Multiplicative reasoning professional development programme : evaluation

BOYLAN, Mark <http://orcid.org/0000-0002-8581-1886>, DEMACK, Sean <http://orcid.org/0000-0002-2953-1337>, WILLIS, Benjamin <http://orcid.org/0000-0002-4346-0459>, STEVENS, Anna <http://orcid.org/0000-0002-9450-890X>, ADAMS, Gill <http://orcid.org/0000-0003-2088-8708> and VERRIER, Diarmuid <http://orcid.org/0000-0002-4148-3948>

Published version

BOYLAN, Mark, DEMACK, Sean, WILLIS, Benjamin, STEVENS, Anna, ADAMS, Gill and VERRIER, Diarmuid (2015). Multiplicative reasoning professional development programme : evaluation. Project Report. London, Department for Education.

Copyright and re-use policy

Department for Education

Multiplicative reasoning professional development programme: evaluation

Research report

June 2015

Mark Boylan, Sean Demack, Ben Willis, Anna Stevens, Gill Adams and Diarmuid Verrier – Centre for Education and Inclusion Research (CEIR), Sheffield Hallam University

Contents

List of figures

List of tables

Acknowledgements

The authors would like to thank:

- the participants in the multiplicative reasoning project and school leaders and headteachers in the control schools
- the school examination officers and teachers in both intervention and control schools who arranged for testing to be done
- Alderwood recruitment, Aspire People, Randstad Education and Vision for Education for providing invigilators
- the invigilators who invigilated the tests
- the teachers and pupils who took part in case study interviews
- colleagues from DfE and NCETM, especially Jo O'Shea the project manager, and Natasha Chippendale, Bob Maunder, Rob Tait and John Westwell
- Professor Andrew Izsák of the University of Georgia who analysed test items in relation to the content of the multiplicative reasoning project
- CEIR colleagues who contributed to the project, data management and reporting

Glossary

CDT - Core development team: The CDT comprised of NCETM national leads and curriculum developers
CRT - Clustered randomised trial
CT - Core teacher - teachers actively participating in the professional development events in intervention schools
DfE - Department for Education
DNL - Double Number Line a mathematical model used in the MRP project materials
GLA - GL assessment, provided and marked the Progress in Mathematics Tests
HoD - Head of Department
FSM: families with an income below a certain level are entitled to their children receiving free school meals. Used as a proxy measure of poverty.
KS2 - Key Stage 2: National curriculum phase in compulsory schooling in England, covering Years 3-6 (ages 7-11)
KS3 - Key Stage 3: National curriculum phase in compulsory schooling in England, covering Years 7-9 (ages 11-14)
KS4 - Key Stage 4: National curriculum phase in compulsory schooling in England, covering Years 10-11 (ages 14-16)
Lesson Study - collaborative approach to professional learning
Maths Hubs - Around 30 in total forming a national network, each hub is led by a teaching school and is intended to provide support to all schools in the area and lead on national projects
MDES - Minimum Detectable Effect Size (see Technical Report, Annexe E1)
MEI - Mathematics in Education and Industry
MESH - Mathematics Education Strategic Hubs. Embryonic school led networks, three of which hosted the MRP regional TIME events. The MESH were a pilot for what become Maths Hubs.
MR - Multiplicative reasoning
MRP - Multiplicative Reasoning Project
NAG - National Advisory Group, advised DfE on the project
NCETM - The National Centre for Excellence in the Teaching of Mathematics
NDT - National Development Team, comprised of the CDT, PD leads and university researchers
NPD - National Pupil Database: database containing information on pupils in maintained schools in England, including attainment and pupil characteristics.
OT - other teacher, teacher in an intervention school other than a core teacher
PD Leads - Professional Development Leads, led the professional development events
PE - Process evaluation
PIM tests - Progress in Mathematics Tests
RCT - Randomised Controlled Trial

TIME team - Teachers Improving Mathematics Education - name for the three regional groups of teachers involved in the project, and those working with them PD leads and university researchers
University researcher - Academic attached to each TIME team

Executive Summary[1]

The Multiplicative Reasoning Project

The Multiplicative Reasoning Project (MRP) delivered by the National Centre for Excellence in the Teaching of Mathematics (NCETM) in 2013/14 focused on developing teachers' understanding and capacity to teach topics that involved multiplicative reasoning to Key Stage 3 (KS3) pupils. Multiplicative reasoning refers to the mathematical understanding and capability to solve problems arising from proportional situations often involving an understanding and application of fractions as well as decimals, percentages, ratios and proportions. The aspects of multiplicative reasoning addressed in the NCETM project were proportional and fractional relationships. The teaching approaches encouraged were the use of mathematical models, visual approaches and problem solving strategies.

Approximately 60 teachers in 30 schools engaged in 3 regional professional development networks, led by professional development leaders, supported by university researchers. Specific project materials were created by a team of curriculum developers.

Evaluation

The impact of MRP on pupil outcomes was evaluated using a 3-level Clustered Randomised Control Trial (CRT) research design. The trial ran between October 2013 and June 2014 and involved 8,777 year 7 (Y7), year 8 (Y8) and year 9 (Y9) pupils (level 1) clustered into 418 Y7, Y8 and Y9 mathematics classes (level 2) clustered into 60 secondary schools. Approximately half the schools, teachers and pupils participated in the intervention and half formed a control group. Progress was compared between the two groups of pupils using Key Stage 2 (KS2) data as a base line and GL Assessment Progress in Mathematics Test (PiM) tests as an outcome measure. This is a measure of general mathematical attainment that is correlated with national tests.

Further analysis was completed specifically in relation to multiplicative reasoning. The GLA PiM tests were analysed and items related to the project materials were identified. Outcomes on these items in the intervention and control samples were analysed. In addition a fidelity analysis was carried out. Outcomes were considered

[1] This report has an accompanying Technical Report which should be consulted for further detail on methodology and details of analysis.

for the sub-sample of pupils who had been taught by teachers who had fully participated in the project, for example, by excluding pupils who had moved classes.

In addition, a process evaluation was conducted using a mixed methods approach of surveys, observations and interviews.

Please see the full report and the accompanying technical report for further details.

Headline findings in relation to research questions

Research Question 1

What is the impact of the programme on pupil outcomes on both general mathematical attainment as measured by GL assessment Progress in Mathematics (PiM) tests and on those items in the GL PiM associated specifically with multiplicative reasoning?

During the timescale of the trial (13 October 2014 to May 2015) the programme did not have any statistically significant impacts on general mathematical attainment as measured by PiM tests or on items on the tests specifically associated with multiplicative reasoning.

Research Question 2

What are the impacts (if any) on: pupils' relationships to mathematics; teacher beliefs and practice including on lesson planning; teacher knowledge of multiplicative reasoning pedagogy; capacity of core teachers' to lead professional development?

The project had a positive impact on pupils' relationship with mathematics as reported by teachers in surveys and interviews and pupils in focus group interviews.

A range of changes in teacher beliefs and practices were identified:

- thinking more deeply about mathematics prior to teaching
- using models when teaching MR and extending models to other areas
- developing questioning
- promoting independence
- focussing on student learning
- questioning assumptions

Many participants, who responded to surveys or participated in case study visits, reported an increased awareness of the importance of multiplicative reasoning and its relationship to other areas of mathematics. Some discussed specific issues about the relationship between additive and multiplicative understanding.

With regard to core teachers leading professional development, many participants shared materials with other members of their department. However, opportunities to lead more extended professional development activities were limited. Nevertheless approximately a quarter of participants involved another member of the department in the first Lesson Study activity. There was some evidence of development of leadership capacity and it appeared the project developed a desire to lead professional development in others.

Research Question 3:

How was the programme conducted and how did this differ from the planned programme in what way and why?

The programme was conducted as planned at national and regional levels involving collaboration between NCETM leads, curriculum developers, university researchers and professional development (PD) leads. At school level the programme was implemented by two core teachers in each school. Originally it was planned to have a stronger focus on core teachers leading PD in their own departments. During the recruitment phase this was judged by the NCETM and Department for Education (DfE) to be more appropriately something to be encouraged rather than required. There was a degree of loss of fidelity at regional level in relation to the extent to which materials were a resource to be drawn on or a recommended coherent curriculum. The extent to which teachers were able to fully participate in the programme depended on support in schools. In some schools the project was actively supported by senior and departmental leadership and there was a good fit with departmental approaches to mathematics teachers and/or strategic developments in the subject. In other schools there were various issues that meant alignment was lower and this created constraints on implementation.

Research Question 4

What are the views of teachers/development teams on the programme including its effectiveness?

Participants responding to surveys or participating in interviews judged both the curriculum materials and professional development activities to be, on the whole, effective. Aspects of the professional development activities that were viewed positively were:

- discussion of lessons
- time to plan with the other core teacher
- input on the pedagogical approach
- engaging in mathematics by trialling lesson activities

Aspects of the materials and resources that were produced that were viewed positively were the variety of materials; the use of realistic contexts, and visual models.

Issues were raised for improvement such as having the materials in advance of the training events and having the events nearer to the participants' schools. However, these may be related to the pilot nature of the project.

Research Question 5:

If/how was the programme effective and what lessons can be learnt for scalability?

The programme was effective as a curriculum design project and pilot project for teacher professional development. It did not meet the intention to impact on pupil attainment (which was measured by the GL PiM tests within the timespan of the project).

Given this, there is no rationale, in terms of an intention to impact on pupil attainment in the short term, for repeating the project in the same form and for the same duration with a larger number of schools. With further development or refinement a similar approach might lead to impact on those aspects of pupil learning measured by PiM tests. One possibility would be to increase the project length to allow time for professional learning to impact on pupil learning and then to re-evaluate.

A project with all the features of the MRP may not be scalable. For example, a positive feature was the opportunity for face to face dialogue between teachers and curriculum developers. This may difficult to repeat at scale. However, support could be given to collaborative professional development communities based on the TIME model or different forms of teacher networks using specific curriculum materials in relation to MR or other areas of the curriculum. There are examples in both recent and medium term past of this being done successfully.

Features of the project that were found to be effective and are reproducible at scale in future projects are:

- teacher professional development focused on research-informed curriculum materials
- teacher collaborative learning communities that draw on the expertise of teacher leaders, curriculum developer and researchers

The MRP brought together a team of curriculum developers whose designs were informed by different but complementary pedagogical principles. Developments were supported by other university researchers and teacher professional development leaders. This was recognised by the national development team members, including

teacher leaders, as a powerful and productive approach. Teachers interviewed also valued the different contributions.

The project aimed to address concerns about pupil understanding of multiplicative reasoning in KS3. Teachers found the models used in the projects as useful approaches to teaching multiplicative reasoning. Given this, there is the potential to evaluate specific materials in a more focused way - for example using the bar model and ratio table. This would then allow for an assessment of whether these models should be adopted more widely.

The project materials were found to be a useful resource for KS3 and potentially KS2 curriculum. They were developed to be used as part of a professional development project, however, they have the potential to be adapted as a stand-alone resource or potentially with on-line or similar PD materials. This might follow the Bowland approach[2] or develop alternatives such as video presentation and discussion of materials by developers or through webinars.

Research Question 6:

Are there any patterns of differences in effectiveness for particular groups of pupils, teachers, schools, or across the three TIME teams?

In relation to the impact analysis, there are no significant patterns of differences between particular groups of pupils, teachers, and schools or across the three TIME teams.

Analysis of qualitative data indicates that some of the materials were perceived to be particularly useful for younger or relatively low attaining pupils, with the bar model and ratio tables being accessible ways to develop multiplicative reasoning. Other materials were used with high attaining pupils or adapted for KS4. Teachers who worked in schools in which there was strong senior leadership or departmental support for the programme were able to engage more fully. The project appeared more successful where school cultures supported innovation and problem solving pedagogies.

Research Question 7:

What was delivered through the programme including: activities; quality and quantity of professional development; the reach (teachers and pupils) including beyond those directly involved and impacts on organisational capacity?

---

[2] http://bowlandmaths.org.uk/

The programme delivered a total of 1,250 professional development (PD) days that were potentially accessible by participants. The quality of PD was judged to be high by participants. In addition, a set of curriculum resources and supporting documentation were produced. Materials were organised into 'lessons' and ‘units’. Each lesson could take 2-3 hours of teaching time. Thus approximately 36-54 hours' worth of high quality research-informed curriculum materials were produced. These could potentially form the basis of a coherent KS3 curriculum strand in relation to multiplicative reasoning. Approximately 60 teachers attended PD events, with a further group of 15 teachers in schools participating in Lesson Study activity. In the majority of schools materials were shared with other teachers.

Approximately 3,400 KS3 pupils were taught in classes identified with nominated core teachers in intervention schools. Approximately 2,450 of these pupils were identified through an 'on-treatment’ analysis. These 2,450 pupils had experienced a minimum level of MRP curriculum materials and were taught by teachers who had engaged in PD activity. The other 950 pupils had only experienced one or the other. Additional pupils in KS2 and KS4 experienced some of the materials.

One aspect of the project was the hosting of TIME events by Maths Hubs and the contribution of these schools in providing professional development leaders. The project provided a focus for the Maths Hubs programme and provided lessons that have informed the Maths Hub programme. Teaching schools involved in the programme have indicated they intended to provide further professional development on multiplicative reasoning.

The project also brought together curriculum developers working from different approaches to curriculum design.

Research Question 8: *Was the programme cost effective?*

Given the lack of evidence with regard to impact on pupil attainment as measured by PiM tests, a cost benefit analysis is not appropriate. As an alternative, a cost effectiveness analysis based on a comparison of the MRP with other forms of professional development was undertaken.

It found that the project appeared to be delivered in a cost effective manner. Considering the direct PD costs rather than developmental costs, the programme was cost effective in comparison with alternative PD priced as one day courses. The MRP has offered a range of added value with PD focused on a specific recognised need - multiplicative reasoning and developing research based mathematics materials.

Discussion of results

It is important to recognise limitations of the results of both the RCT and process evaluation. Schools involved in the project were not representative of the national population of schools and in the case of the process evaluation there is likely to be some sampling bias with those schools more favourable to the project being over represented.

A measured impact on pupil attainment using the PiM tests was not found, this may be due to a combination of:

- unobserved significant variables
- attrition
- sensitivity/appropriateness of the test measures
- issues related to the theory of change and/or the intervention

Most significantly, in relation to the latter, were the complexity of the project and the short time scale in which it took place.

Despite the lack of impact on pupil attainment on the PiM tests, the MRP embodied an approach that drew on the evidence base of effective professional development as well as research on multiplicative reasoning.

Characteristics of the MRP that were important were:

- a focus on subject matter
- professional learning focused on student learning
- sustained duration
- collective participation of teachers from the same school

One additional area that could have been included was to directly address teachers' beliefs about mathematics and the teaching and learning of mathematics. Definitions or frameworks could have been introduced at the start of the project and, in particular, the definition of multiplicative reasoning upon which the project was based could have been made explicit.

Important to the theory of change that underlay the project was the need to address teachers' understanding of multiplicative structures and MR pedagogy. There is the potential to develop further targeted interventions and professional development materials related to this. One approach to doing this would be to draw on international developments in relation to understanding and assessing teachers'

multiplicative reasoning knowledge[3]. This could lead to interventions that were assessed by impact on teachers in the first instance as a pre-requisite for evaluating impact on pupils.

The regional collaborative communities were led by schools and hosted by Maths Hubs. The Maths Hubs pilot (initially known as MESH) has evolved into the Maths Hub programme and this has already been funded by the DfE. The MRP evaluation provides evidence that supports this decision. Maths Hubs involved in the project, as well as other teaching schools, recognised the potential of MRP model and materials as an approach to professional development in their hubs.

A design approach to curriculum development has potential to address other areas of concern in mathematics. In addition to the curriculum developers involved in MRP there are other recent and significant examples of curriculum design in mathematics in England[4]. The NCETM has a particular role as broker in relation to different contributors to mathematics professional development, as evidenced by its approach in the MRP.

One of the effective features of the project, as identified by participants, was utilising the expertise of university-based developers and researchers. The Maths Hub programme aligns with this, with higher education institutions (HEIs) involved in strategic bodies within Hubs. The NCETM has also previously encouraged teacher innovators to draw on HEI expertise when awarding direct grants.

Lessons for evaluation design

There are lessons to be learnt for the approach to design evaluations of this type in current educational contexts, as well as in relation to practical matters. A key lesson is the importance of ensuring that any specific intervention has gone through a design and pilot stage, preferably with evidence of positive effect. Longer lead in times would help to ensure greater balance at the point of randomisation.

Longitudinal studies would allow for use of national test data rather than additional testing of pupils, so reducing the burden on schools. Trials lasting longer would also allow for impact over a longer period than nine months to be assessed.

---

[3] For example the Diagnosing Teachers' Multiplicative Reasoning Project http://www.kaputcenter.umassd.edu/projects/dtmr/

[4] For example, Bowland Mathematics, the Effecting Principled Improvement in Science and Mathematics project,the Cre8ate maths project, the NRich project; and the Nuffield applying mathematical processes.

More generally, the appropriateness of randomised controlled trials as an evaluation approach should be considered for complex PD interventions.

Recommendations

Improving pupils' multiplicative reasoning skills

The programme did not have any statistically significant impacts on general mathematical attainment as measured by PiM tests, within the timescales of the project. Given the evidence of previous studies that multiplicative reasoning can be improved through targeted intervention[5], and the possibility that the lack of impact was due to the timescale of the project, further development in this area is justified:

Recommendation 1: For the DfE to consider commissioning smaller scale quasi-experimental trials of specific MR materials or models, aiming for a more consistent approach to material use, with outcome measures designed specifically for the trial to measure MR knowledge, and to allow more time to detect impact

Recommendation 2: For the DfE to consider developing and evaluating interventions that focus specifically on teachers' multiplicative reasoning knowledge and pedagogy using specific measures of impact that are suitable for this purpose.

Recommendation 3: For the DfE to consider lessons learnt for the use of RCT and similar methodologies in relation to: pilot and development prior to testing; early appointment of external evaluators to advise on intervention design; and longer lead in times to allow for greater balance at randomisation and a longer trial period.

Impact on teacher professional development

Teachers judged both the curriculum materials and activities to be, on the whole, effective in terms of their professional development, and useful in the classroom.

Recommendation 4: That the NCETM is encouraged to refine the MRP materials for use in KS3 and potentially KS2, providing guidance on how they could be used effectively, potentially with on-line PD activities.

Recommendation 5: For the NCETM to make available MRP materials and approaches to teaching MR through the Maths Hub network.

Impact on new approaches to PD in the school-led system

---

[5] Hodgen, J. et al. (2014)Improving students' understanding of algebra and multiplicative reasoning: Did the ICCAMS intervention work? Pope, S. (Ed.) Proceedings of the 8th British Congress of Mathematics Education 2014, 167-174

Participants were generally positive about the form of collaborative professional development with teacher leaders being key to delivery and focused on issues of concern to them.

Recommendation 6: For the DfE to support collaborative professional development with features of the MRP model and similar teacher-led professional development networks through the Maths Hub programme.

Recommendation 7: That the NCETM seeks further opportunities to engage curriculum developers with Maths Hubs and other NCETM activities and potentially to develop future curriculum design projects that address the needs of teachers, schools and pupils.

Recommendation 8: For the NCETM to encourage HEI involvement in specific projects developed by Maths Hubs.

Conclusion

The Multiplicative Reasoning Project was a pilot professional development project. Given that it was a pilot and it was conducted over a short time scale it is not surprising that it did not lead to pupil impact as evidenced by external tests. The project was effective in relation to teacher professional development and supported the development of the Maths Hub programme. Those effective features of the project in relation to professional development should be developed and replicated. However, further evaluation of the specific curriculum models is needed to evaluate if they are useful more widely.

1 Introduction

The Multiplicative Reasoning Project (MRP) was commissioned by the Department for Education (DfE) to:

- improve teaching of mathematics at Key Stage 3 (KS3) in which only 38% of lessons were recently rated as good or better by OFSTED and 12% inadequate[6]
- design professional development opportunities for less experienced and non-specialist teachers who are often allocated to KS3 classes
- address the relative drop in rate of progress of English pupils compared to comparator nations with higher than average levels at end of primary school but lower than average at end of KS3[7]

The particular focus on multiplicative reasoning follows from its importance within mathematics. The multiplicative nature of the reasoning used to solve problems arising from proportional situations often involves an understanding and application in different ways of fractions as well as decimals, percentages, ratios and proportions (see Technical Report Annex A and Annex C). Together these constitute a significant part of the mathematics curriculum and, in addition, proportionality is connected to many other curriculum areas and has applications across other subjects. Pupils (and adults) often fail to move on from additive structures and this can lead to many misconceptions and errors in subsequent mathematical study.

The rationale for a project focus on multiplicative reasoning was informed by:

- the recognition for at least 30 years that pupils often fail to understand fundamental concepts of multiplication and division, ratio, proportion and rates of change[8]
- the need to reverse the trend of declining understanding of multiplicative reasoning compared to pupils 30 years ago[9]
- a view that where techniques and methods are learnt they are often misapplied[10]

---

[6] OFSTED (2012) *Made to measure: messages from inspection evidence*. OFSTED
[7] Mullis I. et al. (2012). *TIMSS 2011 International Results in Mathematics*. International Association for the Evaluation of Educational Achievement
[8] Hart, K.M. (Ed) (1981), *Children's Understanding of Mathematics*: 11-16, John Murray
[9] Brown, M. et al. (2010) Attitudes, gender and attainment: evidence from a large scale survey in England. Paper presented at the *34th Conference of the International Group for the Psychology of Mathematics Education*, Belo Horizonte, Brazil:
[10] Watson, A., Jones, K. and Pratt, D. (2013), *Key Ideas in Teaching Mathematics: Research Based Guidance for ages 9-19*, OUP

- an awareness that whilst attempts have been made to address these issues, through National Strategies and other initiatives, there has likely been little genuine progress[11]
- the belief that interventions with a coherent approach and using relevant models can increase pupils' understanding of multiplicative reasoning[12]

The MRP design drew on previous effective forms of professional development undertaken by the National Centre for Excellence in the Teaching of Mathematics (NCETM), such as Mathematics Knowledge Networks[13] and the Primary Mathematics Hosts Schools Project[14].

The MRP was delivered by the NCETM in the academic year 2013/14. MRP was a professional development project, conducted as a one year pilot. It focused on developing teachers' understanding and capacity to teach topics that involved multiplicative reasoning to KS3 pupils. The aspects of multiplicative reasoning addressed were proportional and fractional relationships and the teaching approaches encouraged were the use of mathematical models, visual approaches and problem solving strategies. Approximately 60 teachers in 30 schools engaged in three regional professional development networks, led by professional development leaders supported by university researchers. The programme delivered a total of 1,250 PD days that were potentially accessible by participants.

Specific project materials were created by a team of curriculum developers drawing on three different approaches to curriculum design. In addition to using these materials, teachers engaged in lesson study and other collaborative professional learning activities. Approximately 36-54 hours' worth of high quality research-informed curriculum materials were produced.

The following were significant design features:

a) A focus on: making connections in mathematics where the underlying structure is multiplicative; deepening the understanding of the mathematics related to solving problems where the underlying structure is multiplicative;

---

[11] Hodgen, J., Brown, M., Küchemann, D., & Coe, R. (2010). *Mathematical attainment of English secondary school students: a 30-year comparison.* Paper presented at the British Educational Research Association (BERA) Annual Conference, University of Warwick.

[12] See Increasing Student Competence and Confidence In Algebra and Multiplicative Structures (ICCAMs) project http://iccams-maths.org/

[13] Gousetti, A., Potter J. and Selwyn N., 2011 *Assessing the impact and sustainability of networks stimulated and supported by the NCETM.* London: Londonknowledgelab. Url https://www.ncetm.org.uk/files/7979131/NCETM+Network+evaluation+report+(final)+amended+300611.pdf Accessed October 2012.

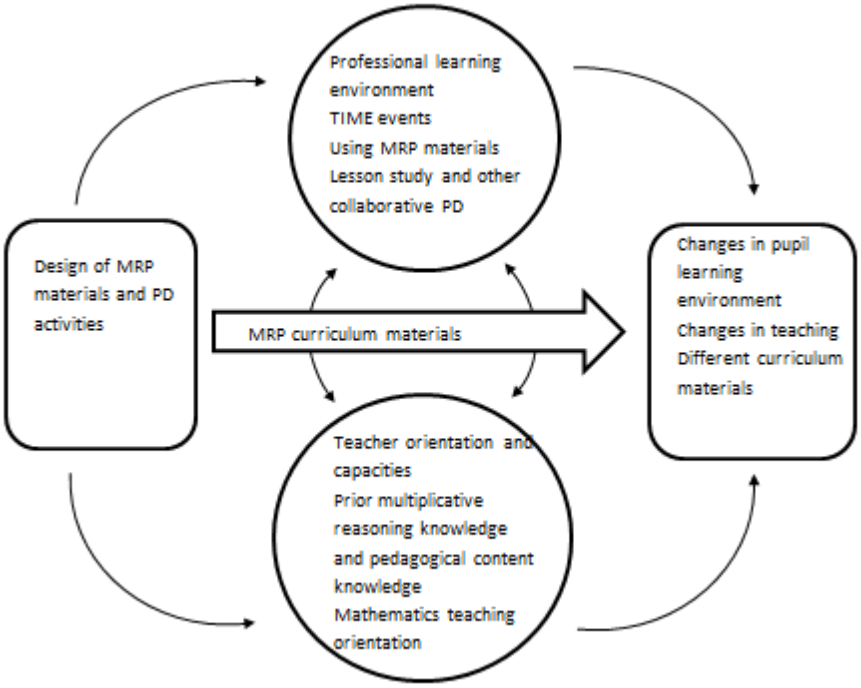
[14] Boylan, M. & MacNamara, A. (2013). *The evaluation of the NCETM primary mathematics host schools project*. CEIR

and exposing misconceptions that might underpin superficially understood procedures.

b) The use of curriculum materials that used visual images to support sense making by pupils and introducing a number of mathematical models related to multiplicative reasoning such as the bar model, the double number line and ratio tables.
c) The resources and accompanying professional development activities were created by a team of national experts to address the fundamental issues that pupils have with multiplicative reasoning and made explicit the research base in lesson commentaries.
d) The project drew on the evidence base on effective PD: a focus on pupil outcomes; sustained duration over a year with careful sequencing of learning experiences; addressing teacher subject knowledge; teachers and/or teacher leaders involved in the design and delivery of the PD supported by external expertise.
e) A collaborative approach at all levels: nationally between developers and PD leads, regionally in the TIME teams, and at school level between an experienced and less experienced teacher in each department and encouragement for teachers to engage in wider collaboration in the department.
f) The use of strategies such as lesson study that had the potential to support teachers to lead departmental training in order that successful developments could become embedded more widely. The initial intention was for the project to be departmental wide.
g) Utilising a range of different forms of expertise, within different roles (see Technical Report Section 3.3.2). PD leads had different backgrounds including practicing teachers. The curriculum developers drew on different approaches to curriculum design and approaches to mathematics pedagogy. The project structure encouraged interaction between the roles.
h) School leadership through the use of teacher PD leaders and the Maths Hubs programme with Maths Hubs lead schools hosting the PD events.
i) An attempt to effect cultural change to move from PD as being something distinct to seeing PD as part of teachers' work.
j) Internal evaluation of the pilot by the project team that built on teacher reflection on materials and learning that both support teacher development as well as providing feedback to the development team.

The intervention had both professional development and curriculum development aspects. Figure 1 provides a model of a theory of change, indicating how changes in pupil learning might occur.

Figure 1: MRP theory of Change



The theoretical basis for the project supposed that improvements in pupils' MR capacity might lead to a more general improvement in mathematical attainment.

The remainder of this report describes how the project was evaluated and findings in relation to impact and wider issues. A separate technical report provides further details of the project, evaluation, analysis and detailed findings.

2. The evaluation methodology

The Centre for Education and Inclusion Research (CEIR) was appointed to externally evaluate the project using a randomised controlled trial (RCT) methodology measuring impact on pupil outcomes and supported by a process evaluation.

2.1 Randomised Controlled Trial design and methodology

The impact of MRP on pupil outcomes was evaluated using a 3-level Clustered Randomised Control Trial (CRT) research design. The trial ran between October 2013 and June 2014 and involved 8,777 year 7 (Y7), year 8 (Y8) and year 9 (Y9) pupils (level 1) clustered into 418 Y7, Y8 and Y9 mathematics classes (level 2) clustered into 60 secondary schools.

In October 2013, 30 schools were randomly selected to become the 'intervention group' and to receive the MRP and the remaining 30 schools formed a 'business as usual' control group. Randomisation was stratified by school level GCSE attainment and geographical region. Prior to randomisation, all 60 schools provided details on two core teachers of KS3 mathematics and the pupils within the Y7, Y8 and Y9 mathematics classes that these core teachers taught. Within each of the 30 intervention group schools, these two core mathematics teachers received the Multiplicative Reasoning Project PD programme.

At baseline, the intervention group had 4,367 pupils clustered into 182 Y7, Y8 and Y9 mathematics classes clustered into 30 secondary schools. The control group had 4,410 pupils clustered into 236 Y7, Y8 and Y9 mathematics classes clustered into 30 secondary schools. At this point, pupil details were extracted from the National Pupil Database (NPD) including their attainment in mathematics at Key Stage 2, gender, free School Meals (FSM) and Special Educational Needs (SEN) status. In addition to these pupil-level details, school-level details were also compiled using the 2012/13 Annual School Census. The trial concluded in June 2014 with participating pupils being assessed using the GL Progress in Mathematics 12 (Y7), 13 (Y8) or 14 (Y9) test.

The trial adhered to CONSORT best practice RCT guidelines[15], was publically registered[16] prior to the collection of the PiM outcome data and adopted an 'intention to treat' approach. 'Intention to treat' means that all participants who were present at

[15] See CONSORT website
[16] See controlled trials website

baseline (and prior to randomisation) are included in the final analyses: pupils remain in the study regardless of whether they actually experience any of the intervention. If an intervention is observed to have a statistically significant and positive impact under these conditions it provides the strongest evidence that this impact was a result of the intervention itself and that a similarly positive impact would be likely to occur if the programme was rolled out more widely.

The final data file included complete baseline and Progress in Mathematics (PiM) test outcome data for 6,565 pupils clustered into 349 Y7, Y8 and Y9 mathematics classes clustered into 58 secondary schools - an overall response rate of 75%.

At the design stage, a power analysis[17] was undertaken and predicted that the minimum detectable effect size (MDES) for the proposed design with an 80% statistical power was estimated at between (d=) 0.22 (for the combined Y7 to Y9 sample) and 0.26 (for the Y9 sample) standard deviations. The actual achieved MDES were higher and ranged between (d=) 0.29 (Y7 to Y9 combined) and 0.48 (Y9) standard deviations with a statistical power of 80%. The reason for this unanticipated lower sensitivity was the variance structure for the outcome PiM measures and not the 25% sample attrition. From the statistical models, the proportion of the PiM variance found to be clustered at the classroom level was greater than had been predicted (over 50% in many cases). This unanticipated class-level variance structure resulted in reducing the MDES that the trial could detect with a statistical power of 80% (d=0.22 predicted, 0.29 achieved). However, the achieved sensitivity remained within the same 'moderate' range as described by the Educational Endowment Foundation (EEF), and can roughly be converted into between 3 and 4 months progress at school[18].

Variation can be observed at three hierarchical levels; school, class and individual pupil levels. Scores will differ between pupils and (average) scores will differ at both classroom and school levels. For each PiM measure, the proportion of overall variation that is found clustered at the classroom and school levels can be calculated. The stronger the clustering effect at the class / school level, the greater the proportion of variation will be held there. In other words, the more similar pupils are within classes, pupils within schools and classes within schools, the higher the proportion of variation will be present at these higher levels. The strong clustering effect observed at the class level is likely to reflect ability group setting or streaming practices common in teaching KS3 mathematics.

---

[17] See Technical Report section 5 and Annexe E for specific details on this power analysis.

[18] Higgins, S., Katsipataki, M., Coleman, R., Henderson, P., Major, L.E., & Coe, R. (2014). The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit . October 2014. London: Education Endowment Foundation.

A multilevel linear regression analysis approach was adopted to assess whether progress in mathematical attainment amongst pupils who were taught by one of the two core mathematics teachers in the 30 MRP intervention schools was significantly greater or less than pupils taught by one of the two core mathematics teachers in the 30 control schools.

A total of 12 PiM outcome variables were used; four relating to the overall mathematics attainment and eight relating to three PiM subscales constructed from test items identified (by Professor Iszák) as being related specifically to multiplicative reasoning[19].

The main analyses adopted an 'intention to treat' approach in which all pupils were included. A fidelity analysis was also conducted to explore whether engagement with MRP was statistically associated with higher PiM attainment. Following this fidelity analysis, the main analyses were repeated adopting an 'on treatment' approach. Specifically, the original 'intention to treat' sample was reduced so that all control and intervention group pupils remained within the same class they were in at the start of the trial (i.e. October 2013) and were taught solely by a core mathematics teacher who was named at the start of the trial. Additionally, within the intervention group, the sample was reduced further so that all pupils were taught by a core teacher who attended at least 2 of the TIME event days and reported to use at least some of the MRP materials with their pupils.

This 'on treatment' sample equated to 5,253 pupils clustered into 269 Y7, Y8 and Y9 classes clustered into 56 secondary schools.

Opt-out consent was sought from pupils for participation in the RCT.

2.2 Process evaluation methodology and data collection

The process evaluation was conducted through data collection at national, regional and school level.

National data collection activities were: telephone interviews (n=4) and on-going communication with key NCETM personnel, observation of an NDT meeting, and documentary analysis of NCETM documents and internal evaluation data.

[19] For more detail on the PiM outcome variables and subscales identified as being related to multiplicative reasoning, see section 3.2 below and section 5.5.3 of the Technical Report.

Regional data collection activities were: telephone interviews with professional development leads (n=3); observation of TIME team events (n=4); analysis of attendance data; and surveys of TIME team participants about school practices (125 over 3 TIME meetings).

School level data collection activities were: baseline data collection on application for 66 schools (132 teachers) but 4 schools were excluded due to not meeting necessary criteria for inclusion, meaning data on 8 teachers was not retained; visits to 9 schools and interviews with core teachers/other teachers (n=19), heads of department (n=8), senior leaders (n=5), and pupil focus group interviews containing a variety of abilities and gender (12 groups – with approximately 80 pupils in total).

Teacher level data collection included baseline data capture and exit surveys of intervention (n=21) and control school teachers' responses (n=17).

Consent was sought from interviewees.

3 Impact analysis

This section provides a brief overview of the multilevel analyses of the PiM outcome variables that were used to quantitatively assess the impact of MRP in three sections. The first section summarises the impact findings. The second summarises the analysis approach and variables used. Finally, this section explains the statistical analyses conducted as part of the impact analyses and where to find them within the Technical Report.

For much more detail on the analyses summarised here please refer to the Technical Report (section 7) and to the Technical Annex (Annex E).

3.1 Summary of the impact analysis

For both the intention to treat and the on treatment analyses, the difference between the MRP intervention and control group pupils was found to not be statistically significant. This leads to the overall conclusion that no statistical evidence has been found that the MRP teacher PD intervention had an impact on pupil outcomes in mathematics during the period of the trial (October 2013 to June 2014).

Table 1 presents a summary of the effect size statistics derived from the multilevel analyses presented in section 7 of the report for the 12 PiM outcome measures. The effect size statistics for the 'intention to treat' analyses are shown alongside those for the 'on treatment' analyses.

Table 1: Summarising effect sizes and directions for main analyses

MRP clustered RCT multilevel analyses summary

Comparing intention to treat and on treatment findings

		Year 7	Year 8	Year 9	Year 7 to Year 9 combined
PiMs Overall (General Maths Attainment)	Intention to treat	-0.02	-0.08	-0.11	-0.07
	On treatment	0.00	+0.02	0.00	-0.01
PiMs Subscale 1 (items with some relationship with multiplicative reasoning)	Intention to treat	0.00	-0.02	-0.15	n/a
	On treatment	+0.02	+0.08	-0.04	n/a
PiMs Subscale 2 (items with some relationship with multiplicative reasoning that are weakly or strongly connected to topics in MRP project)	Intention to treat	0.00	+0.01	-0.09	n/a
	On treatment	+0.01	+0.09	0.00	n/a
PiMs Subscale 3 (items with some relationship with multiplicative reasoning that are weakly or strongly connected to topics in MRP project)	Intention to treat	n/a	+0.02	-0.08	n/a
	On treatment	n/a	+0.12	-0.01	n/a

Effect sizes observed within the samples are shown to range from weakly positive (g=+ 0.12 in Y8 subscale 3, on treatment) through zero and on to weakly negative (g= - 0.15 in Y9 subscale 1, intention to treat) but across the whole of Table 1, none were found to be statistically significant.

The small effect sizes are all below the original predicted MDES estimates (0.22 to 0.26). Therefore, the higher than expected variance structure at the class-level has not impacted on the validity of these findings.

The maximum observed difference between the MRP intervention and control group pupils was observed to be an effect size below 0.15 which is a 'low' effect size that can be roughly converted to around two months of progress at school. These differences are simply too small to be able to validly conclude that the MRP had an impact on pupil mathematics outcomes during the short period of the trial.

3.2 Analysis Approach and Variables Used

A staged multilevel regression analysis approach[20] was taken for the impact analysis within the MRP evaluation. Three levels were included into the models; pupils (level 1) clustered into maths classes (level 2) which are clustered into secondary schools (level 3). Analyses were conducted using STATA and SPSS software and Excel was used to present these analyses.

The quantitative 'impact' of MRP focused on post-test (or outcome) measures from scores on the GL Assessment Progress in Mathematics (PIM) tests: PiM12 for Y7, PiM13 for Y8 and PiM14 for Y9 tests. The majority these tests took place in school on the 4th of June or within that week.

In addition to variables that identified each of the three hierarchical levels (pupil<class<school), the key 'impact' explanatory variable was located at the school level - which distinguished pupils within MRP intervention schools from those within control schools.

Analyses of the baseline sample identified some imbalance between the MRP intervention and control group samples[21] and so a further 5 school-level and 5 pupil-level explanatory variables were included into the model to help correct for this observed 'imbalance'.

To summarise, in addition to identifiers at the pupil, classroom and school levels, the following 11 explanatory variables were included into the multilevel models:

…at the **pupil-level**:

- KS2 Maths Attainment (Fine Point Score) - centred around the mean

[20] See section 7.1.3 of the Technical Report for details on this 'staged' multilevel regression approach.
[21] See section 7.2.2 of the Technical Report for details on this baseline imbalance.

- FSM status (1=FSM, 0=not FSM)
- gender (1=Female, 0=Male)
- SEN (1=SEN; 0=not SEN)
- age in months (only used for PiM subscale outcomes)[22]

…at the **school-level**:

- group (1=Intervention; 0=control)
- school GCSE attainment (taken from the 2011/12 school census; 1=above average, 0=below average)
- school admissions policy (selective = 1, not selective=0)
- school IDACI score (based on school postcode)
- school percentage of FSM (taken from the 2011/12 school census)
- school OFSTED rating (1= grade 1 or 2, 0=grade 3 or 4).

The main post-test measure used was the standardised age score (SAS) calculated from the results of the PiM test. This measure is based on the underlying raw score so that it takes into account the age of pupils and allows the score to be compared to the national average. In addition to overall attainment in mathematics captured by the PiM SAS, three PiM subscales were used that focused on items within the PiM test that were identified as being related to multiplicative reasoning. Overall and across the three year groups, a total of 12 outcome variables were used as shown below in Table 2 below.

Table 2: Summarises the GL PiM outcome variables by year group

	PiM standardised age score	MR related but not necessarily related to the MRP	MR related and weakly or strongly connected to the MRP.	MR related and strongly connected to the MRP
Year 7	PiM12	PiM12_MR	PiM12_MRW	
Year 8	PiM13	PiM13_MR	PiM13_MRW	PiM13_MRS
Year 9	PiM14	PiM14_MR	PiM14_MRW	PiM14_MRS
Year 7 to 9 combined	PiM12to14			

[22] The overall PiM attainment measure was a 'standardised age score' that used age in months of pupils to help determine the eventual attainment whilst the three MRP focused subscales were based on raw attainment without age standardisation. Therefore, for the three MRP subscales, pupil age in months was included into the group of pupil-level explanatory variables.

Essentially, for the main impact analyses there are three key outcomes (PiM12, PiM13 and PiM14) with one additional combined outcome (PiM12 to14). In addition to this, subscales of the PiM test scores were created from specific PiM test items that were identified as being related to multiplicative reasoning. The derivation of these subscales drew on the work of Andrew Izsák in consultation with NCETM[23]. No items on the PiM12 (Y7) paper were identified as being related to multiplicative reasoning that were also strongly connected to topics covered by MRP and so a PiM12_MRS outcome measure was not possible.

From the multilevel models, the coefficient of the 'group' variable was used to capture the statistical impact of the MRP programme on the 12 PiM outcome measures. These coefficients were converted into effect size statistics (hedges g - as shown in Table 1 above) with 95% confidence intervals[24].

3.3 Analyses Conducted

The impact analyses outlined in section 3.2 above were conducted using the 12 PiM post-test measures twice: first, adopting the CONSORT advised 'intention to treat' approach and second, adopting an 'on-treatment' approach[25].

Prior to the multilevel impact analyses, descriptive analyses were undertaken at baseline and outcome[26].

The decision to conduct an on-treatment analysis was taken following an analysis that explored trial fidelity. This fidelity analysis examined evidence of a relationship between engagement with MRP and attainment on the PiM post-test measures. Higher levels of engagement were found to be associated with higher attainment[27]. Additionally, both intervention and control group samples were limited so that they contained pupils who remained within the same class and taught by the core KS3 maths teacher who was named prior to randomisation.

Whilst the 'intention to treat' approach results in a higher standard of evidence it is less likely to find a statistically significant impact. This is because this approach must include all pupils / teachers in the analyses regardless of whether they experience or

---

[23] See Technical Report section 5.5.3 for more detail on this and Technical Report Annex D for the report by Professor Andrew Izsák on the PiM test items.

[24] See Technical Report section 7.3, 7.4, 7.6 & 7.7 and Annex E for details on Hedges g effect sizes and how they were calculated.

[25] These are outlined in section 2.1 above and in more detail in sections 7.1, 7.3 and 7.4 of the Technical Report and Annex E.

[26] See section 7.2 of the Technical Report for these descriptive analyses.

[27] See section 7.5 of the Technical Report for the fidelity analyses.

engaged with MRP. If an intervention is found to have a statistically significant impact under these conditions, it provides the best evidence that this is likely to happen if the intervention was rolled out more widely - because in reality pupils do move and engagement will be variable. The 'on treatment' approach does compromise on the standard of evidence but in doing so results in a higher chance of identifying a statistically significant impact. The compromise does bring some practical advantages as the 'on treatment' analyses are not distorted by things such as pupil movement or teachers who did not use the MRP materials.

For both the 'intention to treat' and 'on treatment' analyses the difference between the intervention and control group samples was not statistically significant on all 12 post-test PiM outcome measures.

4 Process evaluation findings

4.1 The project narrative: national and regional

Schools were recruited by the NCETM, with 110 expressing an initial interest and 62 recruited to the trial. Following randomisation 30 schools worked in the three 'TIME Teams'. A core development team and a national development team worked together to produce project materials.

Regional events involved discussion of issues in teaching multiplicative reasoning and research evidence, planning for teaching of MR using project materials and subsequent review and reflection and evaluation activities. TIME teams were led by PD leads who were, in most cases, currently teaching in schools. PD leads were supported by university researchers as well as the curriculum developers. The approach to professional development was flexible and responded to participant feedback. Activities included consideration of issues in relation to multiplicative reasoning, consideration and reflection on teaching materials, professional development on Lesson Study and time for joint planning.

Participants identified enabling features related to organisational aspects and specific features of professional development. Organisational aspects were identified as: good communication from the NCETM; funds to pay for supply cover; and involving two core teachers for mutual support and collaboration. Aspects of the professional development events/activities which enabled engagement and support realising objectives were identified as: the developers introducing material they had written; time for teachers to plan together; and using the same materials with different groups, enabling the teachers to understand the materials and background ideas.

Constraining features at the national level identified in case studies or through other data collection activities were identified as: issues with the organisation of the project related to it being a pilot; a lack of overview at the start of the project about the materials as a whole; not enough time in TIME meetings to look through materials and time only given at the end of the day; and the geographic spread of schools, noted by some participants as an inhibitor to networking.

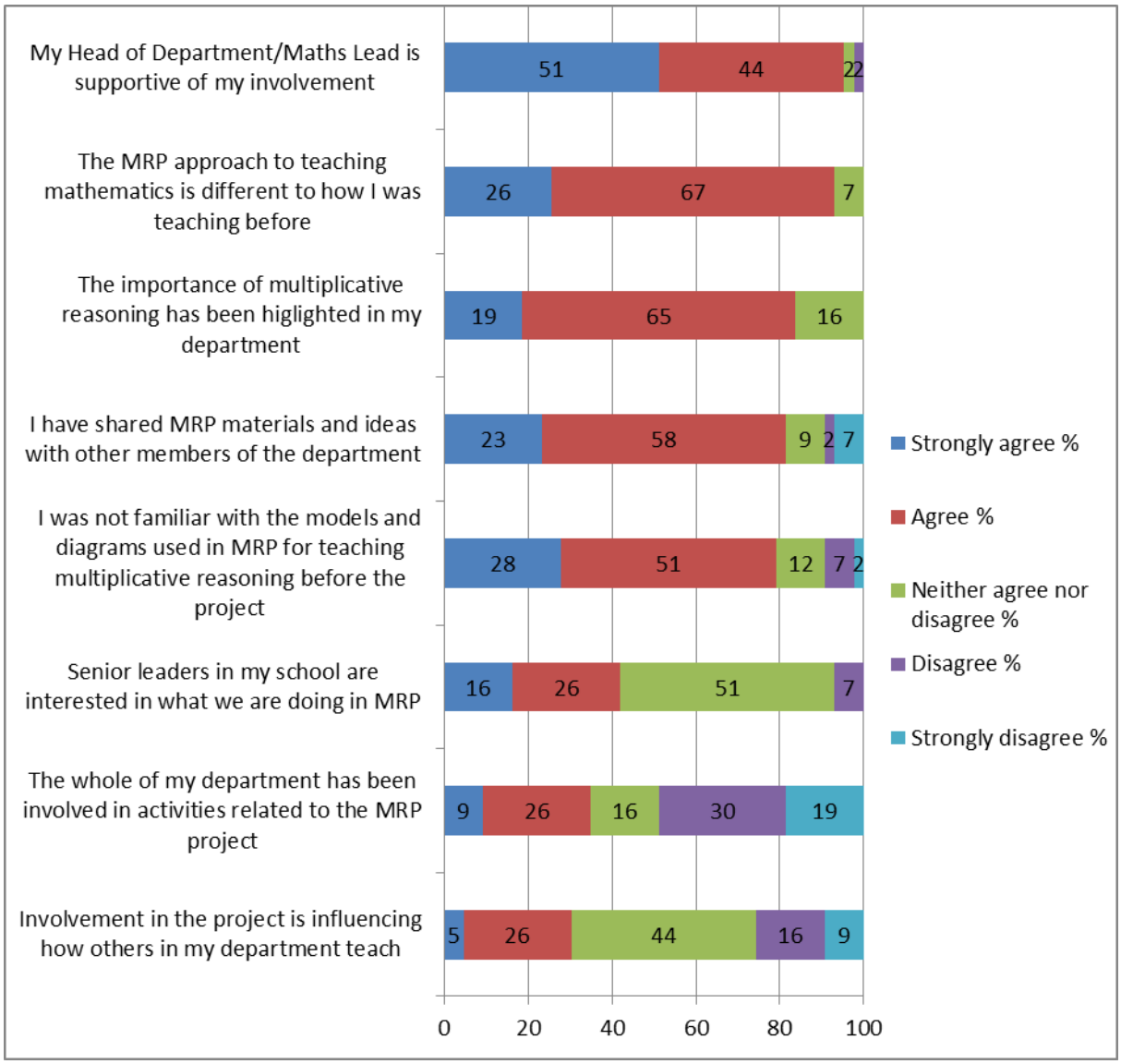
In some schools there was some variation in and changes in who were nominated as core teachers. Teachers engaged in Lesson Study activities, although this varied considerably across the three regions. There was some involvement from other teachers in the departments, generally through sharing of materials and discussion in departmental meetings.

4.1.1 School and departmental leaders - support and views on the project

4.1.2 Support of school and department leaders

At the March/April TIME session, teachers were asked to indicate how far they agreed with a set of statements about the MRP in their school.

Figure 2: Departmental relationship with the MRP



Respondents were generally positive about support for their involvement in the project and in terms of its integration into the school with nearly all (96%) respondents agreeing that their HoD/maths lead was supportive of their involvement; 84% agreeing that the importance of MR had been highlighted in their department; and 81% indicating that they had shared the materials with other members of the department. The majority indicated that the MRP approach to teaching was different

to how they had taught before (93%) whilst 79% stated that they were not familiar with the models and diagrams used in the MRP for teaching before the project. Fewer respondents (42%) agreed that senior leaders in the school were interested in what they were doing in the MRP whilst just over a third (35%) agreed that all of their department was involved in the MRP activities, with a similar proportion (31%) indicating that it had influenced the teaching of others in their department.

4.1.3 School leaders/Head of Department views

Motivations for school's getting involved

Headteachers (HTs) and Heads of Department (HoDs) were generally receptive towards being involved in the project, typically acquiescing to a request from an enthused and proactive core teacher (CT) that had identified the project themselves. Schools primary motivation for involvement was to utilise an opportunity for some of their staff to receive high quality PD. One HT stated they were reassured about being involved because of the reputation of NCETM:

> *'It's [the MRP] something that has been thought through by a professional body. A respected body'. (A13, HT).*

How readily school leaders and HoDs were to participate in the project was largely dependent on the extent to which the MRP complimented existing ways of working and/or aligned to the school's vision for the future. In the case of A46, the MRP linked to their ethos and overarching approach to maths in an especially meaningful manner. The following quotations evidence the amount of support CTs received from senior leaders to teach pupils lessons in a very deep and conceptually informed way - without undue pressures to account for short term progress.

> *Do stuff slowly, deeply, and actually understand it. Part of our philosophy for the last few years is just don't think you've got to belt through the scheme of work. It's much better to go more slowly, and this is what they're doing in maths now, they will not move on unless children are getting an understanding where they are. (A46, HT)*

> *I just think in terms of how it's changed my thinking, in that I'm convinced that we need to do more work on conception. That won't necessarily reflect in levels initially, and I don't know how we address that. I'm not concerned about it because if we explain that to HT he'll be fine, but I can't imagine that would be the same in every other school. (A46, HoD)*

By way of contrast one HoD from a different school claimed they knew of staff from other schools that wanted to apply for the MRP but their headteachers blocked them from getting involved. This was said to be particularly the case with schools that were in special measures or had lower overall attainment.

Practical arrangements

Senior leaders (SLs) were largely satisfied with the financial provision of the project to compensate for staff cover costs. However, some schools remained mindful of key teaching staff having significant chunks of time out of school. Ultimately this meant making a judgement about whether to fully commit to the project or not.

Although SLs were generally supportive and committed, there were isolated instances when sustained involvement was problematic. For example, 2 CTs explicitly mentioned difficulties in terms of convincing departmental heads to grant cover - in one instance permission to attend a TIME team event was withheld due to cover complications. One HoD (B36) lamented the amount of time required for TIME event attendance, particularly when the amount of travelling was taken into consideration. Analysis of registration sheets confirm that attendance at the CTs at TIME events was variable and tended to tail off slightly over time.

Positive features of the MRP

SLs and HoDs were collectively able to isolate a range of valued features from the MRP that had had a demonstrable impact within the school. They were largely impressed by the quality of the lesson plans and materials themselves. A recurring theme was how the project had facilitated a far more reflective and questioning approach from the pupils towards mathematics, and prompted a desire amongst them to genuinely understand the concepts that led them to certain answers.

> *[Big focus on approaches to] engender understanding, rather than just mechanically solving problems. (A13, CT1/HoD)*
>
> *When you do observations on maths and you investigate what is taking place in the classroom, the impact it is having on children just in terms of their ability to question themselves, reflect, converse with each other, develop their speaking and listening, develop their literacy… (B1, SL)*

Increased dialogue and debate was also said to have been extended amongst the staff as well as the pupils. For example, the following HT outlines how the MRP had prompted greater discussion and understanding about what pupils did and did not understand amongst the staff.

> *What the project has actually done is spurred debate and dialogue and action with maths teachers. HoD used this as part of her department meetings – she's used it with colleagues. There is in maths now a much deeper dialogue about what it is children do and don't understand. (A46, HT)*

Engagement in lesson study was a feature of the project pinpointed by certain schools as being of central importance to meaningfully uncovering how pupils were

learning. This was recognised as being quite a fundamental shift to what they would ordinarily be focusing on during lesson observations.

> *Probably the biggest difference has been through the lesson study, actually looking at lessons differently. As Head of Maths if I went into a lesson it's are they teaching the curriculum? How are they teaching it? How effective is it? How is the teacher delivering it? But to go in and start looking at it in terms of how are the children learning and how effective is it, that's a totally new slant on it, and it's something that I'm trying to get the management team to incorporate into their own lesson observations. (A13, CT2/HoD)*

A further benefit identified by one HoD was that they felt that involvement in the project had provided a 'push' for them to improve teaching more widely, getting the department to work together to discuss student learning.

> *I think what this has kick-started almost is that debate on what makes a good lesson and trying to reflect on what the pupils have learnt, where – if anywhere – there's been that first misconception...it's genuinely kick-started our ambition to share good practice and do it, not just talk about it. (B2, HoD)*

Difficulties and areas that need refinement

Although predominately positive, HoDs and SLs also identified aspects of the project that they had reservations about or that they felt needed refinement. The case studies picked up some concerns about the 'very wordy' (B1, CT1/HoD) nature of some of the materials. Certain senior staff made reference in a quite generic way about lessons being too long and that there was too much content to realistically cover. However, the HoD at school C37 conveyed some more specific concerns. For example they disliked the implications for a pupil that was absent for the introduction of a lesson that spanned a number of consecutive sessions.

> *I think our department in general is quite open to new ideas. […] We don't like things to take more than one lesson. We don't like tasks to be – not open-ended, because we like that – but we don't like it to be that if somebody misses the start they are then behind for the next lesson. So they'll [the MRP materials] probably need a lot of adjusting for us to include them in our scheme of work. (C37, HoD)*

Fundamentally the school seemed to find the very flexible nature of the MRP approaches quite challenging and perceived that the pupils themselves struggled to maintain concentration for such sustained periods of time. As such the HoD could see that the materials would be of limited use in their current format, mainly because the approach and timings did not match the department's preferred way of working.

> *We try to appeal to all the different kinds of learning. I think from a free reign, this is your task, you're going to do it for an hour, that's something completely different. We don't normally let them go for that long on the same thing. I think to start with they did struggle with that amount of concentration, so to try and keep up the same thing for an hour. (C37, HoD).*

A related anxiety to emerge was the perception that the very open nature of lessons could make certain groups of pupil i.e. the 'passive and compliant ones' vulnerable to not focusing sufficiently to permit them to achieve the intended learning objectives.

> *Occasionally it's led to lessons where there are some really engaged students, but some of them can hide. It means that the passive compliant ones can sit quietly at the back doing nothing…I think there's some work to be done there in terms of maximising it for every single student and not holding some back or pushing others forwards who don't quite understand. (B36, HoD)*

Finally, the HT from A13 highlighted the delicate balance needed to reconcile between recognising the good features of the project and 'enthusiasm' of the CTs to the project; with their duty as HT to make a sound judgement on how much to incorporate the materials into next year's schemes of work. There was a strong sense that the HT wanted more extensive dialogue with the HoD before incorporating large amounts of the MR materials into next year's schemes of work. One issue conveyed related to the fact the materials themselves needed some adaptation before they would be comfortable to roll them out further.

> *Needed refinement. The materials on the projector I felt weren't totally up to scratch that's a simple thing to resolve but at the time it was important for the kids - they weren't big enough for the kids to see and it was too woolly…the practicalities prevented I felt quality learning! Which sounds trivial but it's actually an important point. (A13, HT).*

However of greater consequence, was the HT's desire to introduce MR material gradually to avoid overwhelming or putting off wider departmental staff not involved in the MRP. Their view was that while an important area, MR was one aspect of a 'very broad' subject meaning care was needed to ensure it did not come to dominate.

> *As headteacher watching my colleagues like anything that's slightly new, unless it's handled extremely carefully it can be taken the wrong way and people are going I don't know what it's meant to do, why are we spending time on this. It's got to be introduced both gradually, carefully with all the right reasoning behind it and also from what I've seen of it, it is only one part of mathematics which is a very broad subject. It's only one part of it and you can't let this take over! (A13, HT)*

4.2 School case studies

Nine schools were visited as part of the evaluation and case studies developed. Case studies show variation in the level of engagement and support given by school and departmental leaders for the project. This had implications firstly for the outcomes for individual teachers and secondly for the extent to which the project impacted on other teachers in the department.

Most participants reported support at a departmental level for the project. However the extent of this varied. Case study analysis indicates two broad patterns of relationship between, firstly departmental support and existing practice and secondly teacher engagement. In six of schools there was broad alignment with departments and active integration. This provided an enabling environment. In three other schools there were issues that affected project implementation, influencing the potential for project aims to be met. This created a constraining environment. It should be noted that the case studies were not a representative sample and those with greater engagement in the project were over represented. These two sets of schools are now discussed; detailed case studies are provided in the separate Technical Report.

Enabling environments

In total there were six case study schools in which there was a close degree of alignment between departmental practices and/or intended approaches with the project aims and the teaching approach developed in the MRP.

In four case study schools - A13, A46, B1, and C8[28] - involvement in the project was initiated by a HoD and was viewed as part of a strategy for professional development and improving teaching and learning. In two of these schools the Head of Department was an active participant in the project as a core teacher. Reasons for involvement in these schools were strategic.

One case study school was one of the Maths Hubs and provided PD leadership. In another, mathematics practice was well developed and it was appointed as a Maths Hub during the lifetime of the project:

> *As a department we look for conceptual understanding through group work, through student talk and through developing ideas (A46, HoD)*

In other schools MRP fitted with the direction in which departments were already moving:

[28] The schools are referred to here by codes, A13, A43, A46, B1, B2, B36, C 8, C37 and C39. The letter of the code refers to the region.

> *We have been aware of the need to develop teaching approaches that engender understanding, rather than just mechanically solving problems. Anything that could help that, and the idea that if multiplicative reasoning unlocks attainment, then we wanted to be involved (A13, CT1/HoD)*

In other cases one of the core teachers approached the head of department either after seeing an advert for the project or a direct approach from the local Maths Hubs hosts or PD lead.

In these schools, the professional development approach of MRP was aligned with existing practice - with some previous experience of Lesson Study or other collaborative approaches. In a number of these schools there was a history of previous participation in research and development or similar projects within departments. There was also support from school leaders.

However, it should be noted that in these more favourable conditions the active involvement of the departments in the project was still limited during the lifetime of the project. There had been some sharing of materials and in one school all members of the department used a specific set of curriculum materials.

In summary, enabling features at school level included: positive school leadership and culture with support from school leaders; previous culture of research and innovation or similar projects; and a HoD involved as a core teacher.

Constraining environments

In three of the case study schools different forms of constraining environments were experienced that militated against the achievement of the project aims and objectives.

School A43 joined the project after a PD leader approached one of the CTs in the project. At this school three teachers participated in the various TIME events. The original plan was for CT1 and CT2 to attend all the lessons but CT1 was unable to make the first time meeting. This prompted CT1 to approach CT3 as to whether they would like to be involved instead. Having participated in the first time team event and having enjoyed it, the decision was taken for CT3 to attend subsequent time team events instead, then feedback and work as a group of three within the school.

The HoD was said to have been supportive but nevertheless CT2 and CT3 were glad there were three staff members involved as it mitigated the risk.

> *Yes, I really do think it shared the risk. Also there are three of you to go up to the Head of Department. (A43, CT3)*

The school was also said to be supportive, although on one occasion cover was not arranged and so teachers could not attend the TIME event.

CTs generally reported being supported by the SLT -

> *Just get on with it, do whatever you want to kind of thing, more than happy. (A43, CT3)*

However, despite reporting that the HoD was supportive, the CTs still hinted at an underlying unease that they had, to a large extent, left the 'normal' curriculum behind and that the accountability system of having to clearly demonstrate progress remained. This combined with the fact that MR lessons tended to focus more strongly on verbal outputs rather than written outputs, lead CTs to be mindful about being observed teaching an MR lesson for a formal lesson observation, as they were worried that it would not necessarily showcase progress due to the predominance of verbal as opposed to written outputs.

> *You completely came out of the normal curriculum map, as we have, and you were teaching something completely different. So to do that as a standalone teacher within a school, I think you'd really have to hold your nerve that you were doing the right thing and stick with it. The fact that everybody else is doing something completely different (A43: CT3)*

At school C39, the initiative to participate came from CT2 a non-specialist teacher of mathematics who saw the advert on the NCETM website. The HoD supported the teachers' applications but needed to convince senior leaders to agree to release them. The two teachers spoke at department meetings but largely undertook the project independently of the rest of the department.

> *I think because there are a lot of staff here who are nervous about losing control of the class I think, and letting them make the mistakes. We've been trying to encourage them, haven't we, particularly with the NQT. Trying to get them to let them make the mistakes and let them do it. I think they're getting there. I think some people just have their way of teaching and this isn't how they would normally teach a lesson and they were nervous to try it. (C39, CT1)*

The project appeared to have had significant impacts on the CTs' practices and thinking. They made extensive use of materials, all of which were used with at least some pupils. However, the two CTs were quite isolated. The project was accepted rather than really supported by the HoD and the school; therefore it was not clear if materials would be incorporated into schemes of work in the following year.

In contrast, at school C37, a member of an academy chain, engagement was instigated by a senior leader in a cross-institutional role in the chain. Again in contrast with the two cases described above, there was a close alignment with the pedagogical approach of the MRP and the preferred teaching approach of the chain.

This common approach was apparent as CT1 discussed professional development run by the academy chain: both the HoD and CT1 reported that pedagogical approaches used in the MRP project were not very different to those currently used in the department. However it seemed that their usual approach was more structured and they found the more flexible approaches underlying the materials challenging.

> *From a working together point of view then it's quite similar. The students are all used to that, and they are used to having quite a visual approach. We try to appeal to all the different kinds of learning. I think from a free rein, this is your task; you're going to do it for an hour, that's something completely different. We don't normally let them go for that long on the same thing. I think to start with they did struggle with that amount of concentration, so to try and keep up the same thing for an hour. I think the teacher also struggled with that because they're used to having more input than was needed. (C37, HoD).*

The department used open ended tasks but did not usually extend tasks over more than one lesson:

> *[…] We don't like things to take more than one lesson. We don't like tasks to be – not open-ended, because we like that – but we don't like it to be that if somebody misses the start they are then behind for the next lesson. So they'll [the MRP materials] probably need a lot of adjusting for us to include them in our scheme of work. (C37, HoD)*

In summary constraining factors identified in these cases or in other schools were:

- school culture and practices focused on performance, leading to pressure to demonstrate immediate progress
- school leader expectations that pupil progress required particular types of activity that were not aligned with the approach taken in the MRP
- fixed lesson structures in some schools
- time pressures including concern about curriculum coverage, lack of time for core teachers to plan together and lack of time for department to be involved

In addition a number of school organisational issues were identified: core teacher absence for extended periods; the challenge of managing two teachers being out of school on the same day; and senior leaders not organising cover.

4.3 Participant views of the project

Participants were generally positive about the project with some describing it as the best professional development they had experienced.

> *This was the most professionally organised and effective course I have ever attended, the resources were great and the organisers incredibly professional. (Survey respondent)*

This generally positive view was shared by nearly all case study interviewees:

> *I think it was extremely well run. Very effective in delivering the ideas, developing the ideas. Very inclusive. Everyone was involved – I think the group that were working at region A were fantastic, and there's already three of us committed to looking at doing this for Key Stage 2. (A 13, CT1/HoD)*
>
> *I think the project was excellent as professional development, both from a personal point of view, by talking to other teachers, but also by the ideas, the concepts that are thrown at you … thinking a little bit outside the box, (A13, CT2)*

When compared to other professional development MRP was seen very favourably:

> *In all my years, I've been teaching 24 years now, and I have been on some great training courses, and I've delivered training myself. But this by far stands out as being the best professional development I've ever had, because I've gone with a colleague, I've been able to talk it through with a colleague. (B2, CT2)*

Some interviewees did offer some criticisms regarding organisation and made suggestions as to how the programme could be enhanced. For example, some teachers would have liked an overview at the start of the project about the materials and the sequence of development of models used.

> *More of an overview at the beginning, whether that involves just reading through resources, or giving them to people. (A43, CT2)*

It is important to note that this was a pilot project and development was done in parallel with delivery and overall all CTs interviewed were very praiseworthy of the programme and keen to state how favourably it compared to other PD experienced.

> *Would recommend it to anyone to get involved in, because the actual questioning of the material I think is really, really great. It's changed my teaching. Without sounding too arrogant I am a good teacher, but I've now got better because of it. I will redesign everything probably, just to make it*

more pupil-friendly, make the material a bit nicer, but the questioning will essentially stay the same. The slides and the worksheets have got to be changed'. (A43, CT1)

4.4 MRP materials

A large number of materials were produced during the project. The core materials were a series of curriculum units focused on reasoning and making sense of fractions, understanding and identifying proportional contexts, and deepening problem solving skills through application. The table below outlines the curriculum materials produced.

Table 3: MRP curriculum materials

Unit 0	• Diagnostic and formative assessment tasks					
Unit 1	1a. Parts of a shape	1b. Pieces of a cake	1c. Fair Shares	1d. Our survey said	1e. Ordering and equivalence	1f. Milkshakes
Unit 2	2a. Contexts and the bare model	2b. Percentages on the bar model	2c. Identifying proportional scenarios	2d. Directly or indirectly proportional	2e. using the DNL to explore relations	2f. Using the DNL to solve ratio
Unit 3	3a &b Ratio tables		3c & d Using stories and diagrams to model multiplication and division		3 e & f Exploring multiplicative structures	

In addition, professional development and other materials were produced to support teaching and professional learning.

4.5 Use of materials

There was wide variation in the extent to which materials were used. The most frequently used lessons were used with 84% of classes taught by core teachers, with the least used materials used with 22% of classes[29]. There was a preference for a and b materials (see Table 3), all of which were used with more than half the classes. In case study interviews, teachers who named a preferred model stated the bar model, which underlay some of the a and b materials.

[29] Full details are given in Section 7.1.5 of the technical report.

Variations in material use were due to perceived appropriateness for particular classes, the volume of material, and - over time - an encouragement to be selective by PD leaders as well as a preference for using the materials that had been introduced more fully in TIME meetings.

From the case study interviews and surveys completed during the TIME events, there were three patterns of material use.

1. A small number of teachers used or attempted to use all the materials with at least one of their classes. The motivation was mixed, sometimes out of loyalty to the project, and more strategically, in order to trial all the materials available:

> *Yes. Basically we decided that if we were going to be part of the project we ought to do it all, so that we knew it, so that then when we want to roll it out to more teachers we know all of it and we can actually fit it to where it will work best (A13, CT1/HoD).*

2. Some teachers used some materials consistently with specific classes.

3. Other teachers selected materials on a more ad hoc basis to use with specific classes if they considered them to be suitable.

In addition to using materials with KS3 pupils, teachers in some schools also used materials with KS2 (for example in case study A13) and KS4 pupils (A43, B2 C39).

CT's appeared to vary in how much they actively adapted the lessons. CT1 (A43) for instance supplemented lessons with additional worksheets, whereas CT2 (A43) tried to keep as much as possible to what had been provided.

The extent to which materials were used decreased over time with teachers becoming more selective. In part this was due to messages at TIME events and on commentaries that teachers could be selective (see Section 8.1.3) and also teachers getting a better sense of the appropriateness of materials for different classes.

> *In the early stages, I'd probably say for all the lesson 1s and lesson 2s, we would do all of it. That meant a one hour allocation sometimes turned into five hours, which was a little bit silly. At the end, it did say on Lesson 3 you don't have to do all of them, so you can pick. If you did all of them, we'd still be teaching it now because there was something like 60 questions. So the top set I'd go for some of the harder ones. I just looked at what I thought was more important for them to be able to do. (A43, CT1)*

Teachers also adapted materials. There were two aspects of adaptation. Firstly, having decided to use a particular lesson, a teacher might select from the materials and so adapt the lesson as a whole.

I'd say I adapted them in that I certainly didn't go all the way through it, and spent a lot more time generally on probably even the start question, so even before getting onto the main part of the lesson sometimes it was a lesson purely on the discussion going on with the first question. That was the main way I adapted it really. (A46, CT1)

Secondly, teachers adapted the actual material or supplement it with other worksheets. Where materials were used with KS2 they were modified considerably for the following reasons.

- the need to cover the 'usual' curriculum and content
- the suitability of some materials for classes they taught either by attainment set or age
- using materials that had been discussed or presented in more detail in TIME sessions
- some materials were perceived to be repetitive in style or content of the previous lesson and so were potentially disengaging

One core teacher outlined the rationale for adaptation and selection that was common to many participants:

I had a bottom set year 8 with Level 2s and Level 3s, and then a set four year 9 class. Some of the earlier lessons were really good for the year 8s, because actually they didn't know what a half was, so those shapes, splitting things up, that was really useful. Sharing the sandwiches was a really good lesson for them as well. Actually moving on, I didn't use the ratio tables and the ribbon lessons with them, because they couldn't actually double or half, so it all went down at the maths and it just stopped. Whereas the other ones were more conceptual and it was more visual, and they didn't need to know the number work as such. The year 9s I used pretty much throughout. One of the ones in Unit 2 I didn't use – mainly because of time – but then I briefly used it to help them with the ratio tables, so almost condensed the two and then used it in the third unit. (A46, CT1)

Although adaptation of lessons is an issue in terms of fidelity with an RCT, the HoD, C8 makes a case for the importance of adaptation.

No, I think it's more important to still keep the adapt thing. We would refine it ourselves and say this is how we've refined it, but I don't think any of this works if you have it as a take this off the shelf sort of thing. I think everything always has to link back to your students and your school, and I think a lot of the PD would be how do you work with those materials in your school, and how are you developing, and which bits are you going to use? I'd have it as

these are lots of different ideas, but they're not necessarily going to work if you just pick them up off the shelf. (C8, HoD)

The preference for the a and b materials (and other materials) was related to the inclusion of concrete situations and contexts. In addition, there may have been an element of greater familiarity at least with the appearance of the materials, as some of the activities involved a degree of repetition of activity with a gradual increase in cognitive demand.

4.6 Participants views of the materials.

Participants viewed the project materials favourably. Survey responses were in the range of 90-100% of respondents expressing favourable views.

In the final survey participants were asked for their views of materials. The table below presents their responses.

Table 4: Participants' views of materials, n=21

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree	Don't know
The materials had lots of good ideas	57%	43%	0%	0%	0%	0%
Some of the materials worked very well in the classroom	48%	48%	0%	5%	0%	0%
The materials needed adapting to make them accessible for the KS3 pupils I teach	19%	43%	10%	19%	10%	0%
The pupils enjoyed using the materials	14%	71%	14%	0%	0%	0%
I will use the materials again with classes in the future	52%	48%	0%	0%	0%	0%
Some of the materials will be used in departmental schemes of work	48%	43%	10%	0%	0%	0%

Particular positive features highlighted were:

- range of varied, different resources
- realistic contexts
- the use of visual models

Specific materials that were favoured were those that used bar models and ratio tables.

Barriers to using the materials and suggestions for improvement were:

- the quantity of materials
- the amount of time needed for specific lessons
- lack of guidance on time needed for activities
- the amount of time needed to use the materials as a set
- the suitability of the learning demand of the materials
- the presentation of the materials
- repetitiveness and lack of explicit opportunities for groups work
- the lack of explicit formative assessment tasks related to each of the sets of curriculum materials.

4.7 Engagement in professional development activities

Attendance at TIME events

In total 84% of teachers attended at least three events, with 6% attending one event only and 10% attending two events.

Table 5:**Frequency of attendance**

Number of events	**1**	**2**	**3**	**4**	**5**	**Total**
Number of teachers attending	4	7	13	24	19	67
%	6	10	19	36	28	100

Originally there were 60 nominated core teachers, in addition to these, seven other teachers had some attendance at the events and responded to surveys. These additional staff were either there to replace the original core teacher (three original core teachers were replaced near the beginning of the programme) or as additional to the two core teachers attending. The figures given here and for the TIME survey analysis include the additional teachers. A total of seven schools sent an additional teacher to the TIME events.

Table 6 Numbers attending different events

	Number of teachers attending
October/November 2013	59
November/December 2013	57
February/March 2014	52
March/April 2014	50
June/July 2014	30

The numbers attending the final event were lower. This meeting focused on evaluation rather than consideration of new materials.

Using the diagnostic materials

A set of diagnostic materials were used widely by teachers. These were called 'Unit 0' materials. Teachers were asked to particularly focus on 4 out of the 20 tasks provided. The materials were used in a variety of ways. Some 41% of teachers videoed diagnostic interviews to support later reflection. A small number of teachers conducted the assessments in the form of a test rather than interviewing.

Activity prior to teaching materials

Teachers engaged in a variety of collaborative activities prior to teaching materials with the other core teacher. The sustained engagement by many teachers in collaboration indicates the potential for longer terms changes in professional learning practice.

Engagement in Lesson Study

In the first phase of the project, some 85% of teachers engaged in Lesson Study, with 24% engaging another member of the department who was not a core teacher. In subsequent phases Lesson Study was not strongly encouraged in one region with a consequent reduction to approximately 50% of teachers attending the fourth TIME meeting reporting that they engaged in Lesson Study.

Professional development activities with others in the department

In the first round of Lesson Study a quarter of schools involved other members of the department. After this, other members of the department were involved through briefings and sharing of materials rather than more sustained professional development. There were a small number of exceptions to this.

4.8 Quality and effectiveness of professional development activities

Participant views of the TIME events

Participants had a favourable view of TIME events. Different views were expressed about the amount of content covered in the events, with some having a favourable view but others considering that there was too much. Participants identified a number of features of the professional learning activities that they engaged with that were beneficial: discussion of lessons; time to plan with the other core teacher; input on pedagogical principles and approach; engaging in the mathematics by doing lesson activities; and general discussion with others and networking.

Table 7:View of the professional development events, n=21

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree	Don't know
The TIME team events were excellent professional development experiences	57%	33%	10%	0%	0%	0%
The presence of the university researchers and the writers of the materials did not add much to the TIME events	10%	0%	14%	33%	43%	0%
There was too much content during the sessions	0%	19%	10%	48%	24%	0%
I learnt a lot about multiplicative reasoning from attending the sessions	33%	62%	5%	0%	0%	0%
It was important to have two teachers from each school taking part	52%	24%	19%	5%	0%	0%
There was a good balance of different professional development activities in the sessions	24%	43%	29%	5%	0%	0%
The mathematical approaches were already familiar to me	0%	24%	10%	43%	24%	0%

The value of Unit 0 diagnostic assessment tasks

These tasks were seen as valuable by most participants, both for diagnostic purposes but also as valuable as learning activities for pupils. Some teachers did not

value them to the same extent but it appears they may have misunderstood the purpose, considering them as a test.

Participants' views on Lesson Study

Some participants appeared to engage in Lesson Study because they were asked to, rather than seeing it as valuable. In these cases, what was described as, 'Lesson Study' became short observations of other practices. However, other participants undertook repeated cycles of Lesson Study adopting the approach suggested in the professional development events. These teachers found the experience transformative. Positive outcomes related both to understanding of MR and the project materials but also to more generally observing colleagues in a different way.

Suggestions for improvement of professional development activities

A number of suggestions were proposed:

- create opportunities to observe teachers in other schools using materials
- smaller regional networks with TIME events closer to participants' schools
- having materials in advance
- increase opportunities for group sharing and discussion
- bringing resources together as a book
- revise the final reflection activities

4.9 Professional development impacts

Participation in the programme changed teacher beliefs and practices in several ways. The models (ratio tables, bar model, number lines) had an impact on the way teachers thought about multiplicative reasoning and on their practice more generally. Teachers felt empowered to experiment with new approaches, developing questioning skills and promoting independence. They thought more deeply about the mathematics in preparation for teaching and were more likely to ask students to justify their reasoning.

It was apparent that the use of several strategies simultaneously was important, for example using a model coupled with thinking about which questions to ask to promote understanding, or using a model and comparing strategies through discussion. Participants highlighted particular activities that led to professional learning, for example:

> *Thinking about how to phrase a question that is linked to students' depth of understanding about the ratio bar. (Exit survey)*
>
> *It's nice to have something that's quite abstract, percentages and fractions, and deal with it with a bit more of a hands-on approach in that it's discussion,*

it's comparing and building up strategies to deal with the problems. It's actually changed my teaching in that I'm now looking more at having pupils model things more using simpler numbers to check it works. (A43, CT2)

Using models: 'the bar model was a little bit of a revelation'

The use of models to support learning had an impact on many teachers, with some citing these as the most valuable part of the project:

I think the bar model – I'm not actually a maths specialist – but I've been teaching maths quite a few years now, so I've not been teacher trained, so to speak. Just the bar model was a little bit of a revelation really – something so simple and so visual and giving a context. (B1, CT2)

The ratio tables were the most successful resources for one-third of the respondents in the exit survey and nearly half linked the use of the models as leading to the most significant change in their knowledge and beliefs, as described in this illustrative quote:

The use of bar representations, ratio tables and double number lines. I have since used them with lower ability KS4 pupils and they have proved very successful. (Exit survey)

There was some evidence that the use of these models went beyond using them in relation to multiplicative reasoning with teachers generalising strategies and encouraging students to work in different ways in other topics, as in this example where the teacher discusses alternative ways to represent problems:

Drawing a picture to solve a problem rather than approaching it using traditional taught methods…. (A43, CT1)

A similar picture emerged from the exit survey, with one-third of teachers reporting the most significant change to practice as their increased use of models and diagrams, as in this example:

Using models (such as the bar model) to help pupils access work in many different topics across mathematics and become more able problem solvers. (Exit survey)

Developing questioning

The emphasis on questioning was significant for at least four of the case study schools. This was particularly the case for one teacher who was impressed by the type of questions employed within the materials:

> *Just the way it encouraged them to think, and some of the questioning. It's questions that I haven't used before, but it just probed for much better, deeper understanding of the topic. (A43, CT1)*

Promoting independence

Discussion was one aspect of increased student participation in lessons, typically requiring students to justify their responses:

> *Yes, I'm much keener now to say, 'Right, you've got the right answer, now tell me why.' Because they've got a standard method that works, but they can't tell me why. When they show a bar, they can, they can talk it through, which is really good. (B36, CT1)*

One teacher spoke of learning about collaborative learning and managing discussions - *'also standing back and not interfering. Teaching in a more interactive way, managing dialogue'* (C8, CT1). For another, encouraging discussion involved *'more standing back, promoting resilience and independence'* (C39, CT1). This resulted in a shift in practice:

> *I think it's changed my approach quite a bit. I'm less likely to automatically help someone. I will take a step back and think about it first. Try it, write down how you tried it[…]. You need to show me and explain how you got to where you are. I'm encouraging them to be more independent and speak to other people before they speak to me. (C39, CT1)*

One respondent to the exit survey found the most significant change to practice was:

> *Thinking about how to phrase a question that is linked to students' depth of understanding about the ratio bar. Letting the students have a go ! Allowing them more thinking time without getting impatient to intervene! (Exit survey)*

Focussing on student learning

Teachers noted a change in perspective when engaged with programme activities focussing their attention on student learning, which was particularly so when engaged in Lesson Study.

Although organisational issues meant that only two of the case study schools (two-thirds of exit survey schools) were able to do more than one Lesson Study, the schools that did implement it as intended found it of huge benefit. The opportunity to work closely with another teacher to plan, teach, observe and discuss a lesson in depth refocused the teachers' attention away from their own actions and on to student learning:

> *It was very valuable, I think, for both of us to watch the other teach, but it wasn't about looking at what they do. It was very much how are the kids responding to this? Looking at the thought processes of the children, we were having light bulb moments, 'Ah, you're struggling with that because this is missing,' or whatever. It was really good to be able to focus on the children, let somebody else do the work almost and focus on what the children were doing and what their understanding was and where they were getting things muddled, or where they were flying. Sometimes they flew and it was lovely to see. (A13, CT2)*

Working together using a Lesson Study approach appeared to help both core teachers have the confidence to make decisions, and to work more flexibly with the lesson plans. One teacher identified several ways that engaging in the lesson study had changed her practice, as it drew her focus of attention to what the students were learning rather than what they were doing, and supported her to reflect more deeply on her own lesson plans:

> *It's changed how I'm approaching planning now, because I'm starting to think okay, what's the really important bit that I want them to understand here? How do I best go about them understanding it? Rather than thinking this is my learning objective, let's do this activity, this activity and this activity. (B2, CT1)*

'Permission' to experiment

Several teachers felt that they had been given permission to teach in a different way to usual, liberating them from perceived constraints on their practice:

> *I think I'm happier to spend more time making sure they understand, rather than just moving on and moving on and moving on. Because I've almost had the flexibility and the permission of saying, no, we need you to teach these lessons because we've had to do it for the project. I need to take the time to do this, and therefore I'm going to do it and just see how it goes. (B2, CT1)*

One teacher said that although this freedom was appreciated it was not an approach that could be sustained or tolerated by existing accountability structures:

> *It was refreshing, but as to whether I would – well, I certainly wouldn't do it every time because we just can't. (A43: CT3).*

Questioning assumptions

A minority of teachers recognised that participating in the programme had challenged their existing practice, making them reconsider traditional methods that they had used or question assumptions they had made:

> *Because it's quite new I've had to really think about what I'm doing, and it's been quite nice for me actually to have to really think. I've taught reverse percentage problems for 20 years unsuccessfully with a lot of kids, because I've got a method that I think works and they've either applied the method or they didn't understand. To see this bar in context has been really good, so it's been really refreshing to see those sorts of things. (B36, CT1)*

This challenge was also evident at one TIME meeting where teachers questioned their existing understanding of what teaching was:

> *We need to change our idea of what teaching is. Teaching goes into preparation rather than delivery' and that ' lessons were more successful [using the project lessons] – students have more control over what they were doing and the real life context was important (Region B TIME meeting).*

Knowledge of MR and MR pedagogy

Many participants reported an increased awareness of the importance of multiplicative reasoning and its relationship to other areas of mathematics. Of respondents in the exit survey, 20 of the 21 teachers responding agreed or strongly agreed with the statement 'I learnt a lot about multiplicative reasoning from attending the sessions'. One respondent added: *'I now believe that MR is a vital tool in developing successful mathematicians in school. It aids understanding and offers a range of approaches to tackle tasks*'.

Several teachers discussed students' struggle to make the transition from working additively, for example:

> *I think I've begun to realise that I think one of the reasons that kids struggle so much is that they haven't moved away from working additively'. […] I think when we started to look at the start of the project and they'd highlight all the things in the new curriculum that were multiplicative based, I just thought it's everywhere. It's no wonder that the kids who don't understand what they're doing when they multiply, and they can't work with groups of or lots of, then they can't see division as grouping, and all these sorts of things. […] It's almost like you're putting the foundations on sand. It's just getting washed away every year. (B2, CT1)*

Leadership capacity

There was evidence of the development of the core teachers' leadership capacity in two-thirds of the case study schools. In the exit survey, all but one respondent reported leading professional development activity ranging from discussions in department meetings to focussed professional development sessions.

Many participants shared materials with other members of their department. However, opportunities to lead more extended professional development activity were limited. Nevertheless approximately a quarter of participants involved another member of the department in the first Lesson Study activity. It appeared the project developed a desire to lead professional development.

> *We've just done the one twilight, but obviously my department plan is going to be based around it next year in order to do the lesson studies with everyone in the department. We've also done, CT1 and I have been to the junior school and done a transition lesson. (A46, HoD)*

In addition to learning from the way the TIME events were run, the sustained nature of the programme allowed teachers to reflect on their attempts to share their learning and adjust their strategies as the following example shows. In some cases, first attempts to share the project experience with the department were met with initial enthusiasm but this was quickly lost:

> *We showed them the sandwich lesson, I think it was 2a, and said, 'Look, just try this lesson.' We got them to try it in a departmental meeting, just with whiteboards. People went, 'Yeah, yeah, sounds good,' and then it kind of disappeared again. (B2, CT?)*

CT1 recognised that there needed to be sustained support for teachers in the department to work with the project materials, proposing to build on the model of professional development she had experienced through the project:

> *I think we've almost got to try and emulate what we had, which is the constant drip of we've taught this – what happened? Let's talk about it, re-evaluate, teach it again. I think it's that bit that is the most powerful. It's not the fact that these are fix-it-all lessons – they're brilliant. It's the dialogue and it's the team planning as well. (B2, CT1)*

Two teachers felt that the programme was more likely to impact on the work of the department if the Head of Department was one of the core teachers:

> *I think the fact that [name] was there, the head of department, I think that had a big impact. We were both quite enthused about it and therefore she could make the decision, so for example, the last couple of weeks what we've done is concentrated on the multiplicative reasoning tools that we've learnt and cascaded it out to the rest of the department. (B1, CT2)*

Even when a teacher had been the sole participant from a school for much of the project there was evidence of an increased understanding of how to lead professional learning even though the actions taken were limited.

Participation in the programme benefitted teachers with a variety of experience, including experienced teachers in leadership roles. However, the extent of development of leadership capacity was limited by the informal nature or short term nature of the activities that involved other members of the department.

4.10 Pupil impacts

Participants were positive about the impacts on pupils believing that pupils enjoyed the lessons; had increased understanding of MR; were more able to problem solve and do open tasks, and were more resilient. The lessons were seen as prompting discussion and engagement. Pupils in focus groups echoed these themes and were positive about MRP lessons.

The extent varied to which MRP lessons supported improved relationships to mathematics. In schools where problem solving and discussion was already a feature, MRP tended to enhance this. In other schools with different cultures, pupils found the problem-solving nature of the tasks challenging. Pupils in focus groups valued the contextualised nature of the tasks. This positive aspect was strengthened when teachers developed this further, for example one teacher provided ingredients to allow milkshakes to actually be made to build on a curriculum resource. The use of different methods and accessibility of materials was beneficial to pupils.

However, some teachers considered some of the materials repetitive and boring and this was echoed by some pupils.

4.11 School and departmental leaders support and views on the project

Positive features of the project were increased discussion amongst pupils and a focus on understanding. In addition the project had positive impacts in terms of developing dialogue in departments about mathematics teaching and learning. Lesson study was viewed favourably by Heads of Departments interviewed. Criticisms of aspects of materials or the project made by some were: wordy materials; activities that took longer than a single timetabled lesson; open problems might not lead to learning objectives being met; and the quality of the presentation of some materials.

4.12 Organisational and network impacts

One-third of case study interviewees suggested that the impact on their department would be limited. This was the case where there had been a lack of involvement of the department. In another third there were plans to integrate materials into a

scheme of work. In the 3 case study schools that were teaching schools, there were plans to share materials more widely.

The networking dimension was seen as a positive aspect of the project. However, the geographical spread of schools meant that the TIME teams, as a whole, are unlikely to continue as networks.

4.13 Control School Approach

Control school were surveyed to determine if there had been any significant difference to their approach to teaching as a result of involvement in the project. No differences were ascertained.

5. Security of findings

The evidence from the process evaluation and analysis conducted as part of the RCT allow for the security of findings to be rated. Five features are relevant: design, power, attrition, balance and threats to validity. Each of these aspects can be ranked on a scale of 0-5, so developing an overall description of the security of the trial.

The table below provides a summary of the security rating. The overall security is the minimum level on any of the features. Due to the variation in how the project was conducted at regional and school level there are threats to validity (see Technical report Section 9, for more details). The shading indicates the assessment of security in relation to each feature.

1.1.1.1.1.1 Table 8: Summary of security rating[30]

Rating	1. Design	2. Power (MDES)	3. Attrition	4. Balance	5. Threats to validity
5	Fair and clear experimental design (RCT)	< 0.2	< 10%	Well-balanced on observables	No threats to validity
4	Fair and clear experimental design (RCT, RDD)	< 0.3	< 20%		
3	Well-matched comparison (quasi-experiment)	< 0.4	< 30%		
2	Matched comparison (quasi-experiment)	< 0.5	< 40%		
1	Comparison group with poor or no matching	< 0.6	< 50%		
0	No comparator	> 0.6	> 50%	Imbalanced on observables	Significant threats

Thus we propose an overall security rating of 2. This would suggest that repeating MRP in its current form is likely lead to a similar outcome in terms of impact.

[30] EEF (2014) Classifying the security of EEF findings: Note this was developed in collaboration with Stephen Gorrard and Steve Higgins.

However, it does not necessarily mean that elements of MRP could not have an impact or if there was greater fidelity that this might not lead to impact.

6. Cost effectiveness

Given the lack of evidence with regard to impact on pupil attainment as measured by PiM tests, a cost benefit analysis is not appropriate. As an alternative, this section provides a cost effectiveness analysis based on a comparison of the MRP with other forms of professional development. The basis for this is that the MRP was judged by participants to be effective professional development.

The cost of the project in delivering the professional development was £12,821 for each hub. This included costs for development days involving PD leads, a university expert, a university research assistant, a curriculum/resource developer and Maths Hubs co-ordination. The total cost across all three hubs in relation to professional development delivery was £38,462. (Note: we have excluded costs in relation to honoraria to participating schools - since this is a direct payment to schools involved - and central costs for Maths Hubs development, resource development and central NCETM costs from this calculation, since these are developmental not delivery costs).

The number of participants gaining support from the programme is estimated - using NCETM data - at 69 (the number of teachers that attended at least one TIME Team event), representing 31 schools. These staff received an estimated total of 1,250 PD days (calculated by multiplying the number of TIME Team events attended by these teachers by 5 representing at an average 5 hours for each event).

Using the same methodology as used in a previous study (the evaluation of the NCETM Primary Mathematics Host Schools Project - PMHSP) this allows the proportional cost per person and per hour of PD to be calculated. However, it should be noted that in the NCETM Primary Mathematics Host Schools Project we were able to calculate estimates of additional school based PD. Doing the same for MRP would increase the number of PD hours and so reduce the cost per hour. On the other hand in the Primary Mathematics Host Schools Project some payments to schools were included in the costs calculation.

Costs are compared below with the cost of a typical one day PD course with a private provider of £250 (note this does not include possible supply costs), and the cost of the MRP.

Table 9: Comparison of MRP costs and other PD

	Per participant	Per PD hour
Multiplicative Reasoning Project	£557	£31
Primary Mathematics Host Schools Project	£227	£15
One day course	£250	£40

These figures indicate that the MRP project cost less per hour compared with alternative PD, but more per hour than PMHSP. It must be borne in mind of course that the MRP has offered a range of added value with professional development focused on a specific recognised need - multiplicative reasoning and developing research based mathematics materials.

7. Interpretation of results

7.1 Limitations

In interpreting the results of both the RCT and the process evaluation, it is important to recognise the limitations of the study. From the analysis of the sample, the schools involved were not representative of schools nationally, with a greater number of high attaining schools than in the national population and so caution is advised if inferring the findings from the sample of 60 schools more widely. RCT designs commonly pay less attention to external validity (generalisability), particularly at the pilot stage. Instead, RCTs focus on maximising the internal validity of a research design, taking the 'counterfactual'[31] into account through the use of a control group.

There were clearly variations in engagement by schools in MRP. The respondents to the exit survey and those teachers interviewed in the case studies are likely to represent teachers that are more favourably disposed to the project.

7.2 Possible explanations for the impact results

In this section we consider a number of possible explanations for the impact results.

Unobserved significant variables

In a trial of this sort there are necessarily issues in relation to balance. Ideally, the control and intervention group samples will closely mirror each other in every way. This way, the assumption that the control group can capture the 'counterfactual' has greatest credence. This was a *clustered* randomised control trial that randomised at the school level. Despite the use of a school-level stratification scheme, this trial found notable differences between the control and intervention group samples at baseline. This was most evident with respect to attainment in KS2 maths but also found in terms of SEN, FSM and gender make up. These imbalances were taken into account within the resulting multilevel analyses through the inclusion of explanatory variables - but evidence of imbalance does weaken the validity of the findings. Imbalances were also found at the school level, the most striking of which was that three schools that had selective admissions policies were placed within the control group sample. A school-level dummy explanatory variable was used to try to

[31] The counterfactual relates to change over time that would have happened anyway - regardless of the existence of a specific educational intervention (such as MRP). The control group captures this 'counterfactual' and in doing so, allows changes observed in the intervention group to be seen as larger, smaller or equal to what would be expected regardless of the intervention.

disentangle any potential impact that a policy (and pupil population) difference resulting from academic selection might have on the control group being able to capture the counterfactual. In other words, the impact of 'selection' was separated from the impact of being in the control group (imbalance is discussed in detail in the Technical Report section 7.2.2). However, we do not know is if there were significant imbalances on other (unmeasured) factors. For example we know that the randomisation led to three grammar schools being in the control group.

There may be other unmeasured school context differences that are relevant. For example, the analysis shows that there is a strong correlation between prior attainment and PIM outcomes. There also appears to be some evidence that there is an effect of setting in mathematics. Most pupils in the study were grouped into relatively narrow attainment groups. Previous research that examined progress in mathematics in KS4[32] compared students with the same KS3 scores who were placed into different sets. Students placed in higher sets averaged between a third and half a GCSE grade higher than those in a lower set who had the same KS3 scores. In region B in year 8, their sample was unbalanced with control schools having a greater number of sets with pupils with relatively higher attainment. It may be that a linear model of progression is not appropriate to a situation in which pupils are taught in sets.

Attrition

Whilst attrition led to some further imbalance, and this has been taken into account, it may be possible that it led to greater imbalance on other factors that were significant but not measured. However, it should be noted that our analysis of reasons given by schools for attrition does not suggest differences between intervention and control schools.

The test measure

The PIM test was chosen as a test of general maths with some MR items. In the event the PIM11 did not have any closely matched items to the MRP materials and the match in other tests was relatively low. In addition, PIM was not a particularly good test of problem solving, which was identified as a positive outcome in the process evaluation.

[32] Wiliam, D. & Bartholomew, H. (2004). It's not which school but which set you're in that matters: the influence of ability grouping practices on student progress in mathematics. British Educational Research Journal, 30(2), 279-293

Issues with the theory of change/intervention

It is more probable that the explanation for the 'no impact' findings lies in the intervention itself rather than how the trial was conducted. There are number of possible explanations which are not mutually exclusive.

1. The complexity of the project. The project aimed to change practice in a number of different ways and through different mechanisms. These features included: problem solving, Lesson Study plus MR itself. It is conceivable that some features may have promoted a positive impact whilst others had a negative effect. The overall effect may have then been neutral.

2. The short time scale of the intervention. There were seven months or less between first TIME meetings and the test date. This is a relatively short time scale given the project had multiple aspects and was aimed at effecting substantial change in practice. As Professor Izsák points out (Technical report annex D), it make take more than a year of sustained engagement with different ways of teaching for practice to change.

3. Implementation dip. Related to the two points above, in research on school and teacher change it has been found that there can be an implementation dip. Initially results can decline before rising to a level above the original level. It might be expected that any positive effects would take time to work through in practice.

4. Related to 1, the project was conducted as a pilot with materials and activities being developed during the project. Given this was a pilot, materials were being trialled during the project. The materials and models used may not have been effective when presented in this form, even though the models and principles themselves had the potential to be effective.

5. There were wide variations in teacher use of materials and engagement. Where engagement and use was low, impact could not be expected.

6. Conversely, some teachers made very extensive use of the materials. This meant that other aspects of mathematics curriculum were not covered and this may have impacted on test results given it was a general test of mathematics with many items unrelated to MR.

7. The professional development may not have been effective in this form.

8. The intended teaching approaches were enacted but were ineffective.

7.3 The MRP as an approach to professional development: improvements and scalability

The MRP was an approach to professional development that drew on the evidence base on effective professional development as well as research on multiplicative reasoning (see Technical Report Annex A and C). Characteristics of the MRP that were important were:

- a focus on subject matter
- professional learning focused on student learning
- sustained duration
- collective participation of teachers from the same school

One additional area highlighted by Professor Izsák in relation to the MRP was a professional development approach to MR pedagogy. This is in relation to directly addressing teachers' beliefs about mathematics and about the teaching and learning of mathematics.

From the evaluation evidence, it is clear that many teachers' beliefs did change as a result of engagement in the MRP. In particular, there was greater emphasis placed on the importance of discussion, problem solving, the use of models, and independent learning. Different pedagogical approaches were discussed, but solely in relation to specific curriculum materials. Addressing wider issues of how mathematics is taught may have supported teachers to address or at least make sense of dissonance between MRP and some of the existing school cultures.

In addition, Professor Izsák echoes some of the participants in suggesting the need for clearer definitions or frameworks at the start of the project and in particular for the definition of multiplicative reasoning upon which the project was based.

The MRP project required significant resource for the development of materials and to support the Maths Hubs infrastructure. It would be costly and potentially impractical for developers to work with similar teacher network directly on a large scale. However, even within the pilot MRP, participants were not able to work with every developer directly. Internet technology could potentially be used for live or asynchronous interaction with developers or video materials could be developed. A further extension would be for developers to work with teachers as co-authors of materials, an approach that has been used in a recent curriculum development project[33].

A project with all the features of MRP may not be scalable. However, support could be given to collaborative professional development communities based on the TIME

[33] See http://cre8atemaths.org.uk/

model or different forms of teacher networks using specific curriculum materials in relation to MR or other areas of the curriculum. There are examples in both the recent and medium term past of this being done successfully.

7.4 Lessons for evaluation design

There are lessons for both principles of design of evaluations of this type and in relation to practical matters.

Randomised controlled trials are relatively expensive and can place burden on the delivery of interventions, as well as on schools, teachers and pupils. RCTs aim to establish causation where evidence has already been established that a specific intervention is associated with desired outcomes. There are a number of principles to consider when designing trials of this sort:

- ensuring that any specific intervention has gone through a design and pilot stage, preferably with evidence of positive effect
- appointment of independent evaluators before finalising the design of the intervention to advise on compatibility with an RCT approach particularly in relation to securing high levels of fidelity
- ensuring that interventions to be tested by RCTs have specific components that can be isolated as causal factors
- considering designs which gather evidence on intermediate change mechanisms (for example in the case of MRP, using scales that can assess teachers multiplicative reasoning)
- RCTs and in quasi-experimental designs to have longer lead in times so randomisation can take into account a wider range of relevant factors, including retrieved pupil data, school characteristics and, in the case of secondary schools, balance in terms of setting patterns
- consider payments to control schools or wait list designs to support participation
- consider longitudinal studies that can draw on national test data rather than specific testing of pupils.

The issue about pilot and design are particularly important in relation to professional development interventions. This is because of the tension between requirements of an RCT - consistency and high fidelity - and what constitutes effective professional development - adaptability, responsiveness to participants, and support for teacher autonomy. It is noteworthy that these were features that were valued by participants and linked by them to professional learning.

In relation to practical issues the following are issues to consider:

- to reduce burden on schools, appropriate resources should be allocated in relation to the testing of pupils, including appropriate number of invigilators
- incentivising schools in relation to return of information
- time trials to avoid testing during examination periods and allow for more flexibility on test dates

8. Outcomes in relation to research questions

In this section we consider outcomes in relation to each research question

Research Question 1

What is the impact of the programme on pupil outcomes on both general mathematical attainment as measured by GL assessment Progress in Mathematics (PiM) tests and on those items in the GL PiM associated specifically with multiplicative reasoning?

During the timescale of the trial (13 October 2014 to May 2015) the programme did not have any statistically significant impacts on general mathematical attainment as measured by PiM tests or on items on the tests specifically associated with multiplicative reasoning.

Research Question 2

What are the impacts (if any) on: pupils' relationships to mathematics; teacher beliefs and practice including on lesson planning; teacher knowledge of multiplicative reasoning pedagogy; capacity of core teachers' to lead professional development?

The project had a positive impact on pupils' relationship with mathematics as reported by teachers in surveys and interviews and pupils in focus group interviews.

A range of changes in teacher beliefs and practices were identified:

- thinking more deeply about mathematics prior to teaching
- using models when teaching MR and extending models to other areas
- developing questioning
- promoting independence
- focussing on student learning
- questioning assumptions

Many participants, who responded to surveys or participated in case study visits, reported an increased awareness of the importance of multiplicative reasoning and its relationship to other areas of mathematics. Some discussed specific issues about the relationship between additive and multiplicative understanding.

With regard to core teachers leading professional development, many participants shared materials with other members of their department. However, opportunities to lead more extended professional development activities were limited. Nevertheless approximately a quarter of participants involved another member of the department in the first Lesson Study activity. There was some evidence of development of leadership capacity and it appeared the project developed a desire to lead professional development in others.

Research Question 3:

How was the programme conducted and how did this differ from the planned programme in what way and why?

The programme was conducted as planned at national and regional levels involving collaboration between NCETM leads, curriculum developers, university researchers and professional development (PD) leads. At school level the programme was implemented by two core teachers in each school. Originally it was planned to have a stronger focus on core teachers leading PD in their own departments. During the recruitment phase this was judged by the NCETM and Department for Education (DfE) to be more appropriately something to be encouraged rather than required. There was a degree of loss of fidelity at regional level in relation to the extent to which materials were a resource to be drawn on or a recommended coherent curriculum. The extent to which teachers were able to fully participate in the programme depended on support in schools. In some schools the project was actively supported by senior and departmental leadership and there was a good fit with departmental approaches to mathematics teachers and/or strategic developments in the subject. In other schools there were various issues that meant alignment was lower and this created constraints on implementation.

Research Question 4

What are the views of teachers/development teams on the programme including its effectiveness?

Participants responding to surveys or participating in interviews judged both the curriculum materials and professional development activities to be, on the whole, effective. Aspects of the professional development activities that were viewed positively were:

- discussion of lessons
- time to plan with the other core teacher
- input on the pedagogical approach
- engaging in mathematics by trialling lesson activities

Aspects of the materials and resources that were produced that were viewed positively were the variety of materials; the use of realistic contexts, and visual models.

Issues were raised for improvement such as having the materials in advance of the training events and having the events nearer to the participants' schools. However, these may be related to the pilot nature of the project.

Research Question 5:

If/how was the programme effective and what lessons can be learnt for scalability?

The programme was effective as a curriculum design project and pilot project for teacher professional development. It did not meet the intention to impact on pupil attainment (which was measured by the GL PiM tests within the timespan of the project).

Given this, there is no rationale, in terms of an intention to impact on pupil attainment in the short term, for repeating the project in the same form and for the same duration with a larger number of schools. With further development or refinement a similar approach might lead to impact on those aspects of pupil learning measured by PiM tests. One possibility

would be to increase the project length to allow time for professional learning to impact on pupil learning and then to re-evaluate.

A project with all the features of the MRP may not be scalable. For example, a positive feature was the opportunity for face to face dialogue between teachers and curriculum developers. This may difficult to repeat at scale. However, support could be given to collaborative professional development communities based on the TIME model or different forms of teacher networks using specific curriculum materials in relation to MR or other areas of the curriculum. There are examples in both recent and medium term past of this being done successfully.

Features of the project that were found to be effective and are reproducible at scale in future projects are:

- teacher professional development focused on research-informed curriculum materials
- teacher collaborative learning communities that draw on the expertise of teacher leaders, curriculum developer and researchers

The MRP brought together a team of curriculum developers whose designs were informed by different but complementary pedagogical principles. Developments were supported by other university researchers and teacher professional development leaders. This was recognised by the national development team members, including teacher leaders, as a powerful and productive approach. Teachers interviewed also valued the different contributions.

The project aimed to address concerns about pupil understanding of multiplicative reasoning in KS3. Teachers found the models used in the projects as useful approaches to teaching multiplicative reasoning. Given this, there is the potential to evaluate specific materials in a more focused way - for example using the bar model and ratio table. This would then allow for an assessment of whether these models should be adopted more widely.

The project materials were found to be a useful resource for KS3 and potentially KS2 curriculum. They were developed to be used as part of a professional development project, however, they have the potential to be adapted as a stand-alone resource or potentially with on-line or similar PD materials. This might follow the Bowland approach[34] or develop alternatives such as video presentation and discussion of materials by developers or through webinars.

[34] http://bowlandmaths.org.uk/

Research Question 6:

Are there any patterns of differences in effectiveness for particular groups of pupils, teachers, schools, or across the three TIME teams?

In relation to the impact analysis, there are no significant patterns of differences between particular groups of pupils, teachers, and schools or across the three TIME teams.

Analysis of qualitative data indicates that some of the materials were perceived to be particularly useful for younger or relatively low attaining pupils, with the bar model and ratio tables being accessible ways to develop multiplicative reasoning. Other materials were used with high attaining pupils or adapted for KS4. Teachers who worked in schools in which there was strong senior leadership or departmental support for the programme were able to engage more fully. The project appeared more successful where school cultures supported innovation and problem solving pedagogies.

Research Question 7:

What was delivered through the programme including: activities; quality and quantity of professional development; the reach (teachers and pupils) including beyond those directly involved and impacts on organisational capacity?

The programme delivered a total of 1,250 professional development (PD) days that were potentially accessible by participants. The quality of PD was judged to be high by participants. In addition, a set of curriculum resources and supporting documentation were produced. Materials were organised into 'lessons' and ‘units’. Each lesson could take 2-3 hours of teaching time. Thus approximately 36-54 hours' worth of high quality research-informed curriculum materials were produced. These could potentially form the basis of a coherent KS3 curriculum strand in relation to multiplicative reasoning. Approximately 60 teachers attended PD events, with a further group of 15 teachers in schools participating in Lesson Study activity. In the majority of schools materials were shared with other teachers.

Approximately 3,400 KS3 pupils were taught in classes identified with nominated core teachers in intervention schools. Approximately 2,450 of these pupils were identified through an 'on-treatment’ analysis. These 2,450 pupils had experienced a minimum level of MRP curriculum materials and were taught by teachers who had engaged in PD activity. The other 950 pupils had only experienced one or the other. Additional pupils in KS2 and KS4 experienced some of the materials.

One aspect of the project was the hosting of TIME events by Maths Hubs and the contribution of these schools in providing professional development leaders. The project provided a focus for the Maths Hubs programme and provided lessons that have informed the Maths Hub programme. Teaching schools involved in the programme have indicated they intended to provide further professional development on multiplicative reasoning.

The project also brought together curriculum developers working from different approaches to curriculum design.

Research Question 8: *Was the programme cost effective?*

Given the lack of evidence with regard to impact on pupil attainment as measured by PiM tests, a cost benefit analysis is not appropriate. As an alternative, a cost effectiveness analysis based on a comparison of the MRP with other forms of professional development was undertaken.

It found that the project appeared to be delivered in a cost effective manner. Considering the direct PD costs rather than developmental costs, the programme was cost effective in comparison with alternative PD priced as one day courses. The MRP has offered a range of added value with PD focused on a specific recognised need - multiplicative reasoning and developing research based mathematics materials.

9. Recommendations and conclusion

9.1 Improving pupils' multiplicative reasoning skills

The programme did not have any statistically significant impacts on general mathematical attainment as measured by PiM tests, within the timescales of the project. Given the evidence of previous studies that multiplicative reasoning can be improved through targeted intervention[35], and the possibility that the lack of impact was due to the timescale of the project, further development in this area is justified.

In particular, teachers found the models used in the projects as useful approaches to teaching multiplicative reasoning. There is, therefore, the potential to evaluate specific materials in a more focused way - for example using the bar model and ratio table.

Recommendation 1: For the DfE to consider commissioning smaller scale quasi-experimental trials of specific MR materials or models, aiming for a more consistent approach to material use, with outcome measures designed specifically for the trial to measure MR knowledge, and to allow more time to detect impact

Important to the theory of change that underlay the project was the need to address teachers' understanding of multiplicative structures and MR pedagogy. There is the potential to develop further targeted interventions and professional development materials related to this. One approach to doing this would be to draw on international developments in relation to understanding and assessing teachers' multiplicative reasoning knowledge[36]. This could lead to interventions that were assessed by impact on teachers in the first instance as a pre-requisite for evaluating impact on pupils

Recommendation 2: For the DfE to consider developing and evaluating interventions that focus specifically on teachers' multiplicative reasoning knowledge and pedagogy using specific measures of impact that are suitable for this purpose.

The evaluation of the project was conducted in the context of policy concern with evidenced based educational interventions. There are a number of lessons to be learnt in relation to how to best use RCTs to support this, and these would also be relevant to related quasi-experimental designs.

Recommendation 3: For the DfE to consider lessons learnt for the use of RCT and similar methodologies in relation to: pilot and development prior to testing; early appointment of

---

[35] Hodgen, J. et al. (2014)Improving students' understanding of algebra and multiplicative reasoning: Did the ICCAMS intervention work? Pope, S. (Ed.) Proceedings of the 8th British Congress of Mathematics Education 2014, 167-174

[36] For example the Diagnosing Teachers' Multiplicative Reasoning Project http://www.kaputcenter.umassd.edu/projects/dtmr/

external evaluators to advise on intervention design; and longer lead in times to allow for greater balance at randomisation and a longer trial period.

9.2 Impact on teacher professional development

Teachers judged both the curriculum materials and activities to be, on the whole, effective in terms of their professional development, and useful in the classroom.

The project materials were a useful resource for KS3 and potentially KS2 curriculum. They were developed to be used as part of a professional development project. However, they have the potential to be adapted as a stand-alone resource or potentially with on-line or similar PD materials. This might follow the Bowland approach[37] or develop alternatives such as video presentation and discussion of materials by developers or through webinars.

Recommendation 4: That the NCETM is encouraged to refine the MRP materials for use in KS3 and potentially KS2, providing guidance on how they could be used effectively, potentially with on-line PD activities.

Recommendation 5: For the NCETM to make available MRP materials and approaches to teaching MR through the Maths Hub network.

9.3 Impact on new approaches to PD in the school-led system

Participants were generally positive about the form of collaborative professional development with teacher leaders being key to delivery and focused on issues of concern to them.

The regional collaborative communities were led by schools and hosted by Maths Hubs. The Maths Hubs pilot has evolved into the Maths Hub programme and this has already been funded by the DfE. The MRP evaluation supports this decision. Maths Hubs involved in the project, as well as other teaching schools, recognised the potential of the MRP model and materials as an approach to professional development in their hubs.

Recommendation 6: For the DfE to support collaborative professional development with features of the MRP model and similar teacher-led professional development networks through the Maths Hub programme.

The Multiplicative Reasoning Project brought together a team of curriculum developers whose designs were informed by different, but complimentary, pedagogical principles. Developments were supported by other university researchers and teacher professional development leaders. This was recognised by the national development team members,

[37] http://bowlandmaths.org.uk/

including teacher leaders, as a powerful and productive approach. Teachers interviewed also valued the different contributions.

A design approach to curriculum development has potential to address other areas of concern in mathematics. In addition to the curriculum developers involved in MRP there are other recent and significant examples of curriculum design in mathematics in England[38]. The NCETM has a particular role as broker in relation to different contributors to mathematics professional development, as evidenced by its approach in the MRP.

Recommendation 7: That the NCETM seeks further opportunities to engage curriculum developers with Maths Hubs and other NCETM activities and potentially to develop future curriculum design projects that address the needs of teachers, schools and pupils.

One of the effective features of the project, as identified by participants, was utilising the expertise of university based developers and researchers. The Maths Hub programme aligns with this, with HEIs involved in strategic bodies within hubs. The NCETM has also previously encouraged teacher innovators to draw on HEI expertise when awarding direct grants.

Recommendation 8: For the NCETM to encourage HEI involvement in specific projects developed by Maths Hubs.

9.4 Conclusion

The Multiplicative Reasoning Project was a pilot professional development project. Given that it was a pilot and it was conducted over a short time scale it is not surprising that it did not lead to pupil impact as evidenced by external tests. The project was effective in relation to teacher professional development and supported the development of the Maths Hub programme. Those effective features of the project in relation to professional development should be developed and replicated. However, further evaluation of the specific curriculum models is needed to evaluate if they are useful more widely.

[38] For example, Bowland Mathematics, the Effecting Principled Improvement in Science and Mathematics project,the Cre8ate maths project, the NRich project; and the Nuffield applying mathematical processes.

Department for Education

Reference: DFE-RR406A

ISBN: 978-1-78105-487-1

This research was commissioned under the 2010 to 2015 Conservative and Liberal Democrat coalition government. Views expressed in this report are those of the authors. They do not necessarily reflect government policy.

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

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

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