring 2020 interviews took place in autumn 2020 and spring 2021. Resource for case study visits was reallocated in spring 2021 to increase the number of teacher interviews from seven to 12.

Case study visits. To understand change mechanisms a series of in-depth case studies were planned (see trial protocol).[16] However, due to Covid-19, these case studies were curtailed with initial visits to two schools prior to Covid-19 restrictions and a follow up interview with one teacher from one of these schools in autumn 2021.

Pupil survey. In both intervention and control schools a pupil survey was intended to be administered to all pupils on attitudes to maths and experience of maths teaching in summer 2020. Due to Covid-19, this did not take place, with the cost being reallocated to undertake a fourth teacher survey in summer 2021.

Teacher survey. An additional survey covering maths teaching in Year 9 was administered in summer 2021 (delayed from spring 2021 to coincide with outcome testing).

Field visits to CPD days. These were a mix of field visits (having taken place prior to Covid-19 restrictions) and online attendance with one summer CPD online event observed and one Year 9 online CPD event.

MMU data summaries and project material. MMU provided SHU with the additional materials developed due to changes in the project.

Costs

Data on the cost of CPD and materials was collected from the MMU team. MMU was asked to disaggregate any costs for development of materials or internal evaluation.

The intervention was for two years but costs over a three-year period are calculated to allow for comparison with other trials.

[16] https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/realistic-maths-education/

Planned timeline and changes

Table 10: Timeline and changes

Dates	Activity	Staff responsible/ leading
Jan 2018	Recruitment began	MMU
Sep 2018	Baseline teacher survey and Year 7 pupil/class lists	Schools/MMU/SHU
5 Oct 2018	Randomisation	SHU
Oct–Dec 2018	RME CPD Day 1 and 2 RME teaching: module 1, 'number' CPD day survey 1	MMU/teachers/SHU
Jan–Jul 2019	RME teaching: modules 2–5 RME CPD Day 3, 4, and 5 CPD day survey 2, 3, and 4 MMU visits to design schools Teacher interviews intervention and control End Year 7 teacher survey for all teachers	RME teachers
Sep 2019–Dec 2019	CPD catch-up day for Year 8 teachers new to project RME teaching (Year 8): modules 6–8 RME CPD day 6 CPD day survey 5 MMU visits to design schools	SHU MMU/new teachers RME teachers
Jan–Apr 2020	RME teaching (Year 8): modules 9–10 SHU visit 1 to case study schools SHU teacher telephone interviews MMU visits to design schools RME CPD day 7 (low attendance due to pandemic) CPD day survey 6 (focus on modules 8 and 9)	RME teachers SHU/teachers MMU/schools
Apr–May 2020 (planned activities did not take place).	RME CPD day 8 CPD day survey 7 (focus on module 10, 'algebra') Year 8 pupil/class data collection MMU visits to design schools First additional online CPD session	MMU/teachers SHU/teachers SHU/schools MMU/schools
Jun–Jul 2020	End Year 8 survey for all teachers SHU visit 2 to case study schools (did not happen) Pupil survey (did not happen) Second additional online CPD session	SHU/schools
Covid-19 extension		
Nov–Dec 2020	Optional 'capstone' module, 'making realistic connections', delivered Optional additional Year 9 teacher training twilights Teacher interviews Year 9 pupil/class data collection	MMU MMU SHU
May–Jul 2021	Outcome test Year 9 pupils (delayed from summer 2020) Additional teacher survey in spring 2021	SHU
Aug–Sep 2021	Analysis and report writing	SHU

Impact evaluation results

Figure 3: Participant flow diagram

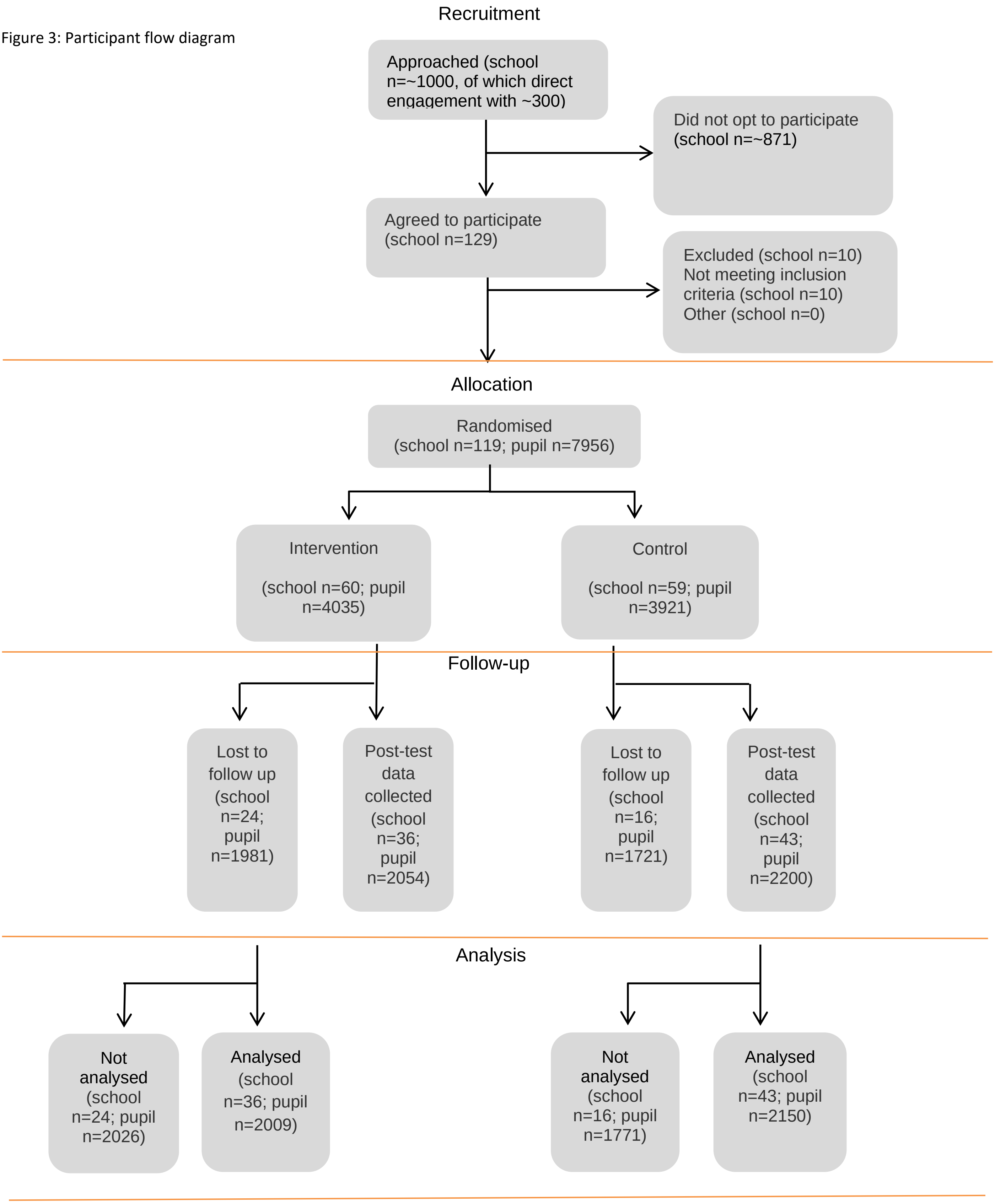


Table 11: Minimum detectable effect size at different stages

		Protocol		Randomisation		Analysis	
		Overall	FSM	Overall	FSM	Overall	FSM
MDES		0.21–0.23	0.22–0.24	0.21–0.23	0.22–0.24	0.15	0.16–0.17
Pre-test/post-test correlations	Level 1 (pupil)	0.70	0.70	0.70	0.70	0.72	0.68
	Level 2 (class)	0.70	0.70	0.70	0.70	0.98	0.95
	Level 3 (school)	0.70	0.70	0.70	0.70	0.00*	0.00**
Intracluster correlations (ICCs)	Level 2 (class)	0.50	0.50	0.50	0.50	0.42	0.42
	Level 3 (school)	0.15–0.20	0.15–0.20	0.15–0.20	0.15–0.20	0.04	0.03
One-sided or two-sided?		Two	Two	Two	Two	Two	Two
Average classes per school		3	3	2.8	2.8	2.7	2.7
Average pupils per class		25	5	24	5	19	6
Number of schools	Intervention	60	60	60	60	36	36
	Control	59	59	59	59	43	41
Number of maths classes	Intervention	180	180	169	169	106	103
	Control	177	177	159	159	110	103
Pupils	Intervention	4,500	900	4,035	840	2,009	515
	Control	4,425	885	3,921	826	2,150	637
	Total	8,925	1,785	7,956	1,666	4,159	1,152

Alpha = 0.05, power = 0.8.
* For the ITT analyses of the GL PTM13 primary outcome, a negative school-level explanatory power of -0.01 was observed for the final precise model. The analysis MDES estimate was undertaken both by setting the school-level explanatory power to zero and then to -0.01. Both resulted in the same estimate of 0.15 SDs.
** For the GL PTM13 analyses for the FSM subsample, a negative school-level explanatory power of -0.15 was observed for the final precise model. The analysis MDES estimate was undertaken both by setting the school-level explanatory power to zero and then to -0.15. The result was MDES estimates between 0.16 and 0.17 SDs.

At the protocol stage, the MDES estimates drew on estimates for school- and class-level clustering and covariate explanatory power. As stated in the SAP, these power analyses were conservative, which accounts for the apparent gain in statistical sensitivity at the analysis stage seen in Table 11 above. The ICC estimates were higher than those observed for the class level (estimated as 0.50, observed as 0.42) and most strikingly for the school level (estimated as between 0.15 and 0.20, observed as 0.04). Estimated covariate explanatory power for the class level ($R^2_{classEst} = 0.70$) was lower than that observed ($R^2_{classObs} = 0.98$). At the school level, estimated covariate explanatory power was much higher ($R^2_{SchEst} = 0.70$) than the negative explanatory power observed ($R^2_{SchObs} = -0.01$). Updating the estimated ICC and explanatory power statistics with what was observed along with the sample size details results in a lower MDES value for both the ITT and FSM subsample analyses than estimated at the protocol stage.

These power analyses corroborate the conclusions of Demack (2019) on class-level clustering and associated class-level explanatory power but should be treated with caution and not as a reliable estimate of the statistical sensitivity of the impact analyses for the RME evaluation. As expected, strong clustering of the GL PTM13 outcome was observed at the classroom level. Demack (2019) indicated that the cause of this strength of class-level clustering related to the extent of setting or streaming practiced across participating schools. Widespread policies of setting or streaming in secondary maths involve segregating pupils into ability groupings based on performance on a maths test, therefore as class-level clustering increases (because of setting/streaming), so will covariate explanatory power of a baseline maths test (KS2 maths score in this instance). These power analyses provide an illustration of the importance of taking class-level clustering and explanatory power into account when designing clustered trials for evaluating programmes focused on secondary maths. However, the estimates do not provide a reliable estimate for the statistical sensitivity of the impact analyses because the sample at analysis stage suffered substantial attrition—48% overall (see below). This means that, at the analysis stage, group membership was determined by something in addition to the baseline randomisation. The MDES estimates at the analysis stage are indicative to illustrate the hypothetical statistical sensitivity of a clustered design with the specified pupil, class, and school sample sizes with the assumption that group membership was randomly determined.

This evaluation also identified an issue for education RCTs that relates to class-level clustering but not directly to the calculation of MDES estimates: pupil and teacher movement between classes during the delivery period. We found that less than half of the ITT sample in RME intervention schools were in a classroom taught by an RME teacher (nominated or replacement) in both years of the programme (43.5%, n = 1,754). At baseline, the 4,035 pupils in the ITT sample in RME intervention schools were located in 169 Year 7 maths classes but by the second year of the evaluation, these pupils were dispersed more widely across 349 Year 8 maths classes.

This is new detail on structural contexts of secondary maths in England but relates to a non-representative sample of 119 schools that participated in the RME evaluation and, therefore, cautious interpretation is advised. However, the realities of pupil segregation and the associated strength of class-level clustering (and hence class-level covariate explanatory power) are a known reality in secondary schools in England (Demack, 2019; Dracup, 2014). It therefore seems reasonable to assume that similar patterns of clustering and explanatory power will be seen in other secondary subjects commonly taught in sets or streams (namely English, science, and modern languages). Further, the introduction and re-configuration of ability groups may go some way to explain the extent of pupil movement we observed for RME and so it seems reasonable to assume similar issues will also be found in other KS3 or KS4 subject areas. While the methodological 'problem' brought by class-level clustering is observed to be mitigated by class-level explanatory power, how the contextual issue of teacher and pupil movement between classes might be mitigated is less clear but does need attention for evaluations of secondary school programmes that have a teacher or classroom focus.

Attrition

Table 12 shows the rate of attrition. In total, 47.7% of pupils at randomisation were not included in the impact analyses. The main reason for this is pupils not taking part in the outcome testing. There are various factors that could explain this. First, as discussed elsewhere in the report, the trial took place in the context of widespread Covid-19 disruption. This led to the extension of the trial, which in turn prompted some schools to end their participation. In middle schools the relevant pupils were no longer enrolled and had moved on to Year 9 elsewhere, whereas other schools were unable or unwilling to continue beyond the two years that were originally planned for the project. While school-level attrition accounted for some of the pupil-level attrition, other issues must also be considered. Covid-19 was still causing high levels of pupil absence when the outcome testing eventually went ahead in summer 2021. It is also expected that a certain number of pupils would move schools during a two-year period even under more usual circumstances; it therefore seems more likely that a greater level of movement would have occurred over what ultimately amounted to a three-year study.

It is also worth noting the extent to which attrition differed between the intervention and control groups: 50% and 45% respectively. These figures relate to attrition at the pupil level. Looking at school-level attrition, the difference is larger, with only 36 of 60 (60%) intervention schools taking part in the outcome testing compared with 43 of 59 control schools (73%). Within schools that supported the outcome assessment, a higher number of pupils took part, which raises the question of why intervention schools were more likely to drop out. Unfortunately, there is no direct evidence on the cause of this disparity. It is possible that the increased burden of additional delivery during the extended intervention period led to some schools declining the opportunity to continue. The higher incentive payments offered to the control group may also have improved retention among those schools. Despite this, the main observation here is that attrition was high primarily due to the pandemic.

Table 12: Pupil-level attrition from the trial (primary outcome)

		Intervention	Control	Total
Number of pupils	Randomised	4,035	3,921	7,956
	Analysed	2,009	2,150	4,159
Pupil attrition (from randomisation to analysis)	Number	2,026	1,771	3,797
	Percentage	50.2%	45.2%	47.7%

Pupil and school characteristics

Table 13 presents pupil- and school-level baseline demographics and attainment information. The number of schools in each geographical area varied according to where recruitment had been most successful, with the largest number of schools located in the North West region, where the developer is also based. Schools in each of the six areas were evenly distributed between the intervention and control groups as this was included as a stratifying variable in the randomisation. The other two stratifying variables, whether a school was secondary or middle and whether a school used setting or streaming, were also evenly balanced between the intervention and control groups. A higher percentage of intervention schools (41.7%) were situated in rural local authority districts than control schools (35.6%).

A higher proportion of intervention schools were rated as outstanding by OFSTED (16.7%) compared to the control group (11.9%), although more control schools were rated as good (40, 67.8%) than intervention schools (32, 53.3%). An element of imbalance is also apparent in the requiring improvement category, with more intervention (12, 20%) than control schools (6, 10.2%) classified as such.

There was good balance in terms of the percentage of pupils defined as disadvantaged (intervention 29.2%, control 30.2%). KS4 (GCSE) maths attainment was also well balanced between intervention schools (mean grade = 4.31) and control schools (mean = 4.26).

Table 13: Baseline characteristics of groups as randomised

School level (categorical)	National-level mean	Intervention group		Control group		
		n/N (missing)	Count (%)	n/N (missing)	Count (%)	
Hub area			60		59	
Yorkshire and Northeast		6/60	6 (10%)	6/59	6 (10.2%)	
North West		17/60	17 (28.3%)	16/59	16 (27.1%)	
West Midlands		10/60	10 (16.7%)	11/59	11 (18.6%)	
East Midlands		6/60	6 (10%)	5/59	5 (8.5%)	
London		11/60	11 (18.3%)	12/59	12 (20.3%)	
South		10/60	10 (16.7%)	9/59	9 (15.3%)	
Sets and/or streams		53/60	53 (88.3%)	53/59	53 (89.8%)	
Middle school		4/60	4 (6.7%)	4/59	4 (6.8%)	
OFSTED *Outstanding*		10/60	10(16.7%)	7/59	7(11.9%)	
Good		32/60	32 (53.3%)	40/59	40 (67.8%)	
Requires Improvement		12/60	12 (20%)	6/59	6 (10.2%)	
Inadequate		2/60	2 (3.3%)	1/59	1 (1.7%)	
Rural		25/60	25 (41.7%)	21/59	21 (35.6%)	
School level (continuous)		n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	
%FSM		59 (1)	29.2 (14.38)	58 (1)	30.2 (15.44)	
KS4 Maths Attainment		51 (9)	4.3 (0.64)	54 (5)	4.3 (0.56)	
Pupil-level (continuous)		n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	Effect size
KS2 maths score		3928/4035 (107)	103.8 (6.96)	3815/3921 (106)	104.2 (6.97)	-0.05

Outcomes and analysis

Primary analysis

Appendix G presents the distributions of the GL PTM13 primary outcome for intervention and control group pupils. From this, two things are apparent. First, the distributions for intervention and control group pupils are similar and, second, the distributions do not follow a bell-shaped Gaussian curve. A sizeable cluster of zeros is observed and the spread between the minimum (0) and maximum (70) is relatively flat. In examining the PTM13 distribution, we were particularly interested in evidence of a ceiling effect but none was observed. The cluster of zeros relate to PTM13 exam papers returned from schools with at least some evidence that the pupil had attempted the test. Checks were conducted manually between receiving the papers from schools and sending onto GL for marking. Additional sensitivity analysis was undertaken for the primary outcome that excluded all pupils with a score of zero. While it seems reasonable to assume that some pupils will have scored zero on the test, it also seems reasonable to assume that some human error might have occurred in the processing of papers that might account for the cluster of zeros. Therefore, we assume that the zero cluster is genuine in our primary analyses but include a sensitivity analysis without the zeros.

Table 14 displays the results for the PTM13 primary outcome analysis and shows that 4,159 of the 7,956 pupils in the ITT sample (52%) were included in the headline model. The impact of RME on maths attainment was observed to be positive but very small (+0.04 SDs, p = 0.44). We therefore conclude from the ITT analysis that no evidence was found that the RME programme had a positive impact on maths attainment.

Table 14: Primary analysis

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
GL PTM Total Raw Score	2009 (2026)	30.8 (30.06; 31.51)	2150 (1771)	30.2 (29.51; 30.94)	4159 (2009; 2150)	+0.04 (-0.06; +0.15)	0.438

Secondary analysis

Follow-on exploratory analyses of GL PTM13 primary outcome—consistent classroom exposure to RME

Of the 4,035 ITT pupils in the 60 intervention schools, 1,754 (43.5%) were observed to be in a maths classroom taught by an RME teacher (nominated at baseline or a replacement) in both Year 7 and Year 8. Taking account of missing outcome and KS2 maths data, of these 1,754 pupils with evidence of consistent RME classroom exposure, 883 (50.3%) were included in the final model. This means that under a quarter of the original ITT sample are known to have experienced RME in both Year 7 and Year 8 and were included in the ITT analyses.

In the 59 control schools, 698 of the 3,921 ITT pupils (17.8%) were observed to be in a maths classroom taught by one of the teachers nominated at baseline in both Year 7 and Year 8.

Reasons for a higher proportion of ITT pupils in intervention schools having Year 7 to Year 8 teacher/class-pupil consistency when compared with ITT pupils in control schools include the use of replacement RME teachers and the encouragement of this consistency in the MOUs.

In response to the difference in observed Year 7 to Year 8 teacher/class-pupil consistency between the intervention and control groups, the multilevel analyses compared the maths attainment of the intervention sample with two control groups. First, the complete 3,921 pupils in the control group with outcome data (n = 2,150; 54.8%) and second, restricting the sample to the 698 control group pupils observed to be in a Year 7 and Year 8 maths class taught by one of the nominated teachers with outcome data (n = 439; 62.9%).

We found no evidence that RME led to a positive impact in pupil maths attainment even when the analyses were restricted to the subsample of pupils known to have been in an RME classroom in both Year 7 and Year 8 (see Table 15). The same finding emerged when the comparison was with the complete ITT pupil control group and when the control group was restricted to control group pupils observed to be in a Year 7 and Year 8 maths class taught by one of the nominated teachers. While the signs differ, both are very small (+0.02 and -0.03) and not statistically significantly different from zero (p = 0.755 and 0.685).

Table 15: Consistent RME classroom exposure in Year 7 and Year 8

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
GL PTM13	883 (3152)	28.6 (27.63; 29.65)	2150 (1771)	30.2 (29.51; 30.94)	3033 (883; 2150)	+0.02 (-0.09; +0.13)	0.755
			439 (3482)	32.1 (30.35; 33.89)	1332 (883; 439)	-0.03 (-0.16; +0.10)	0.685

Follow-on exploratory analyses of GL PTM13 primary outcome—any classroom exposure to RME

In addition to the 4,035 ITT pupils in the 60 RME intervention schools, a further 1,479 pupils outside the ITT sample were observed to be in a Year 8 maths class that was taught by an RME teacher (nominated or a replacement). This combined sample of 5,514 pupils in RME schools are pupils known to have been in a classroom taught by an RME teacher in Year 7, Year 8, or in both of these years. The multilevel analyses compared the maths attainment for this expanded intervention pupil group with the ITT pupil sample in control schools. Results are displayed in Table 16. Once again, no evidence was found that exposure to RME in the classroom led to a positive impact on pupil maths attainment.

Table 16: Secondary analysis—any classroom exposure to RME

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
GL PTM13	2841 (1194)	29.6 (29.0; 30.2)	2200 (1721)	30.0 (29.3; 30.7)	4909 (2759, 2150)	-0.01 (-0.11, +0.09)	0.842

Secondary outcomes (PTM13 subscales)

Exploratory analyses examined the impact of RME using six subscales of the GL PTM13 outcome. The ITT pupil samples in both intervention and control schools were used for these analyses. Results are presented in Table 17.

Across all six subscales we found no evidence that the RME programme led to a positive impact on pupil maths attainment. Effect sizes were all very small, ranging between -0.03 SDs (‘problem solving’ subscale) and +0.06 SDs (‘fluency of concepts’ subscale) with none statistically significantly different from zero.

Table 17: Secondary analysis

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
GLPTM Total excluding Mental Maths Score	2009 (2026)	18.5 (18.00; 19.01)	2150 (1771)	18.3 (17.76; 18.74)	4159 (2009; 2150)	+0.04 (-0.08; +0.15)	0.547
GL PTM selection*	2009 (2026)	11.4 (11.04; 11.65)	2150 (1771)	11.2 (10.90; 11.50)	4159 (2009; 2150)	+0.03 (-0.08; +0.15)	0.557
GLPTM Fluency in Facts and Procedures score	2009 (2026)	3.3 (3.26; 3.42)	2150 (1771)	3.3 (3.22; 3.37)	4159 (2009; 2150)	0.00 (-0.09; +0.08)	0.948
GLPTM Fluency in Concepts score	2009 (2026)	14.7 (14.42; 15.03)	2150 (1771)	14.4 (14.05; 14.66)	4159 (2009; 2150)	+0.06 (-0.04; +0.15)	0.268
GLPTM Problem Solving score	2009 (2026)	1.6 (1.52; 1.70)	2150 (1771)	1.7 (1.60; 1.79)	4159 (2009; 2150)	-0.03 (-0.15; +0.09)	0.616
GLPTM Mathematical Reasoning score	2009 (2026)	11.1 (10.79; 11.43)	2150 (1771)	10.9 (10.57; 11.19)	4159 (2009; 2150)	+0.04 (-0.08; +0.16)	0.476

* MMU selected 11 questions from the PTM13 written test as being aligned with the RME programme and this selection was published in the RME SAP (Appendix IV).

Analysis in the presence of non-compliance

The SAP specified that an instrumental variable approach would be used. The Stata `ivregress` command was used to examine the assumption that compliance to RME was endogenous, using the Wu–Hausman test. If compliance was found to be endogenous, a two-stage, least squares (2SLS) would be required to estimate the Compliers Average Causal Effect (CACE) to avoid breaking a key assumption of standard least squares regression. The Stata `ivregress` command was run twice, first acknowledging clustering at a school level and, second, acknowledging clustering at the classroom level. In both cases, the Wu-Hausman test was not statistically significant.[17] We therefore conclude that there was no evidence that compliance to RME was an endogenous variable and so did not proceed with the 2SLS approach to estimate CACE. Because it is now assumed that compliance to RME was an exogenous variable, CACE is estimated by replacing the group identifier with the complier variable within the multilevel analyses and this is reported below.

The CACE estimate was observed to be positive but very small (+0.03 SDs) and not statistically significantly different from zero (p = 0.77). We therefore conclude that no evidence was found that the RME programme had a positive impact on maths attainment even when pupils were identified as 'compliant'.

[17] When clustering at the school level was acknowledged, the F-test value for the Wu-Hausman was 1.15 with 1,78 df (p = 0.29). When clustering at the class level was acknowledged, the F-test value for the Wu-Hausman was 1.46 with 1,215 df (p = 0.23).

Table 18: CACE estimate

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
GL PTM13	100 (115)	23.0 (20.8; 25.2)	2200 (1721)	30.0 (29.3; 30.7)	2249 (99, 2150)	+0.03 (-0.20, +0.26)	0.769

Missing data analysis

Of the 7,956 pupils in the ITT sample, 3,797 (47.7%) were missing from the analysis of the primary outcome. Of these, 428 were pupils in one of the eight middle schools that had to withdraw from the trial in summer 2020 with the remaining 3,369 of missing data spread across the 111 secondary schools, 79 of which participated in the delayed PTM13 outcome testing. The 79 schools that participated in outcome testing included 5,277 pupils in the ITT sample, 4,254 of whom participated in outcome testing (80.6%). In summary, the majority of missing data comes from the forced withdrawal of pupils in eight middle schools and from pupils in 32 secondary schools that withdrew from the delayed outcome testing.

The missingness in the final analyses related predominantly to missing pupil-level GL PTM13 data (3,702 cases, 46.5%). However, a further 95 cases (1.2%) had outcome data but missing pupil-level KS2 maths attainment. From the multilevel logistic regression model with missing (1) and not missing (0) as the outcome variable, three variables were identified as statistically significant: KS2 maths attainment (lower attainment associated with higher attrition), group membership (pupils in intervention schools more likely to be missing), and geography (pupils in the North West, East Midlands, West Midlands, and South areas were less likely to be missing). Once additional school-level variables were included in the analyses, only KS2 maths attainment remained significant at the 5% level but three additional significant school-level variables were observed: the percentage of FMS pupils (higher %FSM associated with increased chance of missing), the percentage with English as their first language (higher concentrations associated with increased chance of missing), and the percentage with Special Educational Needs (higher %SEN associated with decreased chance of missing). Please see Appendix K for details on these logistic regression models.

To investigate the impact of missing data, the PTM13 analysis was repeated using predicted values obtained via multiple imputation by chained equations (Graham, 2009; Bartlett and Carpenter, 2013). The estimated effect size between intervention and control group pupils following multiple imputation was +0.08 (95% CI: -0.17; +0.34). This was a slightly higher effect size than estimated for the ITT analysis, equivalent to one month's additional progress, but did not reach statistical significance (p = 0.524).

Subgroup analyses

Table 19 shows results from analysis of the FSM subgroup. Multilevel models were estimated for the PTM13 primary outcome and the six PTM subscales.

In addition to the FSM subgroup analysis reported below, an analysis was undertaken for the PTM13 primary outcome that examined whether RME had a differential impact for FSM compared with non-FSM participants. This analysis extended the final ITT model to include the FSM variable and an interaction term FSM*group membership. In these analyses, the interaction between FSM and group membership was not found to be statistically significant (p = 0.368), which led us to conclude we had no evidence to suggest that the impact of RME differed for FSM and non-FSM pupils.

For FSM participants, across the models for the primary outcome and six subscales we found no evidence that the RME programme had an impact on maths attainment. Six of the effect sizes can be considered as very small (ranging between -0.05 SDs for 'fluency in facts' and +0.01 SDs for 'mathematical reasoning') with one observed to be negative and larger (-0.15 SDs for 'problem solving', p = 0.086).

Table 19: Impact analysis of FSM subgroup

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
GL PTM Total Raw Score	515 (726)	24.6 (23.21; 25.96)	637 (668)	24.0 (22.76; 25.27)	1152 (515; 637)	-0.03 (-0.17; +0.11)	0.711
GL PTM total excluding Mental Maths score	515 (726)	14.4 (13.47; 15.31)	637 (668)	14.0 (13.18; 14.83)	1152 (515; 637)	-0.02 (-0.17; +0.13)	0.824
GL PTM selection	515 (726)	9.0 (8.37; 9.53)	637 (668)	8.7 (8.14; 9.17)	1152 (515; 637)	0.00 (-0.15; +0.15)	0.970
GL PTM Fluency in Facts and Procedures score	515 (726)	2.8 (2.61; 2.94)	637 (668)	2.8 (2.63; 2.92)	1152 (515; 637)	-0.05 (-0.17; +0.07)	0.400
GL PTM Fluency in Concepts score	515 (726)	12.2 (11.59; 12.84)	637 (668)	11.8 (11.25; 12.40)	1152 (515; 637)	-0.01 (-0.15; +0.12)	0.829
GL PTM Problem Solving score	515 (726)	1.0 (0.91; 1.17)	637 (668)	1.2 (1.08; 1.36)	1152 (515; 637)	-0.15 (-0.32; +0.02)	0.086
GL PTM Mathematical Reasoning score	515 (726)	8.6 (7.97; 9.14)	637 (668)	8.2 (7.69; 8.71)	1152 (515; 637)	+0.01 (-0.15; +0.16)	0.946

Additional analyses and robustness checks

For the PTM13 primary outcome, we present additional exploratory analyses looking at specific dimensions of fidelity (namely RME training attendance, time spent teaching RME, and use of RME modules) and a combined fidelity indicator. Across the dimensions of teacher fidelity, multilevel fidelity analyses were undertaken using both scale and categorised binary versions of teacher fidelity attached to ITT pupils in intervention schools using the class-level ID. Two sensitivity analyses are also presented: first, a model that excludes pupils scoring zero on the primary outcome assessment and then a model that ignores class-level clustering and adopts a two-level approach, with pupils clustered into schools.

In all cases, the findings from the multilevel models with the scale measures of teacher fidelity did not conflict with those from models including the binary (threshold) measures. We only present fidelity analyses based on the lower thresholds for teacher fidelity to RME in terms of attendance, curriculum time, and module use. Descriptive detail on the scale and upper thresholds are noted below but the multilevel analyses is not. The three teacher-level fidelity thresholds for training, curriculum time, and module use are brought together with one pupil-level threshold of evidence of consistent classroom exposure to RME in Year 7 and Year 8 in the construction of the RME compliers variable (see above).

Please note that the analyses exploring the association between fidelity to RME and maths attainment are distinct from the teacher-level analyses of fidelity to RME discussed below in the IPE section. The teacher-level analyses include all teachers who participated in RME CPD regardless of whether they taught pupils in the ITT sample. The analyses exploring the association between fidelity to RME are based on a subsample of the complete teacher-level fidelity data—a subsample of teachers who could be attached to the ITT pupil sample. At baseline, the complete RME ITT pupil sample comprised of 4,035 pupils in 169 Year 7 maths classes across 60 RME intervention schools; 111 teacher IDs were attached to 147 of these classes in 56 intervention schools accounting for 3,168 ITT sample pupils (78.5%). The analyses that explore the association between teacher fidelity to RME and pupil maths attainment are therefore restricted to this pupil ITT subsample in RME intervention schools.

Attendance at RME training

Attendance registers for the RME training sessions were provided by MMU. These were used to derive a teacher-level scale of attendance between 0 and 9.[18] Additionally, a school-level scale was derived that tallied the number of times two or more teachers were sent to RME training from the same school. Attendance data was obtained for 77 of the 111 RME teachers attached to the ITT sample; the 34 teachers with no attendance data were assumed to have not attended.

Across the 111 RME teachers, the mean attendance was 4.3 days/sessions. The lower threshold for teacher attendance was set at four days: 65 of the 111 RME teachers (59%) attached to the ITT sample met this threshold. They were spread across 93 Year 7 maths classes in 45 RME intervention schools. The upper threshold was set at eight days: 27 of the 111 RME teachers attached to the ITT sample met this threshold (24%) and were spread across 42 Year 7 maths classes in 21 RME intervention schools. The lower and upper thresholds for teacher attendance remained unaltered when the thresholds were re-examined in summer 2020 for the updated SAP.

A school-level variable signifying schools that sent at least two teachers to RME training was requested by MMU but not selected for inclusion in constructing the complier variable (see above). Across the 56 schools where RME attendance data could be attached to the ITT pupil sample, 34 schools were identified as having sent two or more teachers to RME training at least once; 27 schools sent two or more teachers to RME training on at least four occasions and ten schools sent two or more teachers to the full RME training programme (at least eight days).

Table 20 presents the fidelity analyses for RME training attendance lower thresholds. At the teacher level, the unadjusted means show that ITT pupils in a class with a teacher who attended at least four RME training sessions attained higher than ITT pupils in a class with a teacher who attended less than four RME training sessions (unadjusted effect size = +0.20 SDs). However, once KS2 maths attainment, geographical hub, and setting/streaming policy are controlled for in the multilevel analyses, no difference between these two groups of pupils was observed ($g = 0.00$).

At the school level, the unadjusted means show a similar picture with ITT pupils in a school that sent at least two teachers to at least four RME training sessions attaining higher than ITT pupils in a school that did not meet this threshold (unadjusted effect size = +0.25 SDs). Once KS2 maths attainment, geographical hub, and setting/streaming policy are controlled for in the multilevel analyses, this difference was observed to be slightly smaller ($g = +0.17$) but was statistically significant ($p = 0.009$).

[18] Eight RME training days were planned but the final face to face session scheduled in May 2020 was cancelled. Alternative online training was offered to support teachers using RME materials for online or remote teaching. This was offered on a hub basis with two sessions taking place in May and June for each hub, with each session lasting two hours and using a video conferencing platform. This resulted in training attendance covering the seven face to face and two online sessions (nine in all).

Table 20: Fidelity analyses—attendance of RME training (GL PTM13 primary outcome, intervention subsample only)

	Unadjusted means				Effect size		
	Reached threshold		Below threshold				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
RME ITT pupils taught by a teacher who attended 4+ RME training sessions	1120 (831)	32.6 (31.65; 33.51)	889 (1195)	28.5 (27.40; 29.66)	2009 (1120; 889)	0.00 (-0.11; +0.10)	0.973
RME ITT pupils in school that sent 2+ teachers to RME training for 4+ sessions	1043 (810)	32.3 (31.31; 33.37)	966 (1216)	29.1 (28.11; 30.11)	2009 (1043; 966)	+0.17 (+0.04; +0.31)	0.009

RME curriculum time

Teacher surveys were used to collect detail on how many weeks the RME curriculum was followed. Data was obtained for 67 of the 111 RME teachers attached to the ITT sample. Across these 67 teachers, the mean RME curriculum time reported was 11.7 weeks over the two years.

A lower RME curriculum time threshold was set at ten weeks. As noted in the updated SAP, this threshold was reduced from 12 weeks to reflect the reduced curriculum time available in the last five months of RME delivery between March and July 2020. The lower threshold for reported RME curriculum time was met by 41 of the responding 67 RME teachers (61%) spread across 61 Year 7 maths classes in 29 RME intervention schools. An upper threshold was set at 20 weeks. Nine of the responding 67 RME teachers (13%) met this threshold (13%) and were spread across 15 Year 7 maths classes in seven intervention schools.

Table 21 below presents the fidelity analyses for the RME curriculum time lower threshold. At the teacher level, the unadjusted means show that ITT pupils in a class with a teacher who reported using RME for ten or more weeks attained higher than ITT pupils in a class with a teacher who reported a lower RME curriculum time (unadjusted effect size = +0.10 SDs). However, once KS2 maths attainment, geographical hub, and setting/streaming policy are controlled for in the multilevel analyses, the difference is notably reduced (g = +0.01 SDs) and not statistically significant (p = 0.819).

Table 21: Fidelity analyses—RME curriculum time (GL PTM13 primary outcome, intervention subsample only)

	Unadjusted means				Effect size		
	Reached threshold		Below threshold				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
RME ITT pupil taught by a teacher who reported using RME for 10+ weeks in maths classroom	665 (564)	31.9 (30.77; 33.09)	1344 (1462)	30.2 (29.30; 31.14)	2009 (665; 1344)	+0.01 (-0.10; +0.13)	0.819

RME module use

Teacher module surveys were used to collect detail on the number of RME modules teachers reported using in their Year 7 and Year 8 maths classrooms. Data was obtained for 87 of the 111 RME teachers attached to the ITT sample. Across these 87 teachers, a mean of 3.2 RME modules were used over the two years.

A lower RME module-use threshold was set at six modules. As noted in the updated SAP, this threshold was reduced from seven to reflect the reduced curriculum time available in the last five months of RME delivery between March and July 2020; 18 of the 87 responding RME teachers met this threshold (21%). These 18 teachers were spread across 24 Year 7 maths classes in 15 RME intervention schools. The upper threshold was set at all ten RME modules but none of the responding RME teachers met this.

Table 22 presents the fidelity analyses for the RME module-use lower threshold. At the teacher level, the unadjusted means show that the attainment of ITT pupils in classes where teachers reported using materials from at least six RME modules was similar to classes where fewer than six RME modules were used (unadjusted effect size = 0.00 SDs). Once KS2 maths attainment, geographical hub, and setting/streaming policy were controlled for in the multilevel analyses, a very small effect size was observed (g = -0.02, p = 0.819).

Table 22: Fidelity analyses—RME module use (GL PTM13 primary outcome, intervention subsample only)

	Unadjusted means				Effect size		
	Reached threshold		Below threshold				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
RME ITT pupil taught by teacher who reported using materials from 6+ RME modules	279 (195)	30.9 (29.06; 32.66)	1730 (1831)	30.8 (29.99; 31.57)	2009 (279; 1730)	-0.02 (-0.18; +0.14)	0.803

Overall teacher fidelity to RME

A composite measure drew together the three teacher-level RME fidelity lower thresholds: this identified that 14 of the 111 RME teachers met the lower thresholds for training (four days), curriculum time (ten weeks), and module use (six modules). These 14 teachers were spread across 11 RME intervention schools (eight schools where a single teacher met all thresholds; three schools where two teachers met all thresholds). In all but one of these schools, two teachers were commonly sent to the RME training sessions together.[19] These 14 teachers taught 400 ITT pupils in 19 Year 7 maths classrooms across 11 RME intervention schools—9.9% of the original ITT pupil sample.

The 11 schools included two middle schools that had to withdraw from the trial because their Year 8 pupils left for secondary school in summer 2020. It is worth highlighting that these two middle schools were two of only three schools where two teachers were found to meet all three RME fidelity lower thresholds.

In eight of the remaining nine secondary schools, a single teacher met all three RME fidelity thresholds. The one remaining school had two teachers who met all three thresholds. These ten teachers taught 302 ITT sample pupils across 14 Year 7 maths classrooms in nine RME intervention schools—7.5% of the original ITT pupil sample.

Table 23 presents the fidelity analyses for the composite teacher-level fidelity measure based on lower thresholds. At the teacher level, the unadjusted means show that the attainment of ITT pupils in a class with a teacher who met all RME fidelity thresholds was similar to those in a class where the teacher did not meet one or more of the thresholds

[19] In these ten schools, two teachers were sent to training on four or more occasions and in seven of these schools on six or more occasions.

(unadjusted effect size = 0.00 SDs). Once KS2 maths attainment, geographical hub, and setting/streaming policy were controlled for in the multilevel analyses, a very small effect size was observed (g = -0.02, p = 0.819).

Table 23: Fidelity analyses—RME module use (GL PTM13 primary outcome, intervention subsample only)

	Unadjusted means				Effect size		
	Reached threshold		Below threshold				
Outcome	n (missing)	Mean (95% CI)	N (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
RME ITT pupils taught by a teacher who attended 4+ RME training sessions and reported using 6+ RME modules using 10+ weeks of maths curriculum time	239 (161)	28.2 (26.40; 30.02)	1770 (1865)	31.1 (30.36; 31.92)	2009 (239; 1770)	-0.02 (-0.18; +0.14)	0.803

Robustness checks, sensitivity analyses

Two sets of sensitivity analyses were undertaken. First, in response to observing a sizeable cluster of zeros in the GL PTM13 primary outcome, the primary analyses were repeated with this cluster of zero scores excluded; second, as specified in the SAP, a two-level analysis was undertaken that ignored the classroom level.

Table 24 presents the analysis of the PTM13 primary outcome when the cluster of pupils who scored zero were excluded. The effect size remains positive and very small (+0.03 SDs) and not statistically significant (p = 0.576) and so we conclude that these zero scores did not distort the impact analyses findings.

Table 24: Sensitivity analysis for PTM13 primary outcome—excluding cluster of zeros in PTM13

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
GL PTM Total Raw Score excluding zeros	1904 (2131)	32.5 (31.80; 33.17)	2009 (1912)	32.4 (31.67; 33.02)	3913 (1904; 2009)	+0.03 (-0.07; +0.12)	0.576

Appendix I presents a summary of ICC values and covariate explanatory power for two- and three-level impact analyses for the RME ITT sample. These estimates were used for the indicative ‘analysis stage’ MDES estimates above.

In terms of the extent of clustering, from the ICC estimates, a two-level analysis had 18.1% of the variation in GL PTM13 at the school level (between schools) with the remaining 81.9% being located within schools. A three-level analysis only had 3.6% of the variation at the school level, 45.1% of the variation at school or classroom levels (therefore an estimated 41.5% at the classroom level) and 51.2% of the variation located within classrooms (between pupils or between ‘bands’ of pupils if within-class attainment banding is used).

In terms of covariate explanatory power, Appendix I examines this for two- and three-level analyses in two ways. First, with the inclusion of KS2 maths at multiple levels and, second, including the raw KS2 maths variable at just the pupil level. School-level explanatory power was notably high within the two-level analysis (R^2_{school} = 0.749) and nearly zero within the three-level analysis (R^2_{school}= - 0.001) where class-level explanatory power was strikingly high (R^2_{school} = 0.953).

This suggests that the pattern observed at the school level for the two-level analyses is indirectly constructed at the classroom level. The same patterns are observed when the KS2 maths covariate is included at multiple levels or just as a pupil-level raw score, however, a slightly greater explanatory power is obtained when the KS2 maths covariate is included at multiple levels (R^2_{Total} = 0.512 compared with R^2_{Total} = 0.508). Assuming clustering relates to segregating pupils into classes based upon their performance on a maths test, class-level explanatory power for prior attainment largely accounts for this, and this seems to be the case if the attainment covariate is included at multiple levels or just at the pupil level. While the gains in explanatory power brought by including an attainment covariate at multiple levels are small, the inclusion of a classroom level in the design is justified by the teacher/class focus of the RME programme. This enabled teacher engagement with RME to be attached to the ITT sample via the classroom and led to a closer scrutiny of teacher/pupil movement between classes during the delivery period of RME.

The ITT analysis of the primary outcome was repeated ignoring the class level, so pupils were clustered into schools in a two-level analysis. This also found a very small positive effect (+0.05 SDs, p = 0.415). Results are displayed in Table 25.

Table 25: Sensitivity analysis for PTM13 primary outcome—two level, pupils clustered into schools

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
GL PTM Total Raw Score	2009 (2026)	30.8 (30.06; 31.51)	2150 (1771)	30.2 (29.51; 30.94)	4159 (2009; 2150)	+0.05 (-0.06; +0.15)	0.415

Implementation and process evaluation results

Organisation of IPE reporting

As previously noted, the trial was disrupted due to the Covid-19 pandemic. The start of the pandemic affected attendance at CPD days. Implementation and process evaluation (IPE) reporting takes account of this by, first, presenting key findings in relation to implementation prior to March 2020 and then, second, considering the period from March 2020 including reporting data collected on arrangements in 2020/2021 when the pupils originally enrolled in the trial when in Year 7 were in Year 9. For simplicity the period before March 2020 is referred to as the ‘pre-Covid-19 period’. In the original trial period, ending July 2020, the planned use of RME materials for summer term 2020 was largely curtailed through school closures for most pupils. The time after the original period is referred to as the ‘trial extension period’ (September to summer 2021).

IPE reporting for the pre-Covid-19 period is structured in relation to the logic model, reporting on

1. professional learning implementation logic—inputs, outputs, and intermediate outcomes; and
2. pupil use of RME—inputs, outputs, and intermediate outcomes.

Departmental impacts are reported and compliance and fidelity issues are summarised. Usual practice prior to March 2020 is then reported both for control schools and for intervention schools prior to engagement in the trial. Following reporting of IPE findings for the original trial period up to March 2020, findings for the period after this are reported for both control and intervention schools.

Professional learning inputs

Findings reported in this section address IPE RQ7, ‘What is the fidelity of participation in CPD and implementation of RME school-based gap tasks, and use of materials by nominated teachers?’

CPD days

Seven of the eight original planned CPD days were undertaken. Based on observations and interviews with teachers and review of materials, engagement by teacher participants with RME materials was central to CPD days. MMU professional development leaders (PDLs) modelled the use of teaching materials and resources. As part of this modelling, different aspects of the RME teaching approach were demonstrated.

In addition, PDLs provided commentary on their approach to using the materials, including the pedagogic choices they made when using the materials, and potential responses by pupils. Naming and explaining key aspects of the RME teaching approach and techniques were important to the commentary. During these periods of modelling, participant teachers took the role of students. Modelling was interspersed with reflection by teachers on the materials. Introduction of new materials through modelling and related activity constituted 60% to 70% of CPD days. The remainder of the time at CPD days was spent reflecting on and discussing the use of previous modules’ teaching materials, reflecting on gap tasks, planning for use of curriculum materials back in school, and on administrative matters. The modifications to the approach to gap tasks meant that generally less time was spent reflecting on these on CPD days later in the programme than at the start.

CPD events took place in regional networks or clusters. The MMU professional development leads reported variability in the level of engagement in different networks.

Survey responses and interviews with teachers reflect a positive experience of CPD days. Teachers who had participated in CPD in 2020 responded to the statement ‘RME days were good professional development’ with 67% strongly agreeing and 31% agreeing (61 respondents, see Table 26).

Table 26: End of Year 8 survey (summer 2020)—response to the statement 'RME days were good professional development'

	N	%
Strongly agree	41	67.2
Agree	19	31.1
Neither agree nor disagree	0	
Disagree	1	1.6
Strongly disagree	0	
Total	61	100.0

Source: SHU teacher survey

RME professional development

The response to training was positive from those teachers interviewed:

> *'We loved it—lots of maths to actually have a go at—how we would go about doing a lesson. Could go really deep into a topic' (intervention teacher 3, interview 2020/2021).*
>
> *'Face to face one was so valuable and helpful—changed the way I teach maths completely … The training that goes alongside the resources is crucial—the hidden maths behind it' (intervention teacher 1, interview 2020/2021).*

Participants particularly praised the modelling of lessons and the opportunity to 'act out' the teaching and learning themselves:

> *'[Trainer] was absolutely amazing—she challenged a lot of what we were doing in those sessions, constantly modelling the RME way of doing things in her questioning … the entire time we were the students and had to think in the RME way' (intervention teacher 9, interview 2020/2021).*

The opportunity to talk to other maths teachers about techniques was also appreciated.

Interviewees who had attended the additional online sessions were generally positive in their feedback, describing sessions as well thought out and organised, however, some noted that they were not able to utilise the training due to the lockdown situations in school. A small number felt the online sessions were not as helpful as face to face due to some repetition of material and the inability to easily talk with others.

Gap task guidance and activities

The original plan was that the first gap task in each year would be a diagnostic assessment task to encourage engagement with pupil thinking. Following this, most gap tasks were intended to consist of activities involving collaborative planning by pairs of teachers in each participating school with lesson observations. This was described in MMU planning documents as 'lesson study'. However, these plans were revised during the trial to give greater flexibility and to encourage at least some in-depth engagement with module materials (see professional learning outputs below).

Table 27: Gap tasks

When	Details
After training day 1	Diagnostic assessment task: • During a lesson, set four diagnostic assessment questions to be done by your students in test conditions. • Collect in and then decide who to interview—individually or in pairs. • Carry out your interview with at least three students/groups of students—audio record or camera on hand of student. • Arrange to look at each other's clips/scripts within school pairings prior to the next training day.
After training day 2	Collaboratively plan, with your colleague, one of the lessons from module 2; observe each other deliver the lesson and focus on 'noticing' the teacher strategies used and planned for.
After training day 3	Annotate the Teacher Guide as you teach module 3, 'Filling the Whole'.
After training day 4	Task 1: To annotate the teacher guide for 'Knowing the Unknown' and 'Sorting it out'. Task 2: When marking exam scripts (any year group) find examples of questions where an RME strategy could help.
After training day 5	Annotate the Teacher Guides before and after each lesson delivered. Make a two minute video or audio clip of you teaching to share in a trio.
After training day 6	Annotate the Teacher Guides before and after each lesson delivered. Produce **four sets of four questions** on your allocated module that you have tried and tested with pupils. Bring these, together with examples of students' work, to the next training day.
After training day 7	Annotate the Teacher Guide before and after each lesson delivered. Collect data on the retrieval question trial, specifically focused on three students, to bring to day 8.
Online additional training gap task 1	In school pairs, choose one year of exam papers for the exam board that you use (start with Foundation) and make a worksheet which consists of all the questions on those three papers which you think could be solved either by drawing a bar model or a ratio table. Keep the question number and the available marks with the question. Please email these 'worksheets' to your trainer by Wednesday 10th June so that we can share this resource as revision material for your students and beyond.
Online additional training gap task 2	Answer the five GCSE questions using either a ratio table or a bar model.

Professional learning outputs

In this section, data on professional learning outputs are reported (RQ7, RQ8).

CPD attendance

Table 28 provides summary data on CPD attendance by days.

Table 28: Attendance at RME teacher training

	Attendance	N	%
Year 7 2018/2019	Day 1	88	85.4%
	Day 2	81	78.6%
	Day 3	66	64.1%
	Day 4	79	76.7%
	Day 5	74	71.8%
Year 8 2019/2020	Day 6	61	59.2%
	Day 7	52	50.5%
	Online training May 2020	55	53.4%
	Online training July 2020	41	39.8%
	Total	103	

Source: MMU monitoring information

Further data is reported in the Outcomes and Analysis—Additional Analyses and Robustness Checks section, above, such as patterns of attendance in terms of number of teachers attending from schools, the extent to which two teachers attended, and continuation of teachers from year to year. Reporting data on CPD is complicated by changes in teachers involved in the RME trial from Year 7 to Year 8 and by schools withdrawing from the trial. Day 7 attendance was affected by the start of the Covid-19 pandemic with some events taking place in the weeks prior to the first national lockdown. Regardless, an overall pattern of falling attendance, lower in Year 8 than in Year 7, is apparent.

Table 29 reports data on the number of training sessions attended including the online CPD sessions. As described above, the original intervention design was for the same teachers to teach classes in Year 7 and then in Year 8. As only 32 teachers attended eight or more CPD sessions, this intention was not realised in practice. As noted in reporting impact analyses, the middle schools involved in the trial were more compliant with the original design, thus only in approximately a quarter of secondary schools did teachers continue to participate in CPD over two years.

Table 29: Number of RME training sessions attended

N training sessions attended	Frequency	%
No sessions	2	1.9
1–3 sessions	20	19.4
4–7 sessions	49	47.6
8+ sessions	32	31.1
Total	103	100.0

Source: MMU monitoring information.

In both the 2019 and 2020 surveys, teachers were asked for reasons for non-attendance at CPD days through a multiple response questionnaire item. Responses are shown in Table 30. The most frequently occurring reason in the 2019 (end Year 7) survey was that a colleague attended in their place, and in the 2020 (end Year 8) survey it was that cover could not be found on that day. Respondents were also able to describe their reasons using free text comments and explanations included joining the programme midway through the year, being busy with an induction for the incoming Year 7 cohort, and personal reasons. As Table 30 indicates, in spite of the Memorandum of Understanding with the school, some teachers experienced decisions by more senior staff that impeded full participation.

Table 30: Reasons for missing RME training

Why missed RME training	2019 teacher survey	2020 teacher survey
A colleague attended as a substitute when I could not come	12 (33.3%)	2 (10.0%)
Senior leadership would not release me	4 (11.1%)	2 (10.0%)
Department leadership would not let me out of school that day	0 (0.0%)	3 (15.0%)
My head of department/faculty lead decided only one person could attend	0 (0.0%)	1 (5.0%)
Senior leaders decided only one person could attend	8 (22.2%)	2 (10.0%)
Cover could not be found for me that day	5 (13.9%)	6 (30.0%)
I was teaching a Y11 class that day and this had to take priority	0 (0.0%)	3 (15.0%)
Too many staff were already out that day	6 (16.7%)	1 (5.0%)
I had a parents' evening that day	1 (2.8%)	4 (20.0%)
I was ill and so not working that day	1 (2.8%)	0 (0.0%)
Other	13 (36.1%)	1 (5.0%)
N	36	20

Source: SHU teacher survey

Completion of CPD gap tasks

Table 31 provides data on the number of teachers who fully completed gap tasks. The percentage row is the percentage of teachers who completed one or more CPD gap tasks using the 103 teachers listed in the MMU monitoring information as the denominator. Only 67 teachers responded to this question in the survey undertaken at the end of each CPD day but the percentages presented use the higher figure as this is likely to be a more accurate representation of compliance. As two teachers per school began the trial, for the first four gap tasks the number potentially completing the gap task was never higher than 28% of teachers in the trial.

However, in the interviews conducted during 2018/2019, seven out of nine completed or partially completed the gap tasks (a possible sample bias); there is, thus, a lack of data on reasons for non-completion. From the data available, and supplemented by reports from MMU, teachers found it challenging to arrange peer observation. Staffing and timetabling also impacted the completion of the gap tasks. Where peer observation did take place, usually only one teacher would observe the other rather than reciprocal observation of each other. Factors that appeared to support completion of peer observation were that this was part of existing practice within schools and timetables that allowed for observation without additional cover. When peer observation did take place, it was a positive experience for teachers. Later gap tasks did not require peer observation.

Table 31: Teacher completion of RME CPD gap tasks

Gap task number	1	2	3	4	5	6	7	8	9	Total
N	22	29	20	19	6	13	10	17	8	103
%	21.4%	28.2%	19.4%	18.4%	5.8%	12.6%	14.9%	16.5%	7.8%	

Source: SHU teacher CPD module survey

Curriculum materials use

As described above, modules in each year were:

- number;
- geometry;
- proportional reasoning;
- data; and
- algebra.

In 2018/2019, all five Year 7 modules were taught following CPD days. In Year 8, the first three Year 8 modules were taught but the teaching of data and algebra modules was disrupted by Covid-19. Details on number of modules taught and so the pupil experience of RME curriculum materials is reported in detail below (see Table 44).

Professional learning intermediate outcomes

Intended intermediate professional learning outcomes were changes in knowledge, beliefs, and practices as well as both teaching and collegial values. Findings in this section address RQ10 related to how teachers changed due to participation in the intervention.

Data was collected about professional learning outcomes by survey and through interview.[20] As described in the IPE Methods section above, for the purposes of assessing changes in beliefs and practices from survey data, Year 8 survey data was excluded due to potentially confounding effects of the survey taking place during lockdown. This means that changes in beliefs and practices are surveyed after one year only and the potential full effect of the two-year programme is unknown. However, given changes in teachers participating from year to year, assessing the effect of one year of participation in CPD and using RME materials is useful in its own right. For the Year 7 teacher survey, participant data was matched to baseline responses, so the sample comprises all 85 respondents (control n = 40, intervention n = 45) that took part in both the pre-randomisation and the end of Year 7 surveys.

Teacher confidence

Table 32 presents responses to a set of survey questions asking about teacher confidence in different aspects of maths teaching identified in dialogue with the MMU RME team as relating to RME teaching practice. Responses were given on a scale of 1 (very confident) to 5 (very unconfident). Data presented here needs to be treated with caution given potential sample bias as data from respondents was analysed only where they completed both the pre-trial and end of Year 7 survey. Overall, there is little difference between the change over time observed in the intervention and control groups.

The largest differences at follow up emerge on the two items that explicitly mention context. First, for the statement, 'I can link maths to contexts that are meaningful to my students', the intervention group has a mean response of 1.72 compared to 2.23 in the control group, yet a similar gap existed at baseline (intervention mean = 2.23, control mean = 2.50), so the main finding is that both groups gained confidence on this over the first year of the programme. This is in contrast to the statement, 'I can start a lesson from a contextualised problem rather than on the basis of lesson

[20] Due to the need in the Year 8 survey to collect data on lockdown teaching, survey data on any changes in collegial values was not collected.

objectives.' The mean control group response fell from 2.43 to 2.33, indicating a very small increase in confidence, yet for the intervention group the change was from 2.39 to 1.63, representing a far greater gain. Other items on which positive change was associated with the intervention were, 'I can meet the needs of all individual students in my maths class' and 'I can use multiple models or representations to teach an idea or skill', although the gains were more modest.

Table 32: Teacher confidence in aspects of maths teaching

	Control		Intervention		Control-RME difference	
Statement (1, 'very confident' to 5, 'very unconfident')	Base line	End Year 7	Baseline	End Year 7	Base line	End Year 7
I can meet the needs of all students no matter what their prior attainment	2.18	2.13	2.23	2.12	-0.05	0.01
I can meet the needs of all individual students in my maths class	2.10	2.13	2.36	2.02	-0.26	0.10
I can build on students' contributions to navigate to my goals	1.78	1.64	1.93	1.72	-0.16	-0.08
I can get my students to understand underlying concepts in maths	2.03	1.87	2.05	1.81	-0.02	0.06
I can manage whole class discussion over an extended period of time	2.28	2.18	2.05	2.00	0.23	0.18
I can deal with the variety of ideas that students come up with	1.68	1.64	1.68	1.70	-0.01	-0.06
I can link maths to contexts that are meaningful to my students	2.50	2.23	2.07	1.72	0.43	0.51
I can get students to engage in mathematical discussion	2.05	1.97	2.05	1.79	0.00	0.18
I can use multiple models or representations to teach an idea or skill	2.03	1.92	2.23	1.65	-0.20	0.27
I can get students to explain their mathematical thinking	1.85	1.72	1.93	1.63	-0.08	0.09
I can start a lesson from a contextualised problem rather than on the basis of lesson objectives	2.43	2.33	2.39	1.63	0.04	0.71
Mean	2.10	1.98	2.05	1.74	0.05	0.24
N	40	39	44	43	-	-

Source: SHU teacher pre-randomisation and end of Year 7 surveys

Pedagogical beliefs

Table 33 displays the mean responses to survey questions concerning teacher pedagogical beliefs. These questions were included in surveys conducted at pre-intervention and at the end of Year 7, which marks the conclusion of the only year of the trial which was unaffected by Covid-19 disruption. For these items, lower response values indicate beliefs seen as more closely aligned with RME principles (see RME evaluation protocol page 35 for detail of the design process) with lower values indicating closer alignment ('strongly agree' scored as 1 and 'strongly disagree' as 5).

These results are reported for respondents that answered both the baseline and end of Year 7 surveys. They show that beliefs are reasonably balanced at baseline. By the end of Year 7, on five of the six items presented here the intervention group displays responses that are more consistent with RME values than the control group, with 'discussing, explaining, and thinking about maths is more important than the amount of the curriculum covered' as the only exception. For four of these items, the intervention group also provided responses considered more supportive of the RME ethos at baseline. However, the difference between the intervention and control groups increased for each of these four items at the second data point, by which point the trial had been running for one year. This indicates that the RME professional development or the use of RME materials in practice may have influenced teacher beliefs as it is associated with increased agreement with the ideas underpinning the programme. Furthermore, taking a simple mean of the mean values across each of the six items shows that, overall, intervention teachers showed a greater increase in agreement with RME principles (pre = 2.06, post = 1.82) than control group teachers (pre = 2.18, post = 2.11).

Table33: Questions on beliefs and pedagogy—agreement with statements aligned with RME

	Control		Intervention		Control-RME difference	
Statement (1, 'strongly agree' to 5, 'strongly disagree')	Base line	End Year 7	Baseline	End Year 7	Base line	End Year 7
Discussing, explaining and thinking about maths is more important than the amount of the curriculum covered	2.15	1.90	2.11	2.00	0.04	-0.10
I prefer problems that can be solved in several different ways	1.90	1.85	1.75	1.50	0.15	0.35
Having a go is meaningful progress even if the students' strategy is not productive	1.58	1.64	1.66	1.43	-0.08	0.21
Students need to think aloud to work through their ideas	2.92	2.82	2.64	2.44	0.29	0.38
Contexts should regularly be used at the start of topics to generate a discussion of strategies	2.38	2.33	2.14	1.70	0.24	0.63
Mean of means	2.18	2.11	2.06	1.82	0.13	0.29
N	40	39	44	44	-	-

Source: SHU teacher pre-randomisation and end of Year 7 surveys.

The teacher surveys contained eight additional statements on maths teaching beliefs where the statement was not consistent with RME pedagogy. These were also structured in Likert format. These are displayed in the Table 34. The results are more mixed for this set of items compared to those discussed above.

The first item, 'Teaching secure use of procedures is more important than facilitating classroom discussion', received greater agreement from the control group (mean = 2.93) than the intervention group (mean = 3.19) at baseline, but while the mean response from both groups became more aligned with RME at the end of Year 7, control responses did so to a slightly larger extent. Given the difference is negligible (0.06), there is no evidence that participation in the programme was associated with beliefs on these statements becoming more in line with RME.

For the second statement in this set, 'Complex/open-ended problems are worthwhile for high attaining students but not for those who find maths difficult', the control group was less aligned with RME principles at baseline (mean = 3.35) than the intervention group (mean = 3.72) but by the second survey this difference had all but disappeared (control mean = 3.72, intervention mean = 3.70). A similar pattern can be found in responses to the statement, 'The most important part of teaching maths is explaining ideas and procedures clearly.' Again, there is no evidence here that participation in the programme was associated with beliefs becoming more in line with RME.

At baseline, responses to the statement, 'Drawing diagrams is useful in solving shape and space problems but can confuse pupils at other times' leant marginally more toward disagreement in the control group (mean = 3.48) than in the intervention group (mean = 3.36), but this pattern was reversed at follow up (control mean = 3.74, intervention mean = 3.80) amid mean increases across both groups.

The statement, 'Context is useful to motivate and initiate a topic but should be dropped once students have grasped the concept/skills required' yielded the same mean response (3.95) from control respondents at both baseline and follow up. This stands in contrast to the intervention group, where agreement with the statement increased between the two data points, with mean values falling from 4.05 to 3.52, indicating beliefs becoming more distant from RME principles.

This finding is noteworthy given the importance of context in RME and may reflect an interpretation by some teachers that the process from context to model to abstraction in RME teaching is a necessary, desirable, and linear pathway to be followed in all instances. The MMU RME approach, in contrast, sees the relationship between context, model, and abstraction as more intertwined, with application of abstraction back into context also being, in some instances, appropriate and desirable. This contrasts with observed CPD training where the importance of relating abstract methods back to context throughout mathematical learning was emphasised. There also appears to be a discrepancy between this finding and those on teacher practices.

Agreement with the statement, 'Students should learn basic skills before being asked to solve non-routine mathematical problems' was fairly strong at baseline (control mean = 2.38, intervention mean 2.50). At follow up, agreement increased slightly among control respondents (mean 2.26) but the intervention group showed less agreement (mean = 3.05), which is consistent with participation in the intervention influencing beliefs towards RME principles.

Responses to the statement, 'There is usually a best method for solving a maths problem and my job is to make sure students learn that method' showed only small differences between the intervention group (baseline mean = 3.80, end Year 7 mean = 3.73) and control group (baseline mean = 3.68, end Year 7 mean = 3.74) in both waves of the survey. The high level of disagreement with the statement and the lack of change over time indicate a potential ceiling effect for this item.

Finally, the statement, 'It is important not to confuse students by teaching too many ways of doing the same sort of problem' received very similar levels of agreement from the control group in both surveys (baseline mean = 3.35, end Year 7 mean = 3.31). For the intervention group, beliefs became more aligned with RME over time (baseline mean = 2.98, end Year 7 mean = 3.18).

Table 34: Questions on beliefs and pedagogy—agreement with statements not aligned with RME

	Control		Intervention		Control-RME difference	
Statement (1, 'strongly agree' to 5, 'strongly disagree')	Base line	End Year 7	Baseline	End Year 7	Base line	End Year 7
Teaching secure use of procedures is more important than facilitating classroom discussion	2.93	3.21	3.19	3.41	-0.26	-0.20
Complex/open-ended problems are worthwhile for high attaining students but not for those who find maths difficult	3.35	3.72	3.68	3.70	-0.33	0.01
Drawing diagrams is useful in solving shape and space problems but can confuse pupils at other times	3.48	3.74	3.36	3.80	0.11	-0.05
Context is useful to motivate and initiate a topic, but should be dropped once students have grasped the concept/skills required	3.95	3.95	4.05	3.52	-0.10	0.43
Students should learn basic skills before being asked to solve non-routine mathematical problems	2.38	2.26	2.50	3.05	-0.12	-0.79
The most important part of teaching maths is explaining ideas and procedures clearly	2.40	2.85	2.64	2.80	-0.24	0.05
There is usually a best method for solving a maths problem and my job is to make sure students learn that method	3.68	3.74	3.80	3.73	-0.12	0.02
It is important not to confuse students by teaching too many ways of doing the same sort of problem	3.35	3.31	2.98	3.18	0.37	0.13
Mean of means	3.19	3.35	3.27	3.40	-0.09	-0.05
N	40	39	44	44	-	-

Source: SHU teacher pre-randomisation and end of Year 7 surveys

In summary, for half the items, responses to statements indicated greater agreement with RME principles and approaches, these were:

- 'I prefer problems that can be solved in several different ways';
- 'Having a go is meaningful progress even if the students' strategy is not productive';
- 'Students need to think aloud to work through their ideas';
- 'Contexts should regularly be used at the start of topics to generate a discussion of strategies';
- 'Teaching secure use of procedures is more important than facilitating classroom discussion;'
- 'Students should learn basic skills before being asked to solve non-routine mathematical problems'; and
- 'It is important not to confuse students by teaching too many ways of doing the same sort of problem'.

For one item there appeared to be a movement away from RME principles in comparison with the control group respondents. It is also notable that there were some changes on a number of items for control teachers in an RME direction. A possible explanation for this may relate to other influences on teachers during that year (see Business as Usual section below). A second possibility is the timing of the surveys, given that the pre-randomisation survey took place in September 2018. There may have been changes in teacher beliefs about practice over the year in relation to some items. For example, some teachers may have viewed prioritising learning basic skills early in Year 7 as being important, with their views reflecting different learning priorities at different times in the year.

Practices and pedagogical values

The sections below consider data from responses to a series of survey questions asking about teaching practices. These relate to six aspects of practice (with statements in the tables that follow in this section):

- multiple strategies;
- mathematising;
- representations;
- redefining progress;
- classroom talk; and
- context.

For each of these aspects, respondents were asked to rank four statements according to perceived level of importance. The rationale for using ranked statements was to present more nuanced options so that a single right answer was not obvious and to understand views on preferences for different practices that might all have been agreed as being worthwhile if using a Likert type response. The supposition is that teachers would use a variety of practices and indicate those that are used more often rather than using a Likert response format. The survey prompt was:

> **'**Consider how important these activities are in your teaching. They may be important because they happen often in your lessons or because you think they are important for student learning, or a combination of both these or other reasons.**'**

During the IPE design and preparation period, the MMU delivery team helped to produce this list of statements and provided their ideal rankings to reflect practice consistent with RME pedagogy. Important to interpreting responses is that two statements were formulated to be consistent with, and are encouraged in, RME pedagogy—with one more aligned than the other—and two practices that would be inconsistent and discouraged. In the tables below, statements are ordered in the MMU RME rank, shown in the left-hand column of the table. In the survey, items were presented in different order for each item with the order determined randomly but consistently presented for each respondent and in the same order each time the survey was administered.

Mean ranks for individual statements

Tables 35 to 40 report analysis of mean ranks for the six areas of practice. These were calculated using the MMU ranking to weight the responses by teachers. For the statements ranked by MMU as ‘1’ and ‘2’, which were intended to represent practice aligned with RME, a negative change score indicates movement towards RME approaches. For statements ranked by MMU as ‘3’ and ‘4’, which were intended to represent practice contrary to RME principles, an increase in ranking indicates that practice became more consistent with RME approaches.

Multiple strategies

The first set of statements pertains to the encouragement of multiple strategies for solving mathematical problems. Results are displayed in Table 35. The mean ranks show that, overall, there has been a change towards increasing alignment with RME practices, although the practice ranked by MMU as the second best fit with RME moved from second place in the baseline survey to first (most important) in the end of Year 7 follow up.

Table35: Multiple strategies—mean ranks, intervention baseline and end of Year 7

Practice	MMU Rank	Baseline (2018)	End of Year 7 (2019)	Change
I encourage students to explore alternative methods for solutions	1	1.98	1.80	-0.18
I ask students to decide on their own procedures or methods to solve non-routine problems	2	2.24	1.50	-0.74
I focus on the most efficient way to solve a particular kind of problem	3	2.22	3.20	0.98
I steer students away from non-standard or inefficient methods	4	3.56	3.50	-0.06
N		45	44	

Source: SHU teacher pre-randomisation and end of Year 7 surveys.

Mathematising

The next set of statements presented to survey respondents related to mathematising (see Table 36). There has been a change towards the MMU ranking of importance, although the first and second ranked practices are reversed compared to the MMU rankings. The main finding is that responses became more aligned with ideal RME practice over the observation period.

Table 36: Mathematising—mean ranks, intervention baseline and end of Year 7

Practice	MMU rank	Baseline (2018)	End of Year 7 (2019)	Change
I create opportunities to make connections between different mathematical topics	1	2.58	2.05	-0.53
I use activities that help students understand how and why a procedure works	2	1.84	1.74	-0.10
I keep students focused on the essential knowledge or skills needed to answer maths questions.	3	2.70	3.00	0.30
I ensure that students know helpful rules, mnemonics and memory aids to answer questions	4	2.88	3.21	0.33
N		43	42	

Source: SHU teacher pre-randomisation and end of Year 7 surveys.

Representations

Teachers responding to the surveys were asked to rank a set of four questions on representations. Like ‘mathematising’, the overall change is towards the MMU ranking with the practice ranked first and second in terms of importance reversed when compared to the MMU order. Again, this shows an increase in alignment with RME approaches over the study period. Results are shown in Table 37.

Table 37: Representations—mean ranks, intervention baseline and end of Year 7

Practices	MMU rank	Baseline	End of Year 7	Change
I encourage students’ own representations or models in maths and provide opportunities for sharing them in class	1	2.79	2.12	-0.67
I reinforce the idea that drawing diagrams or sketches is valuable for understanding the maths	2	1.91	1.83	-0.07
I use a variety of mathematical representations or models for the same concept	3	2.33	2.71	0.39
I start topics by teaching how to interpret the type of diagrams or representations which are most frequently used in a particular topic	4	2.98	3.33	0.36
N		43	42	

Source: SHU teacher pre-randomisation and end of Year 7 surveys.

Redefining progress

For the four statements on the theme of ‘redefining progress’, the overall change is towards the MMU ranking and the end of Year 7 ranks follow the MMU sequence and overall are close to those ranks (see Table 38). The statement identified by MMU as being the most representative of RME practice is, ‘I emphasise that knowing how you arrived at a solution is as important as getting the answer right.’ The mean rank assigned to this by respondents of the end Year 7 teacher survey was 1.28; responses to this set of statements demonstrate the closest alignment to RME of all those in the survey.

Table 38: Redefining progress—mean ranks, intervention baseline and end of Year 7

Practices	MMU ranks	Baseline	End of Year 7	Change
I emphasise that knowing how you arrived at a solution is as important as getting the answer right	1	1.60	1.28	-0.32
I encourage the strategies that students choose to use even though they may not be the most efficient or standard ones	2	2.98	2.15	-0.82
I make sure that students' misconceptions/errors are corrected straight away	3	2.08	2.92	0.85
I move the class on once enough students are using the procedure or rule correctly	4	3.35	3.64	0.29
N		40	39	

Source: SHU teacher pre-randomisation and end of Year 7 surveys.

Classroom talk

The penultimate set of statements to be ranked by teachers covered the topic of classroom talk. For these practices, the item ranked fourth by MMU—'When I ask questions in whole class discussion, if a student gives a wrong answer, I make sure that the mistake is corrected before moving on'—moves from ranked second in importance to third. This shows a degree of positive change over the first year of the intervention in aligning more with the underpinning ideas of RME. It is also noteworthy that the statement designed to be the most consistent with RME—'I encourage discussions where students question each other and explain their thinking'—was identified by respondents as the most important at baseline (mean rank = 1.61) and by an even greater proportion of respondents at end Year 7 follow up (mean rank 1.27).

Table 39: Classroom talk—mean ranks, intervention baseline and end of Year 7

Practices	MMU rank	Baseline	End of Year 7	Change
I encourage discussions where students question each other and explain their thinking	1	1.61	1.27	-0.34
I provide opportunities for students to make conjectures about mathematical ideas	2	2.78	2.61	-0.17
I keep students' talk in whole-class discussion on-topic to make sure key teaching points are made	3	3.41	3.09	-0.32
When I ask questions in whole-class discussion, if a student gives a wrong answer, I make sure that the mistake is corrected before moving on	4	2.20	3.02	0.83
N		44	41	

Source: SHU teacher pre-randomisation and end of Year 7 surveys.

Context

Finally, for practices related to use of context (Table 40), there is further evidence of practice becoming more closely matched with RME ideals over the first year of the intervention. The MMU-ranked more important item moves from second place to being identified as the most important.

Table 28: Context—mean ranks, intervention baseline and end of Year 7

Practices	MMU rank	Baseline	End of Year 7	Change
Context is used as a vehicle to develop the maths, and is used to do this throughout the lesson/module	1	2.49	1.63	-0.86
I link maths to contexts that are part of my students' daily lives	2	1.86	1.86	0.00
I introduce contexts as opportunities to apply procedures and knowledge once key ideas are learnt	3	2.60	3.09	0.49
Students mainly get experience of applying maths in context through doing word problems	4	3.06	3.43	0.37
N		35	35	

Source: SHU teacher pre-randomisation and end of Year 7 surveys.

Comparison of changes with the control sample and relative degree of change

That engagement in the RME intervention was a probable cause of the change in reported practices is evidenced by comparison with the control subsample who completed both the baseline and end of Year 7 surveys. Across 24 practice statements, on 20, the change for the intervention sample was greater in the direction of alignment with RME practice than for the control sample. However, as noted above, change is possible between statements ranked as 1 and 2 in the MMU rankings. That is, respondent may have ranked two of the four statements identified as more aligned with RME as 1 and 2 in the baseline and then as 2 and 1 in the end of Year 7 survey. In both cases, change in the combined mean ranks of pairs of statements aligned or not with RME provide an overall measure of the change in alignment. Table 41 reports the sum of mean ranks of pairs of statements ranked 1 and 2 by MMU as ranked by the sample of teachers who responded to both the baseline and 2019 surveys. The table is organised so that aspects of practice appear in order of size of change for the intervention sample. A negative change indicates a change towards alignment (as identified by MMU).

Table 4129: Change in mean ranks of MMU first- and second-ranked practices for control and intervention from baseline to end of Year 7

Aspect of practice	Control (N)	Change Control	Intervention (N)	Change Intervention
Redefining progress	34	0.01	39	-1.14
Multiple strategies	44	-0.25	40	-0.93
Context	29	0.08	35	-0.86
Representation	42	0.11	35	-0.75
Mathematising	42	-0.12	37	-0.63
Classroom talk	37	0.00	44	-0.50

Source: SHU teacher pre-randomisation and end of Year 7 surveys.

The data in the table shows that in all cases the change in rank for the intervention sample was greater than for the control sample by at least half a rank. This suggests that for this sample of teachers, the reported changes in practice were the result of the RME professional development.

The areas of practice where participation appears to have had most effect were in relation to redefining progress, multiple strategies, and the use of context. These findings might inform further development of CPD materials and activities.

Reported changes in practice

Considering interview data, the two main changes to teaching practices discussed by participants were around moving towards context-driven teaching and the notion of the 'teacher as facilitator rather than direct instructor' (intervention, 7). Exploring the latter point first, teachers explained how they had 'stepped back' somewhat, allowing pupils to explore, debate, and come to conclusions without teachers intervening or correcting quickly:

> *'More comfortable steering the conversation and the idea of students debating each other. Culture of trying to convince each other and me not being quick to show them that they are correct' (intervention teacher 4, interview 2020/2021).*

> *'RME has given us that platform to come away from "what's in the book" and how many questions have they done, now we focus on listening to kids and to look at how probing teachers' questions are' (intervention teacher 11 interview 2020/2021).*

One teacher described how this method had felt somewhat uneasy at first:

> *'It put me out of my comfort zone, trusting to hand over the lesson to the students—bit apprehensive at first … Helped my questioning when a student is stuck, instead of me volunteering the method or asking questions to draw things out of them' (intervention teacher 5, interview 2020/2021).*

Teachers felt the RME training had helped them to contextualise their maths teaching:

> *'I have always tried to find context in maths but it was limited in my knowledge of maths—context has become integral to the learning journey instead of a naff hook' (intervention teacher 1, interview 2020/2021).*

> *'Using contexts—really lend themselves to the maths—starting with more informal maths and then moving to the formal stuff' (intervention teacher 7, interview 2020/2021).*

A move towards context-driven teaching was thought to engage pupils in the learning, allowing them to see the real-world utility of the maths upfront:

> *'The context driven teaching—not being afraid to get to the bottom of what I am doing quicker … swapping things around, starting with context—engages pupils much quicker. Starting with context gets rid of the "when will we ever use this?" questions' (intervention teacher 8, interview 2020/2021).*

There were examples of interviewees who identified changes not fully captured by the practice statement. One commented on a change in pace (intervention teacher 4, interview 2019/2020) and the teacher being less at the front of the class and moving around the room more to interact with different students.

Overall, teachers were positive about their changes to practice in maths, with one teacher commenting that RME had 'massively changed' the way they teach (intervention, teacher 11, 2020/2021).

Pupil use of RME materials—inputs

As described above, RME curriculum materials were organised into materials for Year 7 and then Year 8, with the following modules in each year:

- number;
- proportional reasoning;
- data;
- geometry; and

- algebra.

Materials were designed to provide more than was likely to be used in the recommended time period. Teachers reported selecting from materials and welcomed this aspect of the project. Generally, teachers had a positive view of the curriculum materials as appropriate for the pupils and good quality:

> *'Very different using resources that you would never have the time or energy or mental thought to create yourself' (intervention teacher 12, interview 2020/2021).*

Some teachers found the materials less accessible to begin with:

> *'The resources can be quite intimidating at first—there are times when you think, "I'm not quite sure where this is going or the point in it"' (intervention teacher 4, interview 2019/2020).*

However, the same teacher reported becoming more comfortable with the approach and the materials over time.

Pupil use of RME materials—outputs

Pupil use of RME materials is reported as part of the Impact Outcomes and Analysis section on Further Analyses. Key data is summarised here followed by discussion of reasons for RME curriculum time and module use. The figures presented are based on the teacher fidelity data collected through teacher surveys and CPD attendance data provided by MMU. This sample of respondents is distinct from the ITT sample of teachers included in the analysis of pupil attainment underpinning the impact evaluation.

Table 42 shows that in Year 7 approximately 40% of teachers taught the material for at least half of the intended 20 weeks of teaching. In Year 8, although the intervention was designed to have ten modules, in most schools only eight were taught before the Covid-19 pandemic partial school closure in March 2020. This gives a potential amount of curriculum time as 16 weeks if materials were used for the equivalent of two weeks of teaching. Only approximately 30% of teachers taught RME for at least half of the potential intended teaching time.

Table 30: RME fidelity data—weeks spent teaching RME in Year 7

Number of weeks	Frequency	%	Valid %	Cumulative %
0	3	2.3	4.6	4.6
1	1	0.8	1.5	6.2
2	2	1.5	3.1	9.2
4	2	1.5	3.1	12.3
5	1	0.8	1.5	13.8
6	3	2.3	4.6	18.5
7	2	1.5	3.1	21.5
8	5	3.8	7.7	29.2
9	4	3.0	6.2	35.4
10	15	11.3	23.1	58.5
11	5	3.8	7.7	66.2
12	6	4.5	9.2	75.4
13	4	3.0	6.2	81.5
14	2	1.5	3.1	84.6
15	5	3.8	7.7	92.3
16	4	3.0	6.2	98.5
17	1	0.8	1.5	100.0
Total	65	48.9	100.0	
Missing	68	51.1		
Total	133	100.0		

Source: SHU module survey.

Table 31: RME fidelity data—weeks spent teaching RME in Year 8

Year 8 weeks spent teaching RME	Frequency	%	Valid %	Cumulative %
0	4	3.0	6.7	6.7
1	1	0.8	1.7	8.3
2	8	6.0	13.3	21.7
3	2	1.5	3.3	25.0
4	6	4.5	10.0	35.0
5	3	2.3	5.0	40.0
6	14	10.5	23.3	63.3
7	5	3.8	8.3	71.7
8	7	5.3	11.7	83.3
9	5	3.8	8.3	91.7
11	2	1.5	3.3	95.0
12	1	0.8	1.7	96.7
13	1	0.8	1.7	98.3
14	1	0.8	1.7	100.0
Total	60	45.1	100.0	
Missing	73	54.9		
Total	133	100.0		

Source: SHU module survey. Please note that these are different from the responses to the teacher survey, which had the responses 0, 'not taught/less than 1 week'; 1, 'approximately 1 week'; 2, 'approximately 2 weeks'; 3, 'approximately 3 weeks'; and 4, 'at least 4 weeks'.

Table 44: Use of RME modules

Module use	N	%
Less than 6 modules	112	84.2
6+ modules	21	15.8
Total	133	100.0

Source: MMU monitoring information.

A lower RME module-use compliance threshold was set at six modules. As noted in the updated SAP, this threshold was reduced from seven to reflect the reduced curriculum time available in the last five months of RME delivery between March and July 2020. The upper threshold was set at all ten RME modules but none of the responding RME teachers met this threshold.

As previously noted, the use of the Year 8 'data' module was largely curtailed by Covid-19 partial school closures and the Year 8 'algebra' module was not taught.

Interviewees all spoke of incorporating RME teaching for their Year 7 and Year 8 pupils, with many stating that they had utilised all the resources and training they had received on RME:

> *'[I] taught them as much RME as I possibly could. Taught every module almost in its entirety and squeezed in the normal scheme as well' (intervention teacher, 12, interviews 2019/2020).*

One teacher described the teaching in this period as being 'completely immersed' in RME. Another, who was fairly new to the school, found that the RME materials were less easily assimilated into their topic area in terms of ordering and as a result found that some of their pupils in a particular set had found the RME materials to be 'a bit too much for them, the pupils were not ready for them by time I got round to it' (intervention 7). However, as noted previously, the interview sample is biased towards teachers who taught more RME materials and so data is lacking on reasons for not using modules or for lower amounts of teaching time.

Pupils—intermediate outcomes

Due to Covid-19 restrictions, as noted, the planned survey of pupils at the end of Year 8 did not take place. Therefore, data on intermediate outcomes is based on interviews with teachers with the risk of sample bias. The logic model posits two intermediate outcomes: the development of an RME culture and development of RME capabilities.

In the section on background above, an RME classroom culture might be marked by:

- time spent talking about context, which might appear to be non-mathematical;
- time spent generating/discussing various representations of a context;
- systematic provision of spaces in which students are invited to talk about their strategies for solving even apparently straightforward questions;
- teacher questions which are open rather than closed; and
- student willingness to initiate/share/discuss/question/explain strategies.

From discussion with the MMU RME team, RME capabilities are understood as the capacity to engage RME classroom culture and to engage with RME practices (as ranked 1 and 2 by MMU above in the section on practice). Interview data indicates that in at least some teachers' classrooms aspects of RME culture did take place.

> *'Everything is valid, every student feels involved, it's changed from getting the right answer to discussion' (intervention, teacher 3, interviews, 2019/2020).*
>
> *'The RME way of teaching lets the students discover the key ideas and concepts and theorems in maths—the students iron out their own misconceptions—other students challenge each other in the context of a real problem, so when students talk, they talk in the context of the problem, which helps their understanding … made me feel OK about horrible hanging silences and the power of a "why?" question and a blank face' (intervention, teacher 9, 2020/2021.)*

Covid-19 disrupted period—summer 2020

In this section, we report data on the Covid-19 disrupted period during summer 2020 where schools were partially closed. Teachers were surveyed at the end of the summer term.

This was a multiple response question but the design meant that a 'yes' to the first question would preclude a 'yes' to the others. Other responses were used to report more complicated patterns. Results show that RME materials were not used by respondents during the summer term.

Table 4532: Use of RME in summer 2020 with Year 8 pupils

	Selected
Due to school closure for most pupils, I have not used any RME materials in the summer term with pupils	86.7%
I have set materials from the Year 8 data and/or algebra modules as out of school tasks	0
I have used materials from the Year 8 data and/or algebra modules in online remote teaching	0
I have used other RME materials with Year 8 pupils as out of school tasks	3.3%
I have used other RME materials with Year 8 pupils in online remote teaching	3.3%
Selected Choice Other (please specify)	13.3%
N	30

Source: SHU end Year 8 teacher survey.

Teachers were also asked about home learning approaches. This question was asked of both control and intervention teachers. Responses were made on a 'choose any that apply' basis. Differences between intervention and control schools suggested some variation in pupil experience. Live lessons, online learning platforms, and materials in more traditional formats were all used by a higher number of control teachers.

Table 46: Pupil home learning experience spring and summer 2020

	Control	Intervention	Total
Live lessons taught by a teacher from your school using platforms such as Google Classroom, Microsoft Teams, or Zoom or similar	37.9%	30.0%	20
Online teaching sites/platforms such as Hegarty Maths, MathsWatch, Oak Academy, NCETM	93.1%	80.0%	51
Worksheets, textbook exercises or resources set as work for individual study	65.5%	53.3%	35
Other (please explain)	13.8%	16.7%	9
N respondents	29	30	59

Source: SHU end Year 8 teacher survey.

As a follow up question, teachers were asked about home learning resources used. Responses are provided in Table 47. The findings show that a wide variety of resources were used by teachers in both the intervention and control groups. MathsWatch, MyMaths, and White Rose were the most popular choices, with more than 20 of the 59 respondents indicating that they used these. Despite this, the most commonly selected response was school developed or downloaded materials, with 28 of the 59 teachers reporting that such resources had been used. Only one teacher said that RME had been used as a Year 8 home learning resource. Of course, this was not available to teachers from the control group; the fact that it was not used by most intervention school teachers perhaps reflects that other materials have been designed for home learning purposes.

Table 47: Year 8 home learning resources

	Control	Intervention	Total
BBC Bitesize	27.6%	16.7%	13
Hegarty Maths	27.6%	30.0%	17
Maths mastery	6.9%	0.0%	2
MathsWatch	41.4%	40.0%	24
MyMaths	44.8%	30.0%	22
NCETM	10.3%	0.0%	3
Oak Academy	17.2%	10.0%	8
RME	0.0%	3.3%	1
School developed or downloaded worksheets/activities	51.7%	43.3%	28
Usual Year 8 textbook	0.0%	6.7%	2
White Rose	41.4%	26.7%	20
Other (please explain below)	34.5%	26.7%	18
N respondents	29	30	59

Source: SHU end Year 8 teacher survey.

With regard to teacher-reported frequency of contact, similar patterns were reported by intervention and control teachers with the overwhelming majority of respondents reporting that maths teachers had been in daily or weekly contact with Year 8 students to assign work, provide support, or give feedback. However, it is also worth noting that a small number of respondents said that such contact occurred less frequently, in some cases only monthly. This was the case in both intervention and control schools.

Table 33: 'How often are maths teachers in touch with Year 8 students about maths work during changes to school provision? This could be to assign work, provide support, or give feedback'

	Control	Intervention
Daily	48.3%	40.0%
Weekly	44.8%	50.0%
Fortnightly	0.0%	3.3%
Monthly	6.9%	6.7%
Total	29	30

Source: SHU end Year 8 teacher survey.

To supplement survey data, teachers were asked in interviews about their use of RME materials during the period of partial school closures. Only one of the teachers interviewed had used RME materials or approaches during this time. This teacher had used some RME materials 'to keep the ideas alive', setting one RME question a week. Teachers generally reported setting work for pupils online, such as watching a video and completing a worksheet, Hegarty Maths and MathsWatch were utilised by some schools. Limited use of live lessons were said to have been used by some as the school closure period extended. The types of activities used were not aligned with RME principles or the design of RME curriculum materials. Teachers were not always able to gauge the level of engagement from pupils in their maths learning over this period but some stated they felt it was either variable or low for reasons related to home learning environments, learner motivation, and difficulty for teachers to monitor engagement.

As noted, to substitute for the planned summer CPD days, online training sessions were offered instead. Interviewees who attended had different views on their value.

> *'I enjoyed the online sessions—it was good to go into smaller groups and have those discussions virtually—it worked well' (intervention teacher 1, interview 2021/2021).*
>
> *'Online was useful—it was good but not as effective or useful as the one we had in person' (intervention teacher 2, interview 2021/2021).*
>
> *'Remote sessions not a great use of my time, repetitive—nothing new. Online—can see why they are there but stuff we have delivered before so didn't feel that pertinent to myself and my colleagues who have done the training' (intervention teacher 4, interview 2021/2021).*

Departmental impacts

The intervention focused on two teachers per school per year engaging with RME. However, it was anticipated that there might be wider impacts at the departmental level[21]. Interviewees were asked about their use of RME approaches with other (non RME) pupils in the school. The extent to which this was said to be happening varied between schools, with one participant explaining they repeated the RME approaches to teaching with their current Year 8 classes, and others saying they used general strategies and approaches in their maths teaching with other classes such as their foundation groups in Key stage 4, including Y11.

Tables 49 and 50 report data from the end of trial survey on planned use of RME materials with Year 7 and Year 8 pupils in the future. For Year 7, 40% of teachers report integration of materials into schemes of work or plans to do so. For Year 8, 35% of schools are doing so or intending to do so. Further, approximately 20% of respondents reported that RME approaches were used with other year groups. Interview data also indicates that the longer-term integration of RME materials has been impacted by Covid-19, with less time for developmental work and concerns for 'Covid catch up'.

Table 49: Use of Year 7 RME materials

	N responses	% of responses
All or nearly all of the RME Year 7 modules are integrated into the Year 7 scheme of work	10	22.2%
Some of the RME Year 7 modules are integrated into the Year 7 scheme of work	8	17.8%
All or some of the RME Year 7 modules will be integrated into the Year 7 scheme of work in the future	2	4.4%
Individual teachers are using RME Year 7 modules as units of work in their teaching	3	6.7%
Teachers are selecting individual activities from RME Year 7 materials in their teaching	10	22.2%
We are not currently using RME Year 7 materials	12	26.7%
Total	45	100.0%

Source: SHU end Year 8 teacher survey.

[21] For middle schools, the use of 'department' is misleading and such impacts would be better understood as school-level impacts.

Table 50: Use of Year 8 RME materials

	N responses	% of responses
All or nearly all of the RME Year 8 modules are integrated into the Year 8 scheme of work	9	20.0%
Some of the RME Year 8 modules are integrated into the Year 8 scheme of work	5	11.1%
All or some of the RME Year 8 modules will be integrated into the Year 8 scheme of work in the future	2	4.4%
Individual teachers are using RME Year 8 modules as units of work in their teaching	6	13.3%
Teachers are selecting individual activities from RME Year 8 materials in their teaching	10	22.2%
We are not currently using RME Year 8 materials	13	28.9%
Total	45	100.0%

Source: SHU end Year 8 teacher survey.

The following quotes from interviews illustrate impacts on departmental practice:

> *'We have made a move away from rote learning, and that narrative runs through KS3' (intervention teacher 11, interview 2020/2021).*
>
> *'[There is] greater emphasis on questioning, and this has gone across the department and so has influenced the Year 9 experience. My HoD thinks the same: questioning across the department is much better' (intervention teacher 7, interview 2020/2021).*

However, other participants were more isolated and their participation was 'against the grain' of the department professional learning culture rather than with it and challenges were reported in sharing approaches with colleagues. For example:

> *'It's not easy—trying to share practice with colleagues within our very small maths departments' (intervention teacher 4, 2019/2020).*

Compliance and fidelity

Compliance and fidelity were reported in the impact analysis and professional learning outputs. Data was collected on reasons for missing RME training (see Table 30, above). From interviews and discussion with the MMU team, contextual factors were identified that affected participation (RQ8). As would be expected, some of these were of varying relevance to different teachers and in different cases.

School staffing and timetabling issues
School staffing and timetabling issues impacted on teacher release to attend CPD sessions (Table 28) during each year and the continuation of Year 7 teachers participating into the second year of the intervention. This latter issue was also affected by staff changes and in a small number of cases was cited as a reason for not continuing with the intervention. As noted above, staffing and timetabling also impacted the completion of the gap tasks and reduced lesson-study-style collaboration in school.

Departmental cultures
A small number of schools that took part in interviews had existing approaches to professional development and departmental cultures that supported the intervention. Such schools reported continuing to engage in forms of peer observation, which was aided by this already being part of departmental culture. Some schools cascaded materials through the department and in such cases were more likely to then incorporate RME into ongoing schemes of work (see Departmental Impacts above).

School cultures

A small number of interviewees felt that some of the RME materials took longer to go through than the ‘usual’ schemes of work, breaking down some of the learning for a deeper understanding. Although this was not in itself problematic, concerns were raised about taking a slower approach as well as how to evidence pupils’ learning in the context of school expectation:

> *‘There is a deep culture of wanting to progress quite quickly and for students to learn things almost immediately, which does not always happen. Trying to explain this to senior leaders—you need to be really strong-willed in convincing them that this does work but it takes a bit of time. Always this feeling that there is not enough time; you have to be brave and confident, knowing that when they do learn it it will be more secure’ (intervention teacher 9, 2020/2021).*
>
> *‘How do we engage senior management wondering what is written in books? Class discussion doesn’t lead to evidence in books’ (intervention teacher 5, 2020/2021).*

Usual practice

In this section we report on practice prior to the intervention and teacher perceptions of differences between RME practices and those used previously; in so doing we also consider the differences in reports of practices between teachers participating in RME and those in control schools (RQ9).

Overall, practices in control schools and intervention schools appeared similar at baseline. Table 51 has data from the baseline survey on RME-aligned or non-aligned practices, with the calculation of mean rank. Calculating statistical significance for each item is not appropriate given that with 24 items some apparent statistical significance might occur by chance. However, the mean average of the two ranks are identical at 2.5 by definition, which is the mean of the nominal ranking 1 to 4 scale. However, looking at means for each item provides insight into baseline usual practice.

Favoured practices at baseline across both samples were a mix of those aligned or contrary to RME approaches. There is evidence from surveys and interviews that during the period of the RME trial some schools in the control group were engaged with mastery influenced practices. Such practices, as noted in the Introduction, have some similarities to RME, for example, the use of models and representations.

Table 51: Teaching practice at baseline

	Correct rank	Control N	Control mean	Intervention N	Intervention mean
Multiple strategies					
I encourage students to explore alternative methods for solutions	1	93	2.05	92	1.95
I ask students to decide on their own procedures or methods to solve non-routine problems	2	93	2.18	92	2.26
I focus on the most efficient way to solve a particular kind of problem	3	93	2.35	92	2.26
I steer students away from non-standard or inefficient methods	4	93	3.41	92	3.53
Mathematising					
I create opportunities to make connections between different mathematical topics	1	91	2.29	89	2.6
I use activities that help students understand how and why a procedure works	2	91	1.79	89	1.96
I keep students focused on the essential knowledge or skills needed to answer maths questions	3	91	2.71	89	2.67
I ensure that students know helpful rules, mnemonics and memory aids to answer questions	4	91	3.21	89	2.78
Representations					
I encourage students' own representations or models in maths and provide opportunities for sharing them in class	1	91	2.85	85	2.89
I reinforce the idea that drawing diagrams or sketches is valuable for understanding the maths	2	91	2.16	85	1.89
I use a variety of mathematical representations or models for the same concept	3	91	2.18	85	2.36
I start topics by teaching how to interpret the type of diagrams or representations which are most frequently used in a particular topic	4	91	2.81	85	2.85
Redefining progress					
I emphasise that knowing how you arrived at a solution is as important as getting the answer right	1	78	1.51	82	1.66
I encourage the strategies that students choose to use even though they may not be the most efficient or standard ones	2	78	3.08	82	3.11
I make sure that students' misconceptions/errors are corrected straight away	3	78	2	82	1.9
I move the class on once enough students are using the procedure or rule correctly	4	78	3.41	82	3.33
Classroom talk					
I encourage discussions where students question each other and explain their thinking	1	82	1.93	83	1.78
I provide opportunities for students to make conjectures about mathematical ideas	2	82	2.82	83	2.89
I keep students' talk in whole class discussion on topic to make sure key teaching points are made	3	82	3.16	83	3.35
When I ask questions in whole class discussion, if a student gives a wrong answer, I make sure that the mistake is corrected before moving on	4	82	2.1	83	1.98
Context					
Context is used as a vehicle to develop the maths, and is used to do this throughout the lesson/module	1	81	2.53	74	2.54
I link maths to contexts that are part of my students' daily lives	2	81	2.16	74	1.85
I introduce contexts as opportunities to apply procedures and knowledge once key ideas are learnt	3	81	2.38	74	2.65
Students mainly get experience of applying maths in context through doing word problems	4	81	2.93	74	2.96

Source: SHU, teacher pre-randomisation and end of Year 7 surveys.

Changes to maths teaching in control schools

In interviews with teachers at control schools in both 2019 and 2020/2021, interviewees either reported that they had made no changes to their maths teaching since being involved in the trial or changes made were not due to involvement in the trial. Influences on practices were due to engagement in Maths Hub CPD or mastery approaches and training.

All but one school said they had no experience of RME. One school was linked to an intervention school, but also said that some maths teachers in its departments were graduates of Manchester Metropolitan University who had undertaken and previously taught some of the RME units. This meant an inevitable influence on their teaching:

> *'My head of maths (one of the control teachers) is an MMU graduate and so is using some of the resources but we are not experts; it's our being graduates of MMU that has led to our ways of teaching but not in our schemes of work' (control teacher 6, interview 2019/2020).*

Thus, there was evidence of some potential 'spillover' in one school.

Trial extension period—2020/2021

In this section, findings are reported about the trial extension period in 2020/2021. During this time, pupils who began the trial in Year 7 were in Year 9. In this year, pupils returned to school from September. However, in the autumn term all schools closed for most pupils for a two-week period due to Covid-19 restrictions. In the spring term, schools were similarly closed but for a longer period.

We report here on practices in both the intervention and control schools and, from survey data, compare the two samples.

Key points are:

- control and intervention practices were similar: there were no identified patterns of difference in relation to use of materials, teaching approaches during lockdowns, or allocation of teachers;
- in intervention schools, there was no continuation of pattern of contact between RME-trained teachers and pupils who had experienced RME; and
- in intervention schools, there was also no continued use of RME approaches and materials as the pandemic disrupted their use due to online teaching, work set for independent study, and changes to pupil grouping to reduce social mixing and ensure social distancing.

Intervention schools

Mathematics organisation in Year 9 and teaching by RME-trained teachers

The majority of teachers interviewed in 2020/2021 were teaching Year 9 pupils during that year, however, there was more of a mixed picture about whether these Year 9 pupils were 'RME pupils'—those who had been taught by RME teachers while in Year 7 and Year 8. Only two participants said that their school's 'RME pupils' were being taught in Year 9 by teachers who had attended the CPD; for the others, there was more variability, often due to the disruption to classes caused by Covid-19 and the need to 'bubble' pupils, for example.

Mathematics teaching in the autumn term 2020/2021

Practices varied across schools and were, as may be expected, different to usual for many schools. During this period, individual pupils or whole classes were instructed to isolate at home for periods of time due to potential or actual Covid-19 infections, therefore teachers sometimes had to teach both in-person and online maths classes. Some schools had 'bubbled' their pupils, often in their form groups, meaning that some maths sets were no longer operating as usual. Within this context, different approaches were taken to teaching: some schools began GCSE work in Year 9 and had therefore started these topics, others returned to their usual scheme or changed to a different scheme such as White Rose. Some interviewees talked about the need to 'catch up' following the period of school closure, with a couple of teachers talking about gaps emerging or widening between pupils.

There was some limited reported use of RME approaches in schools during this period, for example, one school had used some of the RME starter activities and another teacher reported that:

> *'Teachers who have been exposed to [RME] use those ideas and methods, so [we use it] to a certain extent' (intervention teacher 4, interviews 2019/2020).*

Another teacher described how some of the pupils had emotional and behavioural needs, which meant they were not using RME approaches 'as much as we would like to, as the behaviour level isn't there' (intervention 1).

Mathematics teaching during the second partial school closure, January 2021 to March 2021

Of the intervention school interviewees, nine were interviewed in March 2021 and were therefore able to comment on teaching during this period. Teachers reported being *much more prepared* for remote learning at this time and therefore better able to deliver teaching online, with more live or recorded lessons taking place. Pupil engagement appeared to be higher overall, although one teacher commented that there was 'the same third of a class turning up' to the maths live lessons and so two thirds not attending (intervention, teacher 4, interview 2020/2021).

Control schools

Maths organisation in Year 9

All but one of the control schools had setting by previous attainment in place for their current Year 9s, however a small number of participants described a more recent move away from setting:

> *'The current Year 9 are the first year group where we reduced the amount of setting we do. Used to fully set but decided to take two parallel top sets and a support set of around 15 students and mixed ability sets in each year half in the middle. Didn't affect the top sets but the rest of the year group had most of the change—from middle sets to mixed sets' (control teacher 6, interview 2020/2021).*
>
> *'We have been working on maths and moving away from setting. They were the first year we moved away from setting, had set 1, and the rest were mixed ability … Lack of aspirations in the lower groups had caused behavioural issues previously' (control teacher 2, interview 2020/2021).*

The schemes of work or textbooks used in these schools varied, with participants using White Rose, Collins, Edexcel, and their own scheme including mastery approaches.

Maths teaching for control pupils in Year 9

Around half of the control teachers reported some difficulties in maths teaching on returning to school in September 2020 with others reporting, for example, that 'Covid hasn't interrupted much at all' (control school participant 2). For those who had faced challenges, the need to 'bubble' had meant a move away from the usual sets for maths; this, along with the need to go over previous learning and identify gaps that had emerged meant that the usual Year 9 approach had to be altered:

> *'Would have started GCSE in Year 9 but decided not to do that and to do KS3, working to make sure GCSE building blocks are in place. They are coming in weaker in the basics: [we] use diagnostic teaching in our work to identify the gaps. They have definitely suffered due to lockdown and it has been a challenge to overcome' (control teacher 6, interview 2020/2021).*

Comparing intervention and control school practices during Year 9

Table 52 reports data comparing practices during Year 9 in control and intervention schools. Given the sample sizes firm conclusions are tentative but broadly the approaches in the schools in the two conditions was similar. Tables 53 and 54 also present other survey findings on teaching in Year 9 during 2020/2021.

Table 52: ‘Thinking about Year 9 as a whole, what was your overall approach to grouping pupils compared to previous years?’

	Control	Intervention
We increased setting by prior attainment	3.1%	13.2%
We reduced setting by prior attainment (more mixed attainment classes and/or wider bands)	21.9%	13.2%
We kept grouping the same	50.0%	63.2%
Other (please explain)	25.0%	10.5%
N	32	38

Source: SHU, end Year 9 teacher survey.

Table 53: School’s overall approach to teaching Year 9 when all pupils started back in September 2020

	Control	Intervention
We followed the previous Year 9 maths curriculum as planned	64.5%	62.2%
We followed the previous Year 9 maths curriculum at a slower pace and/or with reduced content	22.6%	13.5%
We aimed to review and consolidate previous learning	6.5%	10.8%
Other	6.5%	13.5%
N	31	37

Source: SHU, end Year 9 teacher survey.

Table 54: Pupil home learning experience January to March 2021

	Control	Intervention
Live lessons taught by a teacher from your school using platforms such as Google Classroom, Microsoft Teams, or Zoom or similar	100.0%	97.4%
Online teaching sites/platforms such as Hegarty Maths, MathsWatch, Oak Academy, NCETM	62.5%	36.8%
Worksheets, textbook exercises or resources set as work for individual study	37.5%	34.2%
Other (please explain)	9.4%	2.6%
N respondents	32	38

Source: SHU, end Year 9 teacher survey.

Table 55: Home learning resources used with Year 9 pupils January to March 2021

	Control	Intervention
BBC Bitesize	6.3%	10.5%
Hegarty Maths	34.4%	36.8%
Maths mastery	9.4%	2.6%
MathsWatch	40.6%	44.7%
MyMaths	28.1%	26.3%
NCETM	9.4%	5.3%
Oak Academy	40.6%	23.7%
RME	0.0%	13.2%
School developed or downloaded worksheets/activities	50.0%	34.2%
Usual Year 9 textbook	3.1%	15.8%
White Rose	31.3%	31.6%
Other (please explain below)	31.3%	31.6%
N respondents	32	38

Source: SHU, end Year 9 teacher survey.

Cost

The costs of a school running RME are described here based upon the delivery model employed in this trial. In these figures the central assumption is that schools nominate two maths teachers to take part in the training. Table 56 presents the staff time that would be required. A total of 48 hours between the two teachers across the two school years is necessary to complete the training. A further 16 hours of working time would be allocated for completing the gap tasks that form part of the CPD programme. Additional time for preparation is also expected.

In terms of monetary costs, once printing, payment for supply teachers, travel expenses, and fees to the delivery team are accounted for, it is estimated that a school would pay £5,800 for the two-year intervention. There are no extra costs after the initial training period. These figures are based on having a minimum of 12 teachers attend in-person training and include two RME trainers along with venue and catering costs, assuming that in-person training could take place as was the intention with this trial. Our per pupil cost calculations are based on the two maths teachers in each school being able to teach 60 pupils per year between them and a new cohort of Year 7 pupils starting to receive the programme each year, making 180 pupils over a three-year period. This results in a per-pupil cost of £32.22. The details provided here should enable interested readers to calculate costs based on different numbers of teachers and pupils.

Maths departments could roll out the RME approaches and materials across their department and not simply put forward two teachers for the training. In case of such a roll-out, maths department meeting and preparation time would have to be provided for the two trained teachers to disseminate the RME approaches and materials (mirroring the training schedule but possibly in more condensed form, say 2.5-hour sessions instead of 5 hours).

Table 56: RME staff time—training and delivery

Item	Type of cost	N staff	N hours per year
Teacher training, first year of programme 3 face to face and 2 online sessions —full day: 5 hours + 1 hour break time	Start up	2	30 (first year)
Teacher training, second year of programme 1 face to face and 2 online sessions—full day: 5 hours + 1 hour break time	Start up	2	18 (second year)
Teacher training gap tasks interleaved between training sessions—4 hours each x 4 in each year	Start up	2	16
Preparation and delivery time This would be part of the regular preparation and teaching workload of the teachers; the RME materials could cover/take up about 40% to 50% of the Years 7 and 8 maths curriculum/time.	Recurring (through delivery period)	2 or more	40% to 50% of allocated maths teaching and prep time per class

Table 57: RME material costs

Item	Type of cost	Cost	Total cost over 3 years	Total cost per pupil per year over 3 years
Printing training manuals, slides, and activity sheets (for prep)	Start up	Colour printing @ £0.10 per page	£200	
Supply costs to cover (two) teachers on RME training	Start up	£200 per teacher per day	2 x 8 x £200 = £3,200	
Payment to MMU for training	Start up	£100 per teacher per day	2 x 8 x £100 = £1,600	
Travel costs for in-person training	Start up	Estimate £100 roundtrip train fare	2 x 4 x £100 = £800	
Total			£5,800	£32.22

Source: data supplied by delivery team.

Table 58: Cumulative costs of RME—assuming delivery over three years

	Year 1	Year 2	Year 3
RME	£2,900	£5,800	£5,800

Conclusion

Table 59: Key conclusions

1.	Pupils in RME schools made, on average, no additional progress in maths compared to those in the control group equivalent as measured by GL PTM. This result has very low security rating.
2.	There was no evidence to suggest that the programme had a differential impact on the GL PTM results of FSM pupils.
3.	Movement of ITT pupils between Year 7 and Year 8 classes was observed, meaning that only 43.5% of ITT pupils in the RME intervention schools received both years of the programme. This finding seems to present a challenge for the future design of RCTs for evaluating programmes in secondary schools that focus on change at a teacher or class level over two years.
4.	Some teachers who engaged in RME professional development were enthusiastic about RME as an approach to teaching maths and the professional development they experienced. There was some evidence that engagement changed self-reported teacher beliefs and practices when compared to teachers in the control condition.
5.	There were indications that the model for professional development of two teachers attending external professional development workshops over two years was challenging to implement in some schools due to ongoing pressures.

Impact evaluation and IPE integration

Evidence to support the logic model

The logic model (Figure 1, above) proposed central causal mechanisms: professional development leading to changed teaching approaches (when using RME curriculum materials and in other lessons) and pupil learning being impacted by changed teaching and directly through the use of RME curriculum materials. Given the disruption of Covid-19 (see below on limitations), conclusions are tentative. However, there is evidence from the IPE that the RME CPD did influence self-reported teacher beliefs and practice for at least some teachers. However, data on beliefs and practices at baseline suggest that in some areas and for some teachers, practices were already at least somewhat aligned with RME approaches. Changes in practice happened despite a planned aspect of CPD—lesson study through gap tasks—having low fidelity initially and later adapted to be more of an individual activity.

Although there were changes in beliefs and reported practices, this did not lead to a detectable change in pupil attainment, although given the pandemic disruption this should be treated as an absence of evidence rather than evidence of absence.

Interpretation

While the impact evaluation found no evidence that teacher participation in the RME programme led to gains in pupil maths attainment, there are some important methodological limitations that served to undermine the validity of the clustered RCT design. These are discussed in more depth below but can be summarised as attrition (48%) and movement of pupils and teachers between maths classes during the evaluation period.

The sizable attrition is linked to the problems brought by Covid-19, partial closure of schools during the final five months of delivery and need to postpone outcome testing by about a year. Of the schools and pupils that did participate in outcome testing, the increased time between end of delivery (summer 2020) and outcome testing (summer 2021) may have led to some dissipation of the impact of RME on maths attainment. As reported above, the exposure to both RME materials and methods during Year 9 was very limited in general and largely non-existent for some.

Impact analyses did not find a relationship between material use, module use, or CPD attendance and pupil attainment. An association was found between schools having two teachers regularly attend and attainment outcomes. This may be a result of existing school cultures and contexts rather than a consequence of attendance.

This evaluation was the first relatively large trial of Realistic Mathematics Education in England. Engagement in RME CPD and using curriculum materials did lead to reported changes in beliefs and teacher practices. The changes in practices arose in relation to the RME aspects of practice:

- multiple strategies;
- mathematising;
- representations;
- redefining progress;
- classroom talk; and
- context.

Based on self-reported data, the RME intervention appears successful in changing practice of some teachers. This is noteworthy and potentially has implications for other maths professional development innovations.

RME has some similarity to other interventions and innovations that have been trialled in England.[22] It also has some overlaps with practices encouraged in current mastery approaches. Several RME-type practices are advised in recent guidance for KS2 and KS3 maths teaching (EEF, 2017; see also Hodgen, Foster, Marks and Brown, 2018) which recommend use of representations, focusing on problem-solving strategies, developing pupil independence, and using tasks and resources to challenge and support maths learning. Thus, the RME trial provides evidence that it is possible to design CPD to change teachers' practice (at least as self-reported) in ways that align with evidence on practices that impact learning. However, fidelity and compliance data suggest that the type of engagement by schools and teachers for changes in practice to be reported was not universal or even for a majority of teachers and schools. Factors influencing engagement and implementation in the RME trial are like those previously reported (see, Boylan, Adams, Coldwell, Willis and Demack, 2018).

Limitations and lessons learned

Limitations

As noted above, there are limitations of this evaluation in terms of methodology and the impact of Covid-19 on delivery of RME. The first key limitation was attrition with nearly half of the ITT sample not participating in outcome testing. This attrition related predominantly to the withdrawal of whole schools from the evaluation following the arrival of Covid-19. First, eight middle schools withdrew because the participating pupils left for secondary school in summer 2020 at the end of Year 8. Second, a further 32 secondary schools withdrew from the delayed outcome testing even when additional incentives to remain in the evaluation were offered. In the 79 schools where outcome testing took place in summer 2021, pupil level attrition was notably lower at less than 20%. In response to the attrition, we used multiple imputation to address the missing data but still found no evidence of impact.

The second limitation was the extent of movement observed for the ITT pupil sample between maths classes in Year 7 and Year 8. We observed that less than half (43.5%) of the pupil ITT sample in RME intervention schools were in a class taught by an RME teacher (nominated prior to randomisation *or* a replacement teacher) in Year 8 for the second year of delivery. This movement of the ITT sample between classes was observed to have taken place prior to the arrival of Covid-19 and so presents a fundamental limitation to the validity of the ITT analyses we present.

A third limitation was that the GL PTM outcome testing was done with no invigilation. This was because of restrictions in physically entering schools during the Covid-19 pandemic. Schools were provided with guidance and test papers to undertake the outcome testing, which took place within Year 9 maths classes. Whilst invigilation would have been a

[22] For example, Dialogic Teaching: **https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/dialogic-teaching** ICCAMs **https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/increasing-competence-and-confidence-in-algebra-and-multiplicative-structur** and the Multiplicative Reasoning Project: **https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/436832/RR406A_-_Multiplicative_Reasoning_Professional_Development_Programme.pdf**

better option, we do not think that the lack of it (because of Covid-19) resulted in limitations above and beyond the two key limitations discussed above.

Disruption of the implementation and process evaluation leads to two further limitations. Data on pupil use of materials and experience of RME was affected by not undertaking the pupil survey due to partial school closures. This reduced evidence about the experience of using curriculum materials and intermediate pupil outcomes. Planned case studies were designed to collect data on potential changes in classroom practice and to triangulate self-reported data on practices. Classroom observational data was limited to two initial visits and so did not inform analysis of changes in practice. There is evidence of sample bias in the interview samples given the reported rates of implementation of RME practices and use of materials is higher than the compliance data collected from a larger sample of teachers.

Lessons learned

We consider lessons learned in relation to three areas in turn: maths interventions, professional development interventions, and general implications for undertaking trials in the current secondary school context in England.

Mathematics interventions and curriculum materials

The disruption of Covid-19 has meant that the security of the reported trial is low and firm conclusions cannot be drawn about the value of RME approaches in terms of pupil outcomes. The trial did, however, lead to a more detailed description of RME as an approach as implemented in England and the development of a model of RME practices. This gives the basis for future study of specific aspects of RME practice. Aspects of practice, and appropriate professional development focused on aspects such as 'redefining progress' and 'use of context', may lend themselves to smaller design research focused on the benefits of particular mathematical models and representations for different maths topics. Separate from the professional development approach, RME curriculum materials could be tested beforehand to ascertain their efficacy for teaching certain topics.

Professional development interventions

The approach to professional development in the RME trial was centred on a series of CPD days with two or more teachers attending per school. This is a pattern that has been used in a variety of maths education interventions including EEF trials (for example, ScratchMaths and the Multiplicative Reasoning Project). Having two teachers attend per school has a number of benefits: it supports collaborative practice, it helps to mitigate against teacher absence, and it supports wider change in schools. It is likely that the participation of two teachers may have been important to achieving wider impacts in those schools where there was incorporation of RME materials into ongoing schemes of work and curriculum. However, overall, the fidelity of two teachers participating in all sessions was lower than desired and with some indications that those attending more fully were more likely to have practices aligned with RME approaches. Further, given movement of pupils between classes, models of CPD focused on individual teachers potentially leads to dilution of effect when pupils change class.

Given this, alternative models may lead to higher fidelity such as cascade or train-the-trainer models where departments or schools commit to those attending CPD leading in-school training. Alternatively, CPD might be based in school from the outset with CPD leads working directly as coaches and trainers of whole departments. However, such forms of professional development are difficult to evaluate through quasi-experimental methods.

In any future trial of RME specifically, greater clarity is advised at the outset about the extent to which collaborative practice by pairs of teachers in ways similar to Lesson Study will be required. One modification made early in the trial related to gap tasks. This reduced the degree to which the programme was clearly specified in advance.

Trial design

As far as we are aware, this is new contextual detail on the shifting compositions of maths classes during a two-year evaluation and we plan to draw on this for future research and publications (outlined below). The use of a class level in our clustered RCT design along with the collection of pupil-teacher/class lists in Year 7 and Year 8 enabled this pupil-level movement to be observed. The extent of pupil movement will be explained by the introduction of setting or streaming or the reconfiguration of sets/streams/classes between Year 7 and Year 8. This issue may be most evident within secondary maths where setting or streaming is nearly universal, but it seems likely that similar patterns would be observed in evaluations of other subjects where pupils are commonly set or streamed.

In terms of RCT design, the issue of pupil (or teacher) movement between classes during delivery is most critical for evaluations of programmes that run in secondary schools with a focus on change at teacher or class levels. Additionally, pupil movement may be greater for programmes that span more than a single school year (as is the case here).

Evaluations of programmes that focus on change at a departmental, key stage, or year group level will be less affected by the movement of pupils between classes. In these cases, pupil movement might result in practical issues such as the need for schools to facilitate testing of pupils drawn from more classes at outcome than was the case at baseline (that is, potential increased school burden). Similarly, evaluations of programmes targeted at specific pupil groups (such as FSM or EAL) would be less affected by pupil movement. However, for future evaluations of programmes in secondary schools that are targeted at teacher- or class-level change, our findings suggest that some thought will be needed to design RCTs that are not undermined by the movement of ITT pupils between classes. At the very least, our findings emphasise the importance of evaluators being able to observe movement if it occurs and this in turn emphasises the value of including a class or teacher level for evaluations of such programmes.

In terms of lessons learnt, as noted above, the issue of pupil movement between classes was present before the arrival of Covid-19 (early Year 8 in 2020). However, we did not observe this until the start of the impact analyses in the following summer, 2021. If analyses of the extent of pupil movement were undertaken earlier, this would have informed the discussions on whether to extend or end the evaluation. The extent of pupil movement was unknown but given what we observed, future evaluations of programmes in secondary schools that focus on change at the teacher or class level are advised to put resource into closely monitoring the class compositions of their ITT samples.

Future research and publications

For future research, we plan to look more closely at class-level clustering, covariate explanatory power, and pupil movement. This would take the form of a methodological paper for dissemination to the designers of educational trials and a policy focused paper around the segregation of pupils in secondary maths in England.

References

Allen, R., Jerrim, J., Parameshwaran, M. and Thompson, D. (2018) 'Properties of Commercial Tests in the EEF Database', EEF Research Paper 1.

Barmby, P., Dickinson, P., Hough, S. and Searle, J. (2011) 'Evaluating the Impact of a Realistic Mathematics Education Project in Secondary Schools', *Proceedings of the British Society for Research into Learning Mathematics* 31 (3), pp. 47–52.

Bartlett, J. and Carpenter, J. (2013) 'Missing Data', Module 14 on the Learning Environment for Multilevel Methods and Applications (LEMMA) online course run at Bristol University. **https://www.cmm.bris.ac.uk/lemma/**

Boaler, J. (2002) *Experiencing School Mathematics: Traditional and Reform Approaches to Teaching*, New Jersey: Lawrence Erlbaum Associates.

Boaler, J. (2002). 'Paying the Price for "Sugar and Spice": Shifting the Analytical Lens in Equity Research', *Mathematical Thinking and Learning,* 4 (2-3), pp. 127–144.

Boaler, J. (2003) 'Studying and Capturing the Complexity of Practice – the Case of the "Dance of Agency"', in Pateman, N., Dougherty, B. and Zilliox, J. (eds, 2003) *Proceedings of the 27th Annual Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 3–16) Honolulu, HI: PME.

Boaler, J. and Wiliam, D. (2001) '"We've still got to learn!": Students' Perspectives on Ability Grouping and Mathematics Achievement', in P. Gates (ed.), *Issues in Mathematics Teaching*, London: Routledge Falmer.

Boaler, J., Wiliam, D. and Brown, M. (2000) 'Students' Experiences of Ability Grouping-Disaffection, Polarisation and the Construction of Failure', *British Educational Research*, 26 (5), pp. 631–648.

Boylan, M. and Demack, S. (2018) 'Innovation, Evaluation Design and Typologies of Professional Learning', *Education Research*, 60 (3), pp. 336–356.

Boylan, M., Demack, S., Willis, B., Stevens, A., Adams, G. and Verrier, D. (2015a) 'Multiplicative Reasoning Professional Development Programme: Report', London: DfE.

Boylan, M., Demack, S., Willis, B., Stevens, A., Adams, G. and Verrier, D. (2015b) 'Multiplicative Reasoning Professional Development Programme: Technical Report', London: DfE.

Boylan, M., Demack, S., Wolstenholme, C., Reidy, J. and Reaney-Wood, S. (forthcoming) 'ScratchMaths Evaluation Report'. The 2017 protocol for this evaluation is here: https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Protocols/Round_6-_Scratch_maths_amended.pdf

Boylan, M. and Jay, T. (2017) 'Independent Evaluation of "Investigating the Impact of Realistic Mathematics Education Approach on Achievement and Attitudes in Post-16 GCSE Mathematics Resit Classes"', in Hough, S., Solomon, Y., Dickinson, P. and Gough, S. (2017) *Investigating the Impact of a Realistic Mathematics Education Approach on Achievement and Attitudes in Post-16 GCSE Resit Classes*, Nuffield.

De Corte, E., Op 't Eynde, P. and Verschaffel, L. (2002) '"Knowing what to believe": The Relevance of Students' Mathematical Beliefs for Mathematics Education', in B. Hofer and P (Eds), Personal epistemology: The psychology of beliefs about knowledge and knowing (p297-320). Lawrence Erlbaum Associates Publishers.

De Lange, J. (1996) 'Using and Applying Mathematics in Education', in A. J. Bishop, et al. (eds) *International Handbook of Mathematics Education, Part One* (pp. 49–97), Kluwer Academic 101.

Demack, S. (2019) 'Does the Classroom Level Matter in the Design of Educational Trials? A Theoretical and Empirical Review', London: Education Endowment Foundation. https://educationendowmentfoundation.org.uk/public/files/Publications/Does_the_classroom_level_matter.pdf

Dickinson, P. and Eade, F. (2005) ‘Trialling Realistic Mathematics Education (RME) in English Secondary Schools’, *Proceedings of the British Society for Research into Learning Mathematics,* 25 (3), pp. 1–13.

Dickinson, P. and Hough, S. (2012) ‘Using Realistic Mathematics Education in UK Classrooms’, Centre for Mathematics Education, Manchester Metropolitan University, Manchester, UK.

Dickinson, P., Eade, F., Gough, S. and Hough, S. (2010) ‘Using Realistic Mathematics Education with Low to Middle Attaining Pupils in Secondary Schools’, in Joubert, M. (ed.) *Proceedings of the British Society for Research into Learning Mathematics*, April 2010.

Dracup, T. (2014) ‘The Politics of Setting’: https://giftedphoenix.wordpress.com/2014/11/12/the-politics-of-setting/

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

EEF (2017) ‘Improving Mathematics in Key Stages Two and Three’, London: Education Endowment Foundation.

Fosnot, C. T. and Dolk, M. (2002) *Young Mathematicians at Work: Constructing Fractions, Decimals, and Percents*, Portsmouth, NH: Heinemann.

Francis, B., Archer, L., Hodgen, J., Pepper, D., Sienna, B. and Travers, M.-C. (2017) ‘Exploring the Relative Lack of Impact of Research on “Ability Grouping” in England: A Discourse Analytic Account, *Cambridge Journal of Education*, 47 (1), pp. 1–17.

Graham, J. W. (2009) ‘Missing Data Analysis: Making It Work in the Real World’, *Annual Review of Psychology*, (60), pp. 549–576.

Gravemeijer, K., Van den Heuvel, M. and Streefland, L. (1990) *Contexts, Free Productions, Tests, and Geometry in Realistic Mathematics Education*, Utrecht: OW and OC.

Hedges, L. V. and Rhoads, C. (2010) ‘Statistical Power Analysis in Education Research’, NCSER 2010-3006: https://ies.ed.gov/ncser/pubs/20103006/pdf/20103006.pdf

Hodgen, J., Foster, C., Marks, R. and Brown, M. (2018) Evidence for Review of Mathematics Teaching:Improving Mathematics in Key Stages Two and Three: Evidence Review’. London: Education Endowment Foundation.

Hough, S., Solomon, Y., Dickinson, P. and Gough, S. (2017) *Investigating the Impact of a Realistic Mathematics Education Approach on Achievement and Attitudes in Post-16 GCSE Resit Classes*, MMU/Nuffield.

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R. and Kerr, K. (2016) ‘Implementation and Process Evaluation (IPE) for Interventions in Education Settings: A Synthesis of the Literature’, Manchester Institute of Education: https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/IPE_Review_Final.pdf

Jerrim, J. and Choi, Á. (2014) ‘The Mathematics Skills of School Children: How Does England Compare to the High-Performing East Asian Jurisdictions?’, *Education Policy*, 29 (3), pp. 349–376.

Kelcey, B., Spybrook, J., Phelps, G., Jones, N. and Zhang, J. (2017) ‘Designing Large Scale Multisite and Cluster Randomized Studies of Professional Development, *Experimental Education*, 85 (3) pp. 389–410.

Lewis, G. (2011) ‘Mixed Methods in Studying the Voice of Disaffection with School Mathematics’, in Smith, C. (ed.) *Proceedings of the British Society for Research into Learning Mathematics*, 31 (3) November 2011.

Mason, J. (2002) *Researching Your Own Practice: The Discipline of Noticing*, Psychology Press.

Meelissen, M. and Luyten, H. (2008) ‘The Dutch Gender Gap in Mathematics: Small for Achievement, Substantial for Beliefs and Attitudes’, *Studies in Educational Evaluation,* 34 (2), pp. 82–93.

Nardi, E. and Steward, S. (2003) ‘Is Mathematics T.I.R.E.D.? A Profile of Quiet Disaffection in the Secondary Mathematics Classroom’, *British Educational Research*, 29 (3), pp. 345–67.

National Centre for Excellence in the Teaching of Mathematics (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

OECD (2014. *TALIS 2013 Results: An International Perspective on Teaching and Learning*, Paris: Organisation for Economic Co-operation and Development.

PISA (2000, 2006, 2009, 2015) ‘Programme for International Student Assessment’ (OECD): http://www.oecd.org/pisa/

Romberg, T. and Shafer, M. (2005) ‘The Longitudinal/Cross-Sectional Study of the Impact of Teaching Mathematics Using Mathematics in Context on Student Achievement: Implications and Conclusions’, Wisconsin Center for Educational Research: http://micimpact.wceruw.org/

Searle, J. and Barmby, P. (2012) ‘Evaluation Report on the Realistic Mathematics Education Pilot Project at Manchester Metropolitan University’, Durham University.

Spybrook, J., Shi, R. and Kelcey, B. (2016) ‘Progress in the Past Decade: An Examination of the Precision of Cluster Randomized Trials Funded by the U.S. Institute of Education Sciences’, *International Journal of Research and Method in Education*, 39 (3), pp. 255–267.

Stake, R. E. (2013) *Multiple Case Study Analysis*, New York, Guilford Press.

Streefland, L. (1985) ‘Wiskunde als activiteit en de realiteit als bron’ [Mathematics as an Activity and Reality As a Source], *Nieuwe Wiskrant*, 5 (1), pp. 60–67.

Swan, M. (2006) ‘Collaborative Learning in Mathematics: A Challenge to Our Beliefs and Practices’, Leicester: National Institute of Adult Continuing Education.

TIMSS (1999, 2007, 2010) ‘Trends in International Mathematics and Science Study’, TIMSS International Study Centre: https://timssandpirls.bc.edu/

Treffers, A. (1987) *Three Dimensions: A Model of Goal and Theory Description in Mathematics Instruction: The Wiskobas Project*, Dordrecht: Reidel.

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), pp. 9–35.

Van den Heuvel-Panhuizen, M. and Drijvers, P. (2014) ‘Realistic Mathematics Education’, in S. Lerman (ed.), *Encyclopedia of Mathematics Education* (pp. 521–525), London: Springer.

Webb, D., Boswinkel, N. and Dekker, T. (2008) ‘Beneath the Tip of the Iceberg: Using Representations to Support Student Understanding’, *Mathematics Teaching in the Middle School,* 14 (2), pp. 110–113.

Weiss, C. H. (1997) ‘How Can Theory-Based Evaluation Make Greater Headway?’, *Evaluation Review,* 21 (4), pp. 501–524.

Yackel, E. and Cobb, P. (1996) ‘Sociomathematical Norms, Argumentation, and Autonomy in Mathematics’, *Research in Mathematics Education*, 27 (4), pp. 458–477.

Appendix A: EEF cost rating

Figure 2: Cost Rating

Cost rating	Description
£ £ £ £ £	*Very low:* less than £80 per pupil per year.
£ £ £ £ £	*Low:* up to about £200 per pupil per year.
£ £ £ £ £	*Moderate:* up to about £700 per pupil per year.
£ £ £ £ £	*High:* up to £1,200 per pupil per year.
£ £ £ £ £	*Very high:* over £1,200 per pupil per year.

Appendix B: Security classification of trial findings

SECURITY RATING

OUTCOME: *GL PTM13*

Rating	Criteria for rating			Initial score		Adjust		Final score
	Design	**MDES**	**Attrition**					
5	Randomised design	<= 0.2	0-10%					
4	Design for comparison that considers some type of selection on unobservable characteristics (e.g. RDD, Diff-in-Diffs, Matched Diff-in-Diffs)	0.21 - 0.29	11-20%					
3	Design for comparison that considers selection on all relevant observable confounders (e.g. Matching or Regression Analysis with variables descriptive of the selection mechanism)	0.30 - 0.39	21-30%			Adjustment for threats to internal validity [-1]		
2	Design for comparison that considers selection only on some relevant confounders	0.40 - 0.49	31-40%					
1	Design for comparison that does not consider selection on any relevant confounders	0.50 - 0.59	41-50%	1				
0	No comparator	>=0.6	>50%					0

Threats to validity	Risk rating	Comments
Threat 1: Confounding	Low	Randomisation was conducted independently by the evaluation team. Groups were well balanced on baseline attainment (KS2 maths score, effect size -0.05)
Threat 2: Concurrent Interventions	Low	Both treatment and control schools were shown to use other maths interventions during the trial, but there is no suggestion that uptake of concurrent interventions was differential across groups.
Threat 3: Experimental effects	Low	Some control schools reported using mastery approaches which share some similarities with RME, but these were attributed to Maths Hub CPD or mastery training. All but one control school had no experience with RME for the duration of the trial.
Threat 4: Implementation fidelity	High	Due to the Covid-19 pandemic, the full RME programme was not delivered as intended: 1 of the 8 planned CPD days was not undertaken and teaching of the final two modules (of 10) was disrupted by partial school closures. However, even before the pandemic, attendance at CPD days was inconsistent (mean attendance of 4.3 sessions per teacher) and most schools did not deliver all the pre-Covid RME modules (mean of 3.7 modules delivered across two years). Importantly, due to pupil movement across Year 7 and Year 8, only 43% of the pupil ITT sample were observed to be in a class taught by an RME teacher in the second year of the intervention.
Threat 5: Missing Data	Moderate	Largely due to the impact of Covid-19, attrition at post-testing was 48%, and slightly higher in intervention relative to control pupils (50% vs. 45%).

		Models including predicted values obtained by multiple imputation yield a higher effect size (+0.08, 95% CIs: -0.17; +0.34), but the confidence intervals span zero.
Threat 6: Measurement of Outcomes	Moderate	Due to restrictions posed by Covid-19, outcome assessments were not independently invigilated. Post-testing was also conducted one year after originally intended, which may have led to dissipation of impact. Given the delays, pupils were beyond the intended upper age band of the PTM13 when outcome testing was conducted, but no ceiling effects were observed.
Threat 7: Selective reporting	Low	Deviations from the pre-specified SAP are minor and clearly reported.

- **Initial padlock score:** 1 Padlock – Cluster RCT with MDES at randomisation of 0.21-0.23 and attrition of 48%, due largely to the Covid-19 pandemic
- **Reason for adjustment for threats to validity**: -1 Padlocks – Low implementation fidelity, including pupil movement across Year 7 and Year 8 which led to reduced exposure to the intervention, decreases confidence that RME was delivered and received as designed
- **Final padlock score:** initial score adjusted for threats to validity = 0 Padlocks

Appendix C: Sample analysis code

Empty / Null Model:

mixed GLPTM || School_ID: || Class_ID:

Outcome Only:

mixed GLPTM Group|| School_ID: || Class_ID:

KS2 to GL PTM13 Progress:

mixed GLPTM Group KS2Maths_SchC KS2Maths_ClassC KS2Maths_PupC || School_ID: || Class_ID:

Final (headline) analyses:

mixed GLPTM Group KS2Maths_SchC KS2Maths_ClassC KS2Maths_PupC b1.Hub Phase || School_ID: || Class_ID:

Appendix D: Effect size estimation

As specified in the SAP, the unconditional variance from the empty (null) 3-level regression was used to estimate the Hedges g effect size. Appendix Table 2 presents the null and final models for the ITT analysis of the GL PTM13 primary outcome. The calculation of the estimated effect size is detailed below the table.

Appendix Table 1: Effect size estimation

	Null Model		Final Model	
Variable	Coef (95% Cis)	p value	Coef (95% Cis)	p value
Constant	28.3 (26.6; 30.0)	<0.01	27.7 (23.5; 31.9)	<0.001
Group Membership	-	-	**0.72 (-1.10; +2.53)**	0.438
KS2 Maths Attainment at pupil level (class centred)	-	-	1.38 (1.30; 1.46)	<0.001
KS2 Maths Attainment at class level (school centred)	-	-	1.78 (1.69; 1.87)	<0.001
KS2 Maths Attainment at school level (GM centred)	-	-	2.03(1.48; 2.57)	<0.001
School reported to set/stream	-	-	-2.01 (-5.38; 1.36)	0.243
Geographical Hub (reference Yorks & Humber)				
NW	-	-	2.97 (-0.68; 6.61)	0.111
West Mids	-	-	1.75 (-2.51; 6.02)	0.420
East Mids	-	-	3.29 (-0.88; 7.45)	0.122
London	-	-	4.39 (0.46; 8.32)	0.029
South	-	-	3.49 (-0.34; 7.32)	0.074
Variance	Estimate		Estimate	
School level	10.23		10.23	
Class level	119.13		5.64	
Residual	157.48		124.03	
Total	**286.83**		139.9	

The total unconditional variance was observed as 286.83, the square root of this (16.94) provided the standard deviation used to convert the group membership coefficients and confidence intervals into effect sizes:

ITT Effect size = $\frac{Group\ membership\ coefficient}{\sqrt{total\ unconditional\ variance}} = \frac{0.72}{\sqrt{286.83}} = \frac{0.72}{16.94} =$ +0.042 sds

The upper and lower confidence intervals were similarly converted:

95% CI (Lower) = $\frac{-1.10}{16.94}$ = -0.065 sds;   95% CI (Upper) = $\frac{2.53}{16.94}$ = +0.149 sds

Appendix E: Additional trial documentation

Realistic Mathematics Education Trial Evaluation

Parent and carer information update

Dear parent/carer,

As you may remember, your Year 9 child has been taking part in a mathematics education intervention known as Realistic Mathematics Education (RME - see **RME.org.uk**), funded by the Education Endowment Foundation (EEF). This letter is to update you on our plans for evaluating the trial and what this will mean for your child. It reminds you about your right to withdraw your child from the evaluation process as explained in our earlier communications.

The RME intervention - update due to Covid-19 school closures

To determine whether RME leads to improvements in mathematics attainment, the EEF is funding a randomised controlled trial (RCT) of RME, delivered by teacher educators, curriculum designers and researchers at Manchester Metropolitan University. The intervention began when your child was in Year 7 in 2018/19 and continued into Year 8 in 2019/20. It was due to finish with a final test, the GL Progress Test in Mathematics (PTM), in June or July 2020. However, due to the disruption caused by Covid-19, the project has been extended into 2020/21, to ensure that the RME schools have time to cover all of the content.

The evaluation

The RCT has been running in 119 secondary schools in England. Half of these schools (60) were randomly selected to receive the RME programme and the other half formed a control group. Your child is in an RME school. In these schools, two mathematics teachers used Year 7 RME materials with their mathematics class in 2018/19, and then used Year 8 RME materials when the class moved to Year 8 in 2019/20. In 2020/21, Year 9 RME materials have been made available.

Your child's involvement

As part of this evaluation, Year 9 pupils in both control and intervention schools will take the GL Progress Test in Mathematics in April or May 2021 – this test was due to take place in summer 2020 while your child was in Year 8, but school closures due to Covid-19 meant it was necessary to postpone. The test results will be collected by SHU and marked by GL Assessment. The test will be taken in class as part of the normal school day.

By comparing results from the two sets of schools, we will see how well RME works. This has the potential to benefit not just your child, but children in schools across the country.

Your child's data

We take data security very seriously and ensure that your child’s data is handled appropriately and securely. Data will be shared and processed in accordance with the General Data Protection Regulation (GDPR) and the Data Protection Act 2018.

For the purposes of academic research, test results will be linked with information about the pupils from the National Pupil Database (NPD) and shared with MMU, the Department for Education, the EEF, FFT Education (EEF’s data processor for their archive), the Office for National Statistics (as host of the EEF archive) and potentially with other research teams and the UK Data Archive. Your child's test data will also be sent to the school so that it can be used to gauge progress and inform mathematics teaching.

All EEF trial data is stored in the EEF data archive. The archive does not contain direct identifiers like pupil name, contact details and date of birth, but does hold a Pupil Matching Reference (PMR). The PMR is used for further matching to the NPD and other administrative datasets that may be required as part of subsequent research. We will not use pupil names or school names in any report arising from the research. For more detail, please see the evaluation fair processing notice: **https://bit.ly/2LKrzZV**

Your right to withdraw your child from the evaluation process

You are free to withdraw your child from the evaluation process at any time during the trial period. If you do decide to withdraw your child, no data about them would be included in the evaluation and your child would not be required to take the Progress Test in Mathematics in Year 9. If you have previously withdrawn your child there is no need to tell us again.

For further information, or to withdraw your child, please contact the project managers at Sheffield Hallam University using the contact details below.

Martin Culliney (Project manager) Realisticmaths@shu.ac.uk 0114 225 6072

Realistic Mathematics Education (RME) Evaluation

Sheffield Hallam University and Manchester Metropolitan University
Fair Processing Notice

Introduction

This document accompanies the Memorandum of Understanding and outlines the responsibilities of Manchester Metropolitan University (MMU) and Sheffield Hallam University (SHU) in handling personal data collected from participants as part of the Realistic Mathematics Education (RME) trial and evaluation. The trial is being funded by the Education Endowment Foundation (EEF). Participants include: teachers, pupils and parents/carers.

From 25 May 2018 the General Data Protection Regulation (GDPR) replaces the Data Protection Act and governs the way that organisations use personal data. Personal data is information relating to an identifiable living individual.

Transparency is a key element of the GDPR and this Data Protection Statement is designed to inform participants about:

- how and why SHU and MMU will use personal data collected in this evaluation
- what participants' rights are under GDPR, and
- how to contact us to exercise those rights.

This Fair Processing Notice is compliant with the Data Protection Act 2018

Participants' Rights

One of the aims of the General Data Protection Regulation (GDPR) is to empower individuals and give them control over their personal data. The GDPR gives participants the following rights:

- the right to be **informed**
- the right of **access**
- the right to **rectification**
- the right to **erase**
- the right to **restrict** processing
- the right to **data portability**
- the right to **object**
- rights in relation to **automated decision making and profiling**

For more information about these rights please see: **https://ico.org.uk/your-data-matters/** and: **https://www.shu.ac.uk/about-this-website/privacy-policy/data-subject-rights/subject-access-request**

Participants can contact SHU/MMU at any time to:

- request copies of their own personal data held by SHU/MMU (**a subject access request**)
- exercise **other rights** (e.g. to have inaccurate data rectified, to restrict or object to processing)
- **query** how data is used by SHU/MMU
- report a **data security breach** (e.g. if there are concerns that personal data has been lost or disclosed inappropriately)
- **complain** about how SHU/MMU have used personal data.

Details of who to contact are provided at the end of this statement.

Why are we processing participants' personal data?

It is necessary for SHU/MMU to process some personal data, in order to evaluate the impact of RME. This will help to strengthen the evidence base about mathematics teaching in schools in order to inform future policy development.

Retention

After the evaluation with EEF is complete, SHU and MMU will retain participants' data for research and knowledge-exchange purposes, including presentations at professional or academic conferences, or publications in professional or academic journals, for a period of ten years after the last publication arising from the evaluation. After this period, SHU and MMU will review the longer-term archival value of the data.

Respecting confidentiality

In the production of professional or academic publications or presentations, all data will be fully anonymised and no individual or school will be identified or identifiable. Should we wish to present or publish any information where a school may be identifiable, for example an exemplar case study of how a school has improved as a result of participation in a RME project, we will seek the school's consent for this. Schools will be entirely free to refuse this and we would therefore ensure the school remained anonymous.

What is the legal basis for processing activities?

SHU and MMU are joint Data Controllers for the RME evaluation. The processing of personal data through the RME evaluation is defined under GDPR as a specific task in the public interest. The legal basis for processing your personal data is 'Public Task' (Article 6 (1) (e)). **https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/public-task/**

Which Personal Data will we collect and use?

In order to provide our services we need to collect and use some personal data. Below is a list of what this may include for the trial:

Type of personal data	Pupil	Teacher
Names	X	X
Contact details		X
Personal characteristics data including: Pupil Name, UPN, FSM status and gender	X	
Attitudinal survey responses	X	X
Attitudinal interview responses		X
Progress, achievement and attainment data held by schools, and in the National Pupil Database	X	
Outcome test data, to be processed by GL	X	
Sample pupil work	X	
Data on participation in the RME project and use of materials		X

For SHU case study schools:

Type of personal data	Pupil	Teacher	SLT
Interview responses		X	X
Video recording of lesson	X	X	
Focus group data	X		

For SHU telephone interview schools:

Type of personal data	Teacher
Interview responses	x

For MMU Design schools:

Type of personal data	Pupil	Teacher
Interview responses	X	X
Video recording of lesson	X	X
Sample of pupil/teacher work	X	X

Using the information we receive from schools, we will also obtain data from sources such as the DfE Schools Comparison Service and the National Pupil Database.

Who will we share personal data with?

The privacy of personal data is paramount and will not be disclosed unless there is a justified purpose for doing so. Data may be shared between SHU/MMU and:

- **EEF** for the purposes of research and evaluation**.** This includes submitting project data to the archive managed by the Fischer Family Trust (FFT) at the end of the project. At this point, EEF becomes a data controller, and FFT becomes a data processor.
- **GL Assessment** will mark the outcome assessment papers and will have access to pupil names and SHU generated pupil ID numbers as part of this process.
- **Transcribers**, who we may ask to produce transcripts of audio recordings of interviews and focus groups. If this is the case SHU and MMU will ensure that appropriate contracts and/or data-sharing agreements are in place and that the transcribers process personal data in accordance with the GDPR and other applicable legislation.

SHU and MMU NEVER sell personal data to third parties

Security

SHU and MMU takes a robust approach to protecting the information they hold. This includes the installation and use of technical measures including encryption of data, firewalls and intrusion detection and prevention tools on networks and segregation of different types of device; the use of tools on University computers to detect and remove malicious software and regular assessment of the technical security of SHU and MMU systems. SHU and MMU staff monitor systems and respond to suspicious activity. SHU also has Cyber Essentials certification.

Alongside these technical measures there are comprehensive and effective policies and processes in place to ensure that SHU and MMU users and administrators of information are aware of their obligations and responsibilities for the data they have access to. Access to project data is restricted to the research and evaluation teams and administrators associated with the project. Any sharing of the data with other researchers would require approval by the SHU Faculty of Development and Society ethics committee who will ensure that all data protection requirements are met. Training is provided to new staff joining SHU and MMU, and existing staff have training and expert advice available if needed.

Further Information and Support

For further information about how SHU and MMU use personal data see:

https://www.shu.ac.uk/about-this-website/privacy-policy/privacy-notices/privacy-notice-for-research

https://www.shu.ac.uk/about-this-website/privacy-policy/information-governance-policy

http://www2.mmu.ac.uk/data-protection/ (See ‘For Students and the Public’ section)

The Information Commissioner is the regulator for GDPR. The Information Commissioner's Office (ICO) has a website with information and guidance for members of the public:

https://ico.org.uk/for-the-public/

If there are any concerns about the way this project processes personal data, please raise these with the project teams.

Contact details

SHU

Sean Demack (SHU RME Project Director)

Lead Statistician, SIoE,
Sheffield Hallam University S1 1WB

S.Demack@shu.ac.uk

OR

Governance Services

City Campus, Howard Street

Sheffield S1 1WB

foi@shu.ac.uk

0114 225 5555

MMU

Professor Yvette Solomon

Education and Social Research Institute

Manchester Metropolitan University

Brooks Building

53 Bonsall Street

Manchester M15 6GX
Y.Solomon@mmu.ac.uk
0161 247 2500

If you have an ongoing concern, you can contact the Information Commissioner's Office, the body responsible for enforcing data protection legislation in the UK, at **https://ico.org.uk/concerns/**

Appendix F: Sample of teacher survey respondents

Appendix F Table 1: Number of responses each year when sample restricted to teachers that did both baseline and Year 7 survey (as is the case for all RQ9 and RQ10 analyses)

	Control N	Intervention N
Baseline	40	45
End Year 7	40	45
End Year 8	29	30

Source: SHU teacher survey

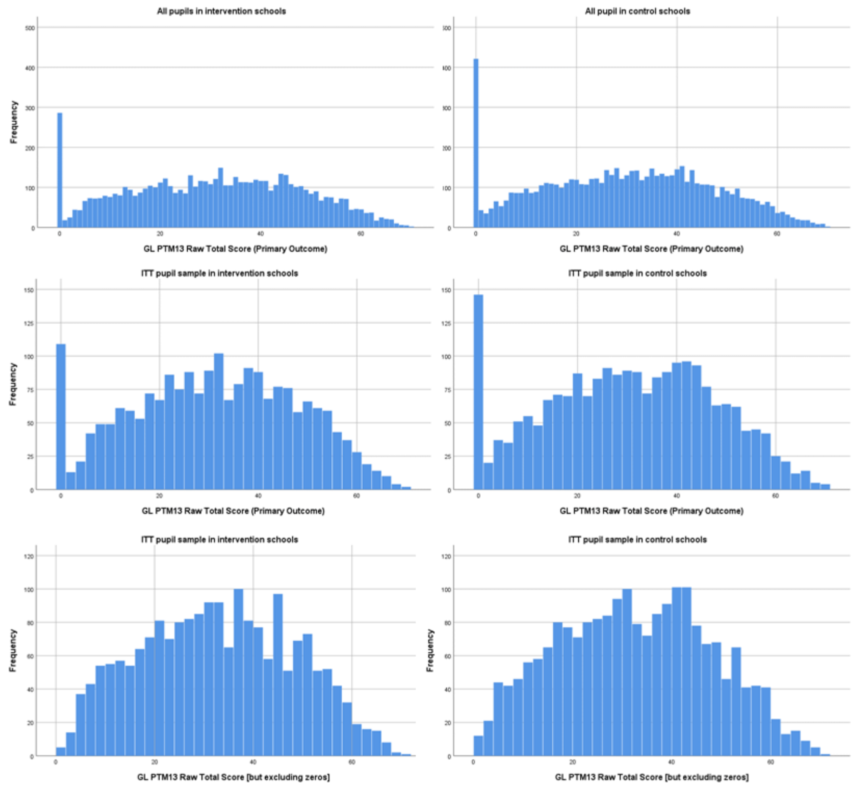
Appendix G: Distribution of GL PTM13 primary outcome

Table and Figure present the pupil level distribution of the GL PTM13 primary outcome for the intervention and control group samples. Three pupil samples are shown; all pupils in the 79 schools that participated in outcome testing; the ITT sample of pupils and the ITT sample excluding pupils who scored zero in the PTM13 test.

Appendix G Table 1: Statistical Summaries for GL PTM13 primary outcome

Sample	Intervention Schools		Control Schools		Effect Size
	N (missing)	Mean (95% Cis)	N (missing)	Mean (95% Cis)	
All pupils	5928 (5654)	30.8 (30.4; 31.3)	6689 (4498)	29.7 (29.3; 30.1)	+0.06 sds
ITT Sample	2054 (1981)	30.6 (29.9; 31.3)	2200 (1721)	30.0 (29.3; 30.7)	+0.04 sds
ITT Sample [excluding pupils with zero score in GL PTM13)	1945 (2090)	32.3 (31.7; 33.0)	2054 (1867)	32.1 (31.4; 32.8)	+0.02 sds

Appendix G Figure 1: Distributions of the GL PTM13 primary outcome



Upon observing the cluster of zero scores we undertook an additional sensitivity of the PTM13 outcome where pupils who scored zero were excluded. This sensitivity analysis agreed with the main ITT analysis and so we conclude that the cluster of zero scores did not result in distorting the estimated impact of RME.

Appendix H: Distribution of KS2 Maths covariate

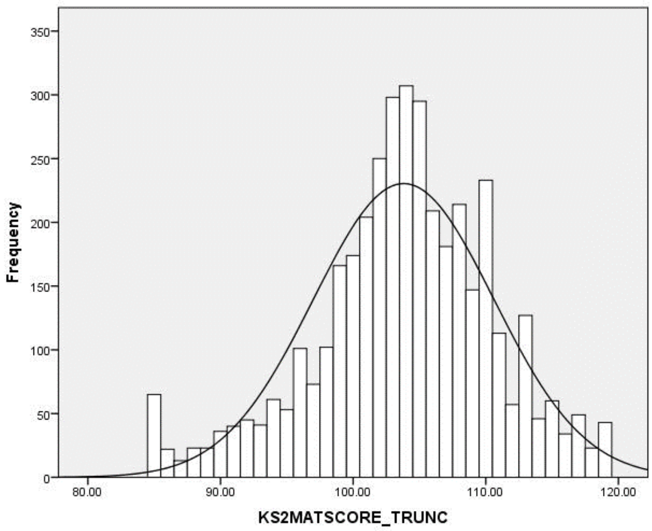
Table and Figure present the pupil level distribution of the KS2 Maths covariate for the intervention and control group ITT samples. The Table presents a statistical summary for the original raw distribution of KS2 maths. The histograms use a truncated version of the KS2 maths variable to comply with the statistical disclosure requirements of ONS SRS (i.e. to ensure none of the bars related to a count of 10 cases or fewer). At the lower end, the truncation happened at 85 (scores between 80 & 85 were combined) and the upper end was truncated at 119 (119 and 120 combined).

Appendix H Table 1: Statistical Summaries for KS2 Maths

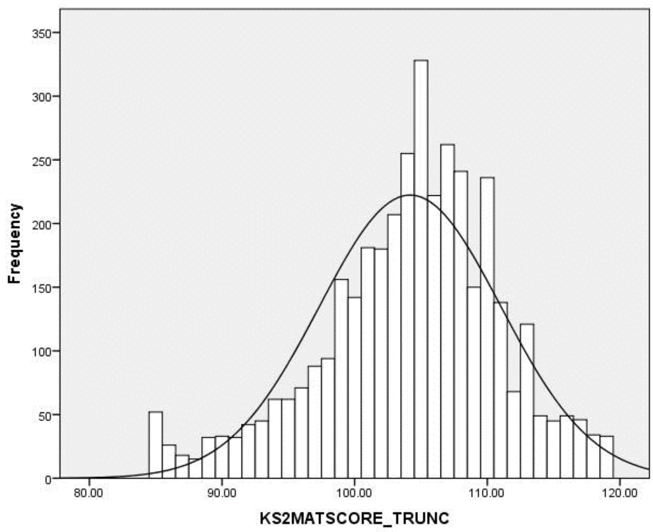
Sample	Pupils in RME Intervention Schools		Pupils in Control Schools		Effect Size
	N (missing)	Mean (95% Cis)	N (missing)	Mean (95% Cis)	
ITT Sample	3928 (107)	103.8 (103.6; 104.0)	3815 (106)	104.2 (103.9; 104.4)	-0.05 sds

Appendix Figure: Distributions of the truncated KS2 Maths covariate

Pupils in RME Intervention Schools



Pupils in Control Schools



Appendix I: ICC estimates and covariate explanatory power

As specified in the SAP, for sensitivity analyses we replicated the ITT model excluding the class level. Table 1 below summarises the extent of clustering observed at school and class levels (from null models) and the covariate explanatory power at school, class and pupil levels (from the final model). Covariate explanatory power is shown twice; first when the KS2 Maths covariate was included at multiple levels (using appropriate centring) and second when just a single raw pupil-level covariate was included.

Appendix I Table 1: ICC estimates and covariate explanatory power for 2- and 3-level analyses

	2 level model	3 level model
Sample Details		
Number of schools	79	79
Number of classes	-	216
Number of pupils	4159	4159
ICC Estimates		
School level	0.181	0.036
School + Class levels	-	0.451
Class level	-	0.415
Residual	0.819	0.549
Covariate Explanatory Power; R^2 including KS2 maths (centred) at multiple levels		
School level	0.749	-0.001
Class level	-	0.953
Residual	0.454	0.212
Total (pupil)	0.508	0.512
Covariate Explanatory Power; R^2 including raw KS2 maths at just pupil level		
School level	0.741	-0.046
Class level	-	0.932
Residual	0.454	0.211
Total (pupil)	0.506	0.501

Appendix J: End of first year survey for intervention teachers

RME teacher end Year 7 survey

This short survey is for all teachers participating in the RME trial and attending the training programme during 2018/19. It begins with questions on your involvement in the trial, your use of RME this year, and your views on RME. These are followed by a set of questions about how you teach mathematics, your beliefs about mathematics teaching, and your confidence in various aspects of teaching. These latter questions are the same as those we asked you at the start of the project. We would like to know how things are changing in your teaching, and at the end of the trial, we would also like to look at any relationship between teachers' survey answers and pupil outcomes. For this reason, we ask you to provide your name and school so that we can link your responses to other data. Personal details will not be used in any reporting of outcomes, and will not be shared with the Manchester Metropolitan University team during the trial; they will only be shared with them after the trial if you have given additional permission. Sheffield Hallam University staff will have access to the full set of responses including your personal details. Manchester Metropolitan University will only have access to anonymised data. Submission of your responses will be deemed as consent to use the data provided for the evaluation of the RME trial. Further information about the trial, along with links to the school information sheet and Fair Processing Notice, can be found here. Please complete the survey by July 12th 2019.

Section 1 - Teacher characteristics

Q1 Your name

____

Q2 School name

____

Q3 School postcode

____

Section 2 - Involvement in Realistic Mathematics Education

Q4 Were you able to attend all the RME training days in the year 2018/19?

- ○ Yes
- ○ No

Display This Question:

If Were you able to attend all the RME training days in the year 2018/19? = No

Q4b If you were not able to attend all of the PD days, please tick all that apply

- ☐ A colleague attended as a substitute when I could not come
- ☐ Senior leadership would not release me
- ☐ Department leadership would not let me out of school that day
- ☐ My Head of Department/faculty lead decided only one person could attend
- ☐ Senior leaders decided only one person could attend
- ☐ Cover could not be found for me that day
- ☐ I was teaching a Y11 class that day and this had to take priority
- ☐ Too many staff were already out that day
- ☐ I had a parents' evening that day
- ☐ I was ill and so not working that day
- ☐ Other ____

Q5 How many different Year 7 classes have you used RME modules with during 2018/19?

- ○ 0
- ○ 1
- ○ 2
- ○ 3 or more

Q6 Below is a table about teaching of different module materials this year. Please complete this for your classes. If it was different for different classes you taught please use the comment box to explain.

	I have taught this module	I did not teach the module	I have not taught the module yet but intend to do so before the end of the year	The module was taught by another teacher	Other (please specify below)
Year 7 Number	○	○	○	○	○
Year 7 Geometry	○	○	○	○	○
Year 7 Proportional Reasoning	○	○	○	○	○
Year 7 Data	○	○	○	○	○
Year 7 Algebra	○	○	○	○	○

Q6a If you have selected other for Q6, please provide details here.

____

Q7 Below is a table about the amount of teaching time used for different modules. Please complete this for your classes. If the module was taught by another teacher please provide the information if you know it, and if it was different for different classes you taught please use the comment box to explain:

	I did not teach the module	Less than 1 week of maths lesson time	Approx. 1 week of maths lesson time	Approx. 2 weeks of maths lesson time	Approx. 3 weeks of maths lesson time	4 or more weeks of maths lesson time	Have not yet taught the module	Don't know
Year 7 Number	○	○	○	○	○	○	○	○
Year 7 Geometry	○	○	○	○	○	○	○	○
Year 7 Proportional Reasoning	○	○	○	○	○	○	○	○
Year 7 Data	○	○	○	○	○	○	○	○
Year 7 Algebra	○	○	○	○	○	○	○	○

Q8 Please add any other comments about the time teaching RME below:

____

Q9 Thinking about your Year 7 classes, do you use RME ways of teaching and learning when you are not using the trial materials *(in your regular teaching in between RME modules)?*

○ No

○ Yes (please use the text box to explain) ____

Q10 Between training days you were encouraged to do 'gap tasks'. For each gap task please indicate the extent to which you undertook the task:

	Did the gap task as suggested or similar to this	Did the gap task in a modified or adapted way	Have not done the gap task but intend to	Have not done the gap task	Other (please explain below)
Using diagnostic assessment questions	○	○	○	○	○
A collaborative planning and observation activity	○	○	○	○	○
Annotating module guides	○	○	○	○	○

Q10a If you have selected 'other' for Q10, please explain here:

____

Section 3 - Mathematics teaching In the next six questions, different activities that you or students might do in your mathematics lessons are listed in sets of four. Think about your mathematics teaching last year. Consider how important these activities are in your teaching. They may be important because they happen often in your lessons or because you think they are important for student learning, or a combination of both of these or other reasons.

Q11 Please rank the activities in order of importance by clicking and dragging to make the MOST important first.

______ I ask students to decide on their own procedures or methods to solve non-routine problems
______ I steer students away from non-standard or inefficient methods
______ I encourage students to explore alternative methods for solutions
______ I focus on the most efficient way to solve a particular kind of problem

Q12 Please rank the activities in order of importance by clicking and dragging to make the MOST important first.

______ I ensure that students know helpful rules, mnemonics and memory aids to answer questions
______ I keep students focused on the essential knowledge or skills needed to answer mathematics questions.
______ I use activities that help students understand how and why a procedure works
______ I create opportunities to make connections between different mathematical topics

Q13 Please rank the activities in order of importance by clicking and dragging to make the MOST important first.

______ I start topics by teaching how to interpret the type of diagrams or representations which are most frequently used in a particular topic
______ I reinforce the idea that drawing diagrams or sketches is valuable for understanding the mathematics
______ I use a variety of mathematical representations or models for the same concept
______ I encourage students' own representations or models in mathematics and provide opportunities for sharing them in class

Q14 Please rank the activities in order of importance by clicking and dragging to make the MOST important first.

______ I emphasise that knowing how you arrived at a solution is as important as getting the answer right
______ I make sure that students' misconceptions/errors are corrected straight away
______ I move the class on once enough students are using the procedure or rule correctly
______ I encourage the strategies that students choose to use even though they may not be the most efficient or standard ones

Q15 Please rank the activities in order of importance by clicking and dragging to make the MOST important first.

______ When I ask questions in whole class discussion, if a student gives a wrong answer, I make sure that the mistake is corrected before moving on
______ I encourage discussions where students question each other and explain their thinking
______ I keep students' talk in whole class discussion on topic to make sure key teaching points are made
______ I provide opportunities for students to make conjectures about mathematical ideas

Q16 Please rank the activities in order of importance by clicking and dragging to make the MOST important first.

______ I link mathematics to contexts that are part of my students' daily lives
______ Context is used as a vehicle to develop the maths, and is used to do this throughout the lesson/module
______ I introduce contexts as opportunities to apply procedures and knowledge once key ideas are learnt
______ Students mainly get experience of applying mathematics in context through doing word problems

Section 4 - My beliefs about maths teaching

Q17 Please indicate the extent to which you agree or disagree with the following:

	Strongly agree	Some what agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
Discussing, explaining and thinking about mathematics is more important than the amount of the curriculum covered	○	○	○	○	○
Teaching secure use of procedures is more important than facilitating classroom discussion	○	○	○	○	○
Complex/open-ended problems are worthwhile for high attaining students but not for those who find maths difficult	○	○	○	○	○
Drawing diagrams is useful in solving shape and space problems but can confuse pupils at other times	○	○	○	○	○
Context is useful to motivate and initiate a topic, but should be dropped once students have grasped the concept/skills required	○	○	○	○	○
Students should learn basic skills before being asked to solve non-routine mathematical problems.	○	○	○	○	○
The most important part of teaching mathematics is explaining ideas and procedures clearly	○	○	○	○	○
There is usually a best method for solving a mathematics problem and my job is to make sure students learn that method	○	○	○	○	○
I prefer problems that can be solved in several different ways	○	○	○	○	○
It is important not to confuse students by teaching too many ways of doing the same sort of problem	○	○	○	○	○
Having a go is meaningful progress even if the students’ strategy is not productive	○	○	○	○	○
Students need to think aloud to work through their ideas	○	○	○	○	○
Contexts should regularly be used at the start of topics to generate a discussion of strategies	○	○	○	○	○

Section 5 - Confidence

Q18 Please indicate your level of confidence in each of the following aspects of teaching maths:

	Very confident	Somewhat confident	Neither confident nor unconfident	Somewhat unconfident	Very unconfident
I can meet the needs of all students no matter what their prior attainment	○	○	○	○	○
I can meet the needs of all individual students in my mathematics class.	○	○	○	○	○
I can build on students' contributions to navigate to my goals	○	○	○	○	○
I can get my students to understand underlying concepts in mathematics	○	○	○	○	○
I can manage whole class discussion over an extended period of time	○	○	○	○	○
When asking open questions, I can deal with the variety of ideas that students come up with	○	○	○	○	○
I can link mathematics to contexts that are meaningful to my students	○	○	○	○	○
I can get students to engage in mathematical discussion	○	○	○	○	○
I can use multiple models or representations to teach an idea or skill	○	○	○	○	○
I can get students to explain their mathematical thinking	○	○	○	○	○
I can start a lesson from a contextualised problem rather than on the basis of lesson objectives	○	○	○	○	○

Q19 Below is a list of things you might do to get the most out of the RME materials if you were advising another teacher, what would you say are the three most important to develop?

- ☐ Getting students to come to the board
- ☐ Using 'pass the pen' effectively
- ☐ Moving myself away from the board to elsewhere in the room
- ☐ Choosing which students’ strategies to work on in whole class discussion
- ☐ Managing students’ contributions about context
- ☐ Managing students’ contributions about their strategies for solving problems
- ☐ Working on context to develop mathematical models or /structures
- ☐ Agreeing ways of sharing and discussing ideas in RME/mathematics lessons
- ☐ Supporting students to listen to others
- ☐ Encouraging multiple responses
- ☐ Remaining neutral about responses or ideas
- ☐ Having students show their work for example by using mini white boards or visualisers

Q20 Have you used RME with other year groups?

◯ No

◯ Yes ______________________________

Q21 Have you used do you use RME ways of teaching and learning with other year groups with or without using RME materials?

◯ No

◯ Yes ______________________________

Q22 Which best describes how you came to be involved in RME project?

◯ I saw information and asked my Head of department/faculty or senior leader if we could apply

◯ I was asked I wanted to be involved by my Head of department/faculty or senior leader

◯ I was told that I would be taking part by a Head of department/faculty or senior leader

◯ Other (please explain in the box below) ______________________________

Section 6 - Experience and qualifications

Q23 Please drag the slider across to indicate how many years' experience you have in the following:

	0 4 8 12 16 20 24 28 32 36 40
Secondary school teaching ()	
Secondary mathematics teaching ()	

Q24 What is your highest mathematics qualification (for degrees, mathematics must be at least 50 per cent of the content)?

- ○ GCSE
- ○ A level
- ○ Bachelors/first degree
- ○ Masters
- ○ PhD
- ○ Other (please explain) ________________

Q25 Is your initial teaching qualification in mathematics?

- ○ Yes
- ○ No
- ○ Other (please explain) ________________

Section 7 - Current role

Q26 Please specify any designated roles or responsibilities that you hold:

- ○ HoD or equivalent
- ○ Teaching and learning responsibility in mathematics (one or more)
- ○ Teaching and learning responsibility in something other than mathematics
- ○ Other (please explain) ____________________

Q27 What are your current working arrangements?

- ○ Full-time
- ○ Part-time

Q28 What percentage of your teaching time is currently spent teaching mathematics?

	0 10 20 30 40 50 60 70 80 90 100
Time spent teaching mathematics	

Q29 Which Key Stages do you currently teach? Please select all that apply:

- ☐ KS2
- ☐ KS3
- ☐ KS4
- ☐ KS5

Q30 Please let us know about anything else that you think is relevant to your participation in the RME trial:

____

____

____

____

____

That was the final question. Please click the forward arrow to submit your responses. Submission of your responses is taken as consent for your answers to be processed and analysed for the purposes set out in the Fair Processing Notice.

Appendix K: Missing Data Analyses

Appendix K table 1: Multilevel Logistic Regression Output: 1= case included in ITT analyses of primary outcome; 0= missing case.

Variable	Coef (SE)	Coef (SE)
Constant	-2.9 (1.44)**	+4.60 (6.29)
Group Membership (RME compared with control)	-1.6 (0.71)**	-1.14 (0.66)*
KS2 Maths Attainment at pupil level (class centred)	+0.02 (0.01)***	+0.02 (0.01)*
KS2 Maths Attainment at class level (school centred)	+0.06 (0.01)**	+0.06 (0.01)***
KS2 Maths Attainment at school level (GM centred)	+0.02 (0.22)	-0.36 (0.26)
School reported to set/stream (1) compared with not	+0.17 (1.19)	+1.66 (1.18)
Geographical Hub (reference Yorks & Humber)		
NW	+2.59 (1.26)**	+0.12 (1.23)
West Mids	-0.83 (1.29)	-2.56 (1.33)*
East Mids	+4.81 (1.71)***	+0.74 (1.56)
London	+1.25 (1.31)	-1.15 (1.35)
South	+3.13 (1.41)**	+1.62 (1.53)
%FSM in last 6 years	-	-10.38 (3.74)***
% English as first language	-	-3.45 (1.91)*
% SEN (with statement or without)	-	+8.81 (4.17)**
% attaining 9-4 in GCSE English & Maths	-	+5.95 (7.96)
Mean Attainment 8 Score	-	-0.10 (0.16)
Sample Details		
Number of pupils	7,743	6,820
Number of maths classes	327	289
Number of schools	119	105
ICC estimates (95% Cis)		
School	0.86 (0.80; 0.90)	0.78 (0.71; 0.84)
Class & School	0.87 (0.81; 0.91)	0.79 (0.72; 0.85)

Statistical significance of model coefficients: * (p<0.10), ** (p<0.05), *** (p<0.01).

The first model shown in Table 1 only includes variables used within the ITT analyses. This first model identified group membership (school level, pupils in RME schools more likely to be missing); KS2 attainment (pupil and class levels; higher attainment associated with lower attrition) and geographical hub (lower attrition seen for pupils in NW, East Mids and South regions) as statistically significant.

The second model shown in Table 1 includes additional school-level variables not included in the ITT analyses. This second model identified group membership (school level, pupils in RME schools more likely to be missing); KS2 attainment (pupil and class levels; higher attainment associated with lower attrition); geographical hub (lower attrition seen for pupils in West Mids regions); %FSM (higher concentrations associated with higher attrition), % English as first language (higher concentrations associated with higher attrition) and %SEN (higher concentrations associated with lower attrition) as statistically significant.

Finally, the ICC estimates from the logistic regression model highlight that missingness was largely due to whole schools dropping out of the trial.

The Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP

https://educationendowmentfoundation.org.uk

@EducEndowFoundn

Facebook.com/EducEndowFoundn