



Education
Endowment
Foundation

Best Practice in Grouping Students Intervention B: Mixed Attainment Grouping Pilot report and executive summary September 2018

Independent evaluators:

Palak Roy, Ben Styles, Matt Walker, Sally Bradshaw, Julie Nelson and Kelly Kettlewell



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

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

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

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus - Private Equity Foundation) and received a founding £125m grant from the Department for Education.

Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.



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

Danielle Mason
Head of Research

Education Endowment Foundation
9th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP
p: 020 7802 1679
e: danielle.mason@eefoundation.org.uk
w: www.educationendowmentfoundation.org.uk

About the evaluator

The project was independently evaluated by a team from the National Foundation for Educational Research (NFER). The principal investigator for this trial was Dr Ben Styles, Head of NFER's Education Trials Unit and the day-to-day trial manager was Palak Roy, Senior Trials Manager from NFER's Centre for Statistics.

Contact details:

National Foundation for Educational Research (NFER)

The Mere
Upton Park
Slough
Berkshire
SL1 2DQ

p: 01753 637385

e: p.roy@nfer.ac.uk

Acknowledgements

We would like to thank the delivery team at UCL Institute of Education for their support throughout the evaluation. We are grateful to Sarah Lynch and Dr Anneka Dawson (in her previous role at NFER) for their help in undertaking some of the interviews in the first year of the evaluation and Jo Morrison for her help during the analysis. We are also thankful to Julie Thompson and Margaret Parfitt at NFER for their administrative support including formatting the report. Finally, we appreciate the support of the team at the Education Endowment Foundation, Emily Yeomans, Camilla Neville, Calum Davey and from 2016 onwards, Anneka Dawson for their guidance throughout this evaluation.

Contents

Executive summary.....	3
Introduction	5
Methods	10
Impact evaluation	24
Process evaluation.....	34
Conclusion.....	43
References.....	45
Appendix A: Recruitment material.....	46
Appendix B: Randomisation syntax.....	65
Appendix C: Theory Of Change	67

Executive summary

The project

The Best Practice in Mixed Attainment (BPMA) intervention trained schools to adopt best practice mixed attainment approaches to grouping Year 7 and Year 8 students for English and mathematics in order to raise attainment. Schools were expected to create classes that contained pupils with a wide range of Key Stage 2 national curriculum results in English and mathematics. Teachers were trained to communicate high expectations of all pupils and to use flexible within class grouping and differentiation techniques. They were asked to minimise the use of marks levels and grades, and to provide formative feedback through comments instead, with a view to creating a growth mindset in pupils. The intervention was developed by a team now based at UCL Institute of Education.

The intervention was developed with three schools between September 2014 and July 2015 and piloted with a further eight intervention schools through a pilot randomised controlled trial (RCT) starting in September 2015 and following Year 7 students for two academic years. The trial was funded by the Education Endowment Foundation (EEF). The aim of the pilot trial was to help inform decisions about a future large-scale RCT and to identify barriers to adopting mixed ability teaching. Preliminary outcome data was collected in order to rehearse for a future trial.

Key conclusions

1. It was challenging to recruit and retain schools to the pilot trial. Despite directly contacting 158 schools and widely advertising the opportunity to another 330, only 18 agreed to be randomly allocated to either receive the training in mixed attainment grouping or to be a control group.
2. Staff had mixed experiences of the intervention; some enjoyed it, whereas others struggled, particularly with differentiation in mixed attainment groups. Schools that continued with the intervention generally adhered to the programme: allocating pupils to mixed attainment classes, applying differentiation techniques in the classroom, and communicating high expectations for all pupils.
3. Most interviewees felt that the intervention had a positive effect on pupil outcomes and that those with low prior attainment particularly benefitted.
4. The pilot RCT was small and designed to test whether a trial was possible rather than to measure the impact of the intervention. The outcome data that was collected did not show a difference in overall maths and English scores between intervention and control schools.
5. Should a future efficacy trial be considered, particular attention must be paid to eligibility criteria, clarity of expectations at recruitment and the teacher workload associated with implementing mixed attainment teaching.

What are the findings?

The recruitment phase of this pilot RCT has demonstrated that, despite evidence that grouping by ability may harm outcomes for disadvantaged pupils, schools are largely reluctant to participate in an evaluation of mixed attainment teaching. Should an efficacy trial of this intervention be considered, particular attention should be paid to eligibility criteria, clarity of expectations at recruitment and the teacher workload associated with preparation for the implementing mixed attainment teaching.

The pilot found that, once in the pilot, schools generally adhered to the programme of mixed attainment teaching. Most of the intervention schools had previously set by attainment, and by the end of the pilot almost all intervention schools reported that students were allocated to classes to ensure that each had a similar range of students in terms of their Key Stage 2 results. All interviewees reported that teachers had high expectations for their students, and that they had systems in place to ensure that these expectations were maintained. Most interviewees reported that their teachers applied effective

differentiation techniques in the classroom. Interviewees also reported cascading the training to the other members of their departments, as required by the intervention. The level of fidelity was rated as moderate.

The final decision as to whether a school joined the intervention or not, was taken at the senior leader level. This meant that some of the non-senior English and mathematics teachers did not fully support teaching in mixed ability groups. Participants' responses to the intervention were mixed, with some enjoying it, and others struggling with it, particularly with differentiation in mixed attainment sets. The teachers had mixed views on the quality of the training sessions. While most attendees enjoyed the initial training session, most interviewees felt that the later sessions were too repetitive. Teachers felt there had been a positive impact on student outcomes. Most interviewees thought that the pupils with low prior attainment had particularly benefitted from the intervention, feeling more comfortable and confident to ask questions and take part in subject specific activities.

The pilot trial was based on only nine schools for the maths outcome and eight schools for the English outcome. It was carried out with an intention of rehearsing the design and analysis for a future trial. The outcome data that was collected did not show a difference between intervention and control pupils' scores in maths and English, but the numbers were too small to reliably detect anything other than an extremely large effect. Preliminary analysis suggests that the intervention may have had a positive impact on less able students for the English outcome. This finding should be treated with caution, due to the size of the pilot trial and because it was one of several statistical tests performed of multiple outcomes and subgroups. There was no evidence of promise for any sub-group of students for the maths outcome, or for pupils who have ever been eligible for free school meals in either subject. Results from the secondary outcomes analyses suggest no evidence of improvements in student self-confidence in maths or English. Many of the schools recruited to the control group for the pilot trial were already using mixed ability grouping. The outcome measures that have been collected are not comparisons of mixed ability grouping with setting or streaming.

How was the pilot conducted?

This was a two-armed cluster pilot RCT. To support recruitment any school was considered eligible to take part regardless of their prior grouping arrangements as long as they were willing to implement mixed attainment grouping. Eighteen secondary schools were randomised. Five were subsequently excluded from the analysis, leaving 13 schools (2,107 pupils) followed up.

A light touch process evaluation ran alongside the pilot RCT. Interviews were used to assess whether the approach is amenable to a full efficacy trial; to understand the difficulties secondary schools have in introducing and implementing a mixed attainment approach; and to identify how these barriers can be overcome and how a mixed attainment approach can be introduced and implemented effectively in secondary schools.

Question	Finding	Comment
<i>Is there evidence to support the theory of change?</i>	N/A	This pilot was designed to test the feasibility of a future trial, rather than to find evidence to support the theory of change.
<i>Was the approach feasible?</i>	Mixed	Recruiting schools to the trial was challenging. However, participating schools did teach pupils in mixed ability classes and apply the differentiation and growth mindset techniques suggested during training.
<i>Is the approach ready to be evaluated in a trial?</i>	Mixed	Should a future efficacy trial be considered, particular attention must be paid to eligibility criteria, clarity of expectations at recruitment and the teacher workload associated with implementing mixed attainment teaching.

Introduction

This report presents findings from a pilot trial and process evaluation of a Best Practice in Mixed Attainment Grouping intervention. This intervention involved training schools to adopt mixed attainment approaches to grouping Year 7 and 8 students for English and mathematics. It was developed by King's College London (KCL) and was later transferred to University College London, Institute of Education (UCL Institute of Education) as part of a larger project called *Best Practice in Grouping Students*, which includes another intervention - *Best Practice in Setting (BPS)* – which was evaluated and reported separately.

The intervention was developed with three schools¹ between September 2014 and July 2015 (feasibility study) and piloted with a further eight intervention schools through a pilot randomised controlled trial (RCT) starting in September 2015 and following children through Years 7 and 8. The evaluation of both phases was conducted by the National Foundation for Educational Research (NFER) between May 2014 and November 2017.

Background evidence

The extent of mixed attainment practice across schools in England at Key Stage 3 is unclear. According to the intervention developers, the practices of setting and streaming have 'steadily increased and now predominate' with mixed attainment teaching becoming increasingly less common (Taylor *et al.*, 2015). According to Ofsted, only around one third of schools visited in 2013 taught students predominantly in mixed-attainment groups through Key Stage 3 (Ofsted, 2013). The same report suggests that setting by ability appears more popular in mathematics compared to other subjects. For English, setting was often undertaken only in the later stages of Key Stage 3 (Ibid). The apparent preference for setting and streaming across schools, and the political rhetoric which also appears to favour these practices (Husbands, 2014), contradict the evidence base, which shows that ability grouping has an overall negative impact on students' attainment, and a particularly negative impact on lower-attaining students (Taylor *et al.*, 2015; EEF, n.d.a).

Evidence about the outcomes of mixed-attainment teaching is inconclusive. We cannot infer that mixed-attainment teaching necessarily produces the reverse effects of ability grouping, either for low-attaining students or for all students. As the UCL Institute of Education team state in their training materials: 'our review of the literature found that the constitution of good practice in mixed attainment classes is currently an under-researched area' (Taylor *et al.*, 2015).

The intervention developed by the UCL Institute of Education team focuses on the pedagogy of mixed-attainment teaching through a range of professional development activities for teachers of mathematics and English. It seeks to test the hypothesis that students with low prior attainment can make better progress in mixed attainment groups than when placed in (low) sets and streams, provided teachers utilise appropriate pedagogies. The intervention has a particular focus on differentiation techniques and establishing a culture of high expectations for all students.

Full details of the intervention are provided below using the TidIER framework as adapted by the EEF (Humphrey *et al.*, n.d.c).

¹ Two pilot schools were located in the South of England and one was in the Midlands.

Intervention

Name:

Best Practice in Mixed Attainment Grouping (BPMA).

Rationale:

Previous research has suggested that young people with low prior attainment make better progress in mixed attainment groups than when placed in (low) sets and streams. This has a bearing for social inequality, as pupils from low socio-economic backgrounds, and from certain black and minority ethnic (BME) groups, are over-represented in low attainment groups (Kutnick *et al.*, 2005). The BPMA intervention and evaluation sought to test this hypothesis, instigating research-informed practices that represent good practice in mixed-ability grouping.

Recipients:

The feasibility study focused on teaching within English and mathematics with a cohort of Year 7 students who were then tracked into Year 8. In the autumn term of 2014/2015, the developers worked with three secondary schools to develop an intervention based on key principles for pedagogy, and exemplar curriculum and assessment materials. In the academic year 2015/16, the intervention was then applied in English and mathematics lessons for the Year 7 student cohort in eight secondary schools as a pilot trial (eight treatment schools, five control schools: 13 schools in total). The same cohorts of students were followed up in the trial in the next academic year, as they progressed to Year 8, hence applying the treatment across two years.

Professional development and materials:

The developers developed a set of guidance for schools around:

- growth mindset and flexible conceptions of intelligence – the notion that teachers' high expectations of students and encouragement of their belief in their ability to overcome problems, can improve outcomes;
- high expectations—the notion that teachers' high expectations of students, communicated clearly through challenging and appropriate learning objectives, is a powerful vehicle to raising achievement for students at all levels of prior attainment;
- Within-class grouping - flexible within-class grouping is preferred to rigid ability-based grouping strategies as it is not detrimental to lower-attaining or disadvantaged students; and
- Differentiation—the key to effective differentiation is to know students well, to avoid labelling them, to diagnose their prior attainment accurately in order to apply differentiation flexibly, and to model exemplar lessons as part of the professional development and training.

These were communicated via workshops (details below) geared towards supporting practitioners to understand the principles, and to begin to plan their application in their own school contexts; and via discussion and observation feedback on classroom practice.

Procedures:

In addition to the professional development on offer, schools were expected to follow a number of procedures in the way they allocated students to classes. To ensure a broad range of attainment in each class, students were allocated to Year 7 classes primarily on the basis of their Key Stage 2 national curriculum results in English and mathematics. Teachers who attended the training were expected to cascade their learning to members of their departments to ensure that all teachers implemented the intervention. Teachers were also expected to develop and maintain high expectations of all students'

attainment in English and mathematics relative to their prior attainment, but regardless of their social backgrounds or characteristics. To encourage a growth mindset, teachers were asked to minimise the use of marks, levels or grades. Instead, teachers were encouraged to provide formative feedback through comments rather than grades as much as possible.

Implementers and mode of delivery:

The developer's original intention was that the BPMA pilot would involve staff attending three full days of CPD (delivered by the KCL (subsequently the UCL Institute of Education) delivery team in school time, as part of regional workshops) and three twilight sessions. However, instead, UCL Institute of Education held four full days and two twilight sessions for the London hub and three full days plus two twilights for the York hub (one twilight was cancelled due to industrial action by a teaching union). The CPD sessions lasted approximately five hours and the twilight sessions lasted for approximately two hours. It was expected that two Year 7 English and two Year 7 mathematics teachers would attend each regional professional development session and that, where possible, this should include the head of English and the head of mathematics.

Tailoring:

BPMA was a manualised (i.e. highly specified) intervention, and the developers emphasised the importance of optimal treatment fidelity. Nonetheless, participating schools were permitted to make limited adaptations (for example, to the process of allocating students to classes, by using teacher or other forms of assessment where Key Stage 2 attainment results were unavailable).

Planning:

Strategies to maximise implementation effectiveness included the use of initial developer-led workshops to develop department-wide approaches to addressing factors such as improving student engagement and attitudes and raising teacher expectations and pedagogy. Later workshops were designed to review earlier learning and to create an arena in which participants could discuss issues arising, monitor progress, and share best practice. The developers' request for heads of department to be present at each regional professional development session was also intended to maximise implementation effectiveness.

Evaluation objectives

This evaluation was set up as a feasibility study and pilot RCT. The main objective of the feasibility study was to develop the intervention with three secondary schools in academic year 2014/2015 and to examine the barriers to using mixed attainment grouping in secondary schools. The pilot RCT was set up to help inform decisions about a future large-scale RCT.

Research questions

The primary research questions from the protocol were:

1. Can secondary schools be recruited to a mixed attainment RCT?
2. What are the barriers to schools adopting a mixed attainment approach?

In addition to these, and since this was a pilot trial, impact analysis was undertaken in order to rehearse for a future trial. The primary outcome measures of the impact evaluation were:

1. What is the impact of mixed attainment grouping on pupils' attainment in mathematics?
2. What is the impact of mixed attainment grouping on pupils' attainment in English?

The delivery team also administered pupil surveys that measured self-confidence in mathematics and English at baseline and at end-point. We used these measures in our secondary analyses.

NFER also carried out a light-touch process evaluation that focused on the training given to schools, the ease of and barriers to implementation, and how scalable the intervention was perceived to be.

Ethical review

KCL obtained ethical approval for the study on 24 July 2014 and NFER's Code of Practice Committee approved the data collection for this study on 5 March, 2015. For the data collection and processing, this trial also followed the EEF's advice on consent and the Data Protection Act (EEF, n.d.b). Schools opted into the trial through the headteacher (or their designated deputy) signing an MoU during recruitment. On this form, they nominated a member of staff to be the main contact throughout the trial. The school information sheet, invitation letter and reply form all contained relevant information about consent and how the data from the National Pupil Database (NPD) would be used. We obtained administrative pupil data (pupil names, dates of birth and UPN) from the participating schools in order to match pupil data such as assessment data, attitudinal data and pupils' group membership within a class to the NPD background characteristics (prior attainment at Key Stage 2 and free school meal (FSM) eligibility). Parental opt-out consent letters were administered by schools prior to them sending this data to us.

Appendix A provides the school information sheet, recruitment invitation letter to schools, parent consent letter, school MoU and reply form.

Project team

The principal investigator for this trial was Dr Ben Styles, Head of NFER's Education Trials Unit. The day-to-day trial manager was Palak Roy, Senior Trials Manager from NFER's Centre for Statistics (who took this role from March 2015). Prior to this date, the trial was managed by Matt Walker. Sally Bradshaw and Jo Morrison (NFER statisticians) supported the impact evaluation. The process evaluation was led by a team of researchers from NFER's Centre for Policy and Practice Research: Matt Walker, Dr Julie Nelson and Kelly Kettlewell. School recruitment and communications were managed by researchers from NFER's Research and Products Operations department: Keren Beddow, Asma Ullah and Kathryn Hurd. The GL Assessment test administration was managed by Shalini Sharma and the tests were administered by trained NFER test administrators.

The intervention was developed and delivered by a team at KCL led by Professor Becky Francis, who was at KCL until June 2016. After this point, the delivery team moved to UCL Institute of Education. The project manager was Dr Becky Taylor. They were supported by a team of researchers from Queen's University Belfast (QUB) and the University of Nottingham (UoN).

NFER was responsible jointly with the delivery team for the pilot trial design, school recruitment and ongoing relationship with the schools. It had sole responsibility for randomisation, administering pupil data collection, conduct, analysis and reporting of the independent pilot evaluation.

The project was supported and guided by EEF staff Emily Yeomans, Camilla Neville, Calum Davey and from 2016 onwards, Dr Anneka Dawson.

Trial registration

The pilot trial was registered as trial number: **ISRCTN10304032**.

Methods

Trial design

The evaluation started with a feasibility study in three schools in the academic year 2014/15. These were schools, identified through networks, that were rated by Ofsted as 'Good' or 'Outstanding' and practising (or willing to practise) mixed attainment grouping. Two were already practising mixed attainment grouping in English, Two were already practising mixed attainment grouping in maths. One school moved to mixed attainment grouping in maths for the feasibility year.

This feasibility study was used to develop the intervention and to examine the barriers to using mixed attainment grouping in secondary schools. Following this phase, the approach was piloted as a randomised controlled trial (RCT), starting in September 2015 and following Year 7 students for two academic years (to July 2017). The target sample size for the evaluation was 20 secondary schools, randomised either to receive the intervention or to be part of a control group. Note that this target was not driven by a minimum detectable effect size (MDES) calculation, but simply to ensure a variety of school contexts and national diversity in location of the ten secondary schools engaging with the intervention. In order to keep the control group engaged with the trial, each school from this group received £1000 once they completed the primary outcomes tests in summer 2017.

Recruitment

NFER and KCL jointly managed the recruitment for this trial, each deploying their own distinct recruitment approach. KCL publicised both the trials in social and traditional media, including but not limited to: TES, *Leader* magazine, networking events run by the delivery team, publicity on Twitter and visits to stakeholder events such as NAMA conference and headteacher meetings. Schools, then, contacted KCL to inquire about the mixed attainment pilot trial as well as the trial of Best Practice in Setting (BPS). In contrast, NFER contacted schools by drawing a stratified random sample of schools to take part in this trial and a separate sample for the BPS trial.

For the BPMA pilot trial, NFER's sampling frame included all English state secondary schools that had both a Year 7 and a Year 8 group containing at least 60 students². As both NFER and KCL were recruiting concurrently, NFER excluded schools from the five Local Authorities (LAs)³ within which KCL focused their recruitment. On this basis, NFER drew a sample of 158 state secondary schools stratified by government office region, school type and proportion of pupils eligible for free school meals (FSM). NFER approached all schools with an initial invitation letter to the headteacher in March 2015. This was followed by an extensive reminder exercise which consisted of a mixture of written reminder letters, phone reminders and emails to named individuals where available. In most cases, NFER staff asked to speak to the headteacher or a senior leader. After a month, we amended the approach slightly, contacting subject heads directly in the first instance, as it was proving to be difficult to establish contact with headteachers. We therefore sent recruitment letters to the head of maths and/or head of English, who themselves sought headteacher buy-in. We also asked to speak to subject heads when phoning the school. Despite all these efforts, only 18 schools (rather than the target of 20) agreed to take part in the trial.

² We needed to test at least 60 students from each school. Please see the 'Sample size' section for more details.

³ These were Kingston, Barking and Dagenham, Luton/central Bedfordshire, Norfolk, York and Hampshire

Selection criteria

During the feasibility study (academic year 2014/2015), the delivery team developed and piloted the intervention in three secondary schools. As per the protocol, selected schools had an Ofsted rating of at least 'Good' and to have some experience of mixed attainment practice.

This trial was originally conceived as an opportunity for schools to embrace mixed attainment teaching as an alternative to setting or streaming. By the time the protocol was written and the feasibility study completed, both the developers and the evaluators realised that it would not normally be possible to change the grouping practices of a secondary school in such a radical manner. Instead, it was acknowledged that the kind of school that would be interested in being involved in this research might already be employing mixed attainment teaching. Therefore, to support recruitment any school was considered eligible to take part regardless of their prior grouping arrangements as long as they were willing to implement mixed attainment grouping if randomised to the intervention group. This means participating schools could already have been setting or streaming, or grouping students by mixed attainment. Since information on schools' grouping arrangements is not routinely available on school-level datasets, schools completed a short proforma at recruitment stage indicating their grouping practices at that time (please see the MoU and reply form in appendix A). Schools' prior grouping arrangements collected via this reply form (or short proforma) are presented in Table 8.

Sixty pupils from each secondary school were selected to complete the primary outcome tests on the basis that half would be tested in maths and half in English. These pupils were randomly selected from the pupil list provided by the schools which was subsequently matched with the NPD. As FSM eligibility represents an important sub-group, this random selection of pupils was stratified by pupil FSM eligibility.

Outcome measures

Primary outcome measures

Whilst this was not a fully powered trial, testing was planned to take place in Year 8 after two years of the intervention to test whether implementation of mixed attainment teaching based on the BPMA intervention impacts on attainment in mathematics or English. As the intervention was aimed at students in Years 7 and 8, testing was necessary as there is no statutory assessment in these years. Testing took place in June and July 2017, right at the end of Year 8. GL Assessment's Progress Test in English⁴ (PTE13) and Progress Test in Mathematics (PTM13) were used to measure outcomes in English and maths.

As the tests we used have a broad coverage of the curriculum, we used the raw total score for each subject, which covers all curriculum content. The maths total score (with a maximum possible score of 70) consisted of fluency in facts and procedures, fluency in conceptual understanding, mathematical reasoning and problem solving. The English total score (with a maximum possible score of 66) consisted of spelling, grammar and punctuation, and reading comprehension: narrative and non-narrative. For both these outcomes, higher scores indicate better performance. Further information for both the tests can be found on the GL Assessment website⁵.

⁴ At the time of the protocol, these tests were being developed by GL assessment. They were being called New Progress in English (NPIE) and New Progress in Mathematics (NPIM). After the development, these tests were named Progress Test in English and Progress Test in Mathematics.

⁵<https://www.gl-assessment.co.uk/media/1346/ptm-technical-information.pdf>

NFER took responsibility for collecting and delivering the PTE13 and PTM13 in paper form using our test administrators. The test administrators had clear guidance on how to administer the test, which also emphasised the importance of the trial methodology. The administrators did not know the group allocation for the schools and the administration guidance specifically asked them to avoid discussion with the school staff about the group allocation. Once the tests were completed, we sent these to GL Assessment for marking which was also blind to group allocation.

To address a lower than anticipated response rate for the GL Assessment test, NFER put in place extensive reminding strategies and also extended the time period for the test administration. We also allowed schools to test their entire Year 8 cohort where this was their preference. Some schools said that this would be more convenient than administering the test with 60 students who could be in different classes. Other schools were already planning to administer the GL Assessment tests with their entire Year 8 cohort and preferred NFER to access their pupil results directly from GL. All the available PTE13 and PTM13 results were used from the participating cohort.

Nine schools took the GL Assessment tests; all nine took the maths test (six intervention schools and three control schools) and eight took the English test (five intervention schools and three control schools). In addition, one school took the online versions of the PTE13 and PTM13 tests. As the online assessments are not directly comparable to the paper versions of the assessment, it was necessary to link the raw scores for these assessments. As the curriculum domains covered by the paper and the online versions are the same and the age-standardised scores were standardised using national samples, we were able to link the age-standardised scores from the online versions to the raw scores for the paper-based versions. Using pupils' age in months and their age-standardised scores on the digital versions, we looked up the raw scores on the paper test versions that would have been obtained by a pupil of the same age with the same age-standardised score. In cases where the exact age-standardised score did not exist for that age⁶, we used linear interpolation⁷. The primary analysis was done using only the paper tests as per the protocol (354 pupils from nine schools took the paper based maths tests and 329 pupils from eight schools took part in paper based English tests). Further details are provided in figure 2. A subsequent sensitivity analysis included the school that used the digital version, with its pupils' scores converted to paper scores.

Secondary outcomes: pupil attitudes

As outlined in the Statistical Analysis Plan (SAP) (EEF, 2017), the secondary outcome measures were:

1. What is the impact of mixed attainment grouping on pupils' self-confidence in mathematics?
2. What is the impact of mixed attainment grouping on pupils' self-confidence in English?

The delivery team measured these by administering a pupil survey at the start of Year 7 in September 2015 (baseline survey administered post randomisation) and at the end of Year 8 in summer 2017 (end-

<https://www.gl-assessment.co.uk/media/1366/pte-technical-information.pdf>

<https://www.gl-assessment.co.uk/media/1384/ptm13-links-to-nc.pdf>

https://www.gl-assessment.co.uk/media/1349/pte13-links-to-national-curricula_0.pdf

⁶ For example, if scores were 98, 100, 101 and 104 for successive raw score points, it was not possible to look up values of 99, 102 and 103.

⁷ Linear interpolation aims to find a value between two existing data points on a line.

point survey). The delivery team was responsible for administration of these surveys. The surveys were administered with an entire cohort from all participating schools. As participation in this trial is for both English and maths, all the pupil surveys were used to measure the secondary outcomes of self-confidence.

In partnership with Queen's University Belfast, the delivery team developed the pupil self-confidence measures in maths and English. Self-confidence composite measures were developed using factor analysis on selected items from the baseline pupil survey data (a combined dataset for both BPS and BPMA trials). These items were drawn or adapted from several existing instruments; SDQII (Marsh 1990); TIMSS questions (IEA, 2011) and PISA questions (OECD, 2012). Please see Table 2 for the list of items included in the principal axis factor analysis. After the factor analysis, all the items were retained and the self-confidence composite measures were created as an average of all constituent items⁸. Self-confidence in maths had an internal reliability (Cronbach's α) of 0.88 and in English of 0.86. Scores for these composite measures ranged from one to five with higher scores reflecting higher self-confidence in the given subject⁹. The delivery team sent us a combined dataset that included baseline and end-point data for individual items and the composite measures. We used the composite measures as the secondary outcomes of pupil self-confidence.

Table 2: List of items included in the secondary outcome measures of self-confidence

Composite measure	Constituent items	Source
Self-confidence in mathematics	Work in Maths lessons is easy for me	Adapted from Marsh (1990) verbal [sic] self-concept
	I am not very good at Maths	Adapted from Marsh (1990) verbal [sic] self-concept
	Maths is one of my best subjects	Adapted from Marsh (1990)
	I hate maths	Adapted from Marsh (1990)
	I do well at maths	Adapted from Marsh (1990) school [sic] self-concept
	I get good marks in maths	Adapted from PISA self-concept in mathematics and Marsh (1990) verbal [sic] self-concept
	I learn things quickly in maths lessons	Adapted from TIMSS self-confidence in learning mathematics and Marsh (1990) verbal [sic] self-concept

⁸ While most pupils had responded to the majority of the constituent items, self-confidence measures were created if a pupil had responded to at least one of the constituent items.

⁹ Note that there are two items in each composite measure which are negatively worded (items in grey cells). Scores for these items were reverse-coded such that higher scores represent positive attitudes.

Composite measure	Constituent items	Source
Self-confidence in English	Work in English lessons is easy for me	Adapted from Marsh (1990) verbal self-concept
	I am not very good at English	Adapted from Marsh (1990) verbal self-concept
	English is one of my best subjects	Marsh (1990) verbal self-concept
	I hate English	Marsh (1990) verbal self-concept
	I do well at English	Adapted from Marsh (1990) school [sic] self-concept
	I get good marks in English	Adapted from Marsh (1990) verbal self-concept
	I learn things quickly in English lessons	Adapted from Marsh (1990) verbal self-concept

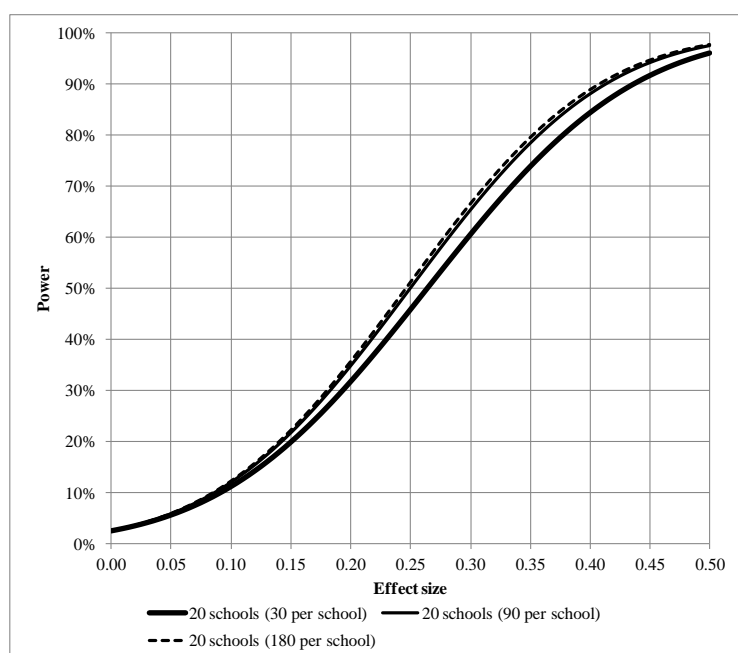
Sample size

The sample size was not driven by a minimum detectable effect size (MDES) calculation, but by the developer. UCL IoE wanted more than a small pilot trial to discern the myriad different approaches of schools to ability grouping. Furthermore, the aim was to measure attainment both to dispel fears of a large negative effect and to carry out further secondary data analysis. Whilst the proposed sample size of ten schools in each group was viewed by NFER to be both larger than a conventional pilot trial and underpowered in terms of an efficacy trial, it was agreed to proceed with the design and analysis as a rehearsal for a future larger trial.

We considered a number of within- school sample sizes for analysis. In order to reduce testing burden per school without sizeable impact on the power, it was decided that NFER would randomly select 60 pupils from the Year 8 school roll from each of the recruited schools. Half of the pupils would sit the mathematics test and half would sit the English test.

The power calculations used the following assumptions: intra-cluster correlation of 0.15¹⁰ (lowered from 0.2 through the use of Key Stage 2 as a covariate); correlation between Key Stage 2 and Year 8 test of 0.7 and average cohort size of 180. Figure 1 demonstrates how there is little difference in power between a random sample of 60 pupils (30 for each of English and maths) in each school, or a far larger sample, such as 360 pupils (180 for each of English and maths). We therefore randomly selected 30 pupils from the cohort to take an English test and 30 others to take a maths test. As this was a pilot trial, it was not fully powered.

¹⁰ The protocol was written before the EEF published guidelines on ICC and pre-tests. All the assumptions in these calculations are based on NFER's unpublished work.

Figure 1: Power curves for ten schools in each group (as per the protocol)

Randomisation

The planned sample size was 20 and the original design meant that, as a result of cluster randomisation, half the schools would be randomly allocated to each group. However, as noted in the recruitment section, it was only possible to recruit and randomise 18 schools. This meant that fewer than ten schools would have been allocated to each group as a result of school-level randomisation. The delivery team was keen to deliver the intervention to as many schools as possible and at least to ten schools. It was, therefore, agreed that NFER would undertake random allocation unequally, with schools allocated to intervention and control group with a 2:1 ratio. This was a protocol deviation.

An NFER statistician carried out the randomisation using a full syntax audit trail. This was done in two blocks due to staggered school recruitment and intervention workshops running concurrently. Randomisation took place in June and July 2015.

In total, 18 schools signed an MoU and were subsequently randomised. Twelve schools were assigned to the intervention group and six schools were assigned to the control group. Of these, one control school was randomised due to an administrative error (the school never intended to take part) and was subsequently removed from the trial. Four intervention schools withdrew participation and were retrospectively classed as 'not eligible'. Three of these schools approached the delivery team to indicate that they were not planning to operate mixed attainment grouping. These schools considered themselves to be operating mixed attainment but had components of attainment based grouping such as top, middle and bottom groups one more school withdrew participation since it had not realised that it was signing up to a trial. Since this is a pilot RCT, it was agreed that these schools would be recognised as 'not eligible' even though they were randomised. As these schools withdrew from the trial, no testing was possible in them anyway, and they were excluded from all analyses (including ITT analysis). For an efficacy trial, these schools would be considered missing at follow-up. Since this was a pilot trial and the main aim of any impact analysis was merely to rehearse for a future trial, it was not necessary to retain them in the analysis and instead to report the drop-out as an important finding. Table 3 details the blocked randomisation. Please see the randomisation syntax in Appendix B.

Table 3: Number and proportion of schools randomised

Block/group	Intervention n (%)	Control n (%)
Block 1 (16 th June 2015)	10 (67%)	5 (33%)
Block 2 (6 th July 2015)	2 (67%)	1 (33%)
Total	12 (67%)	6 (33%)

Analysis

We followed EEF analysis guidelines¹¹ and the published SAP for this pilot trial (EEF, 2017). This section provides an overview of the analysis undertaken and the SAP provides further details.

Primary intention-to-treat (ITT) analysis

The analysis of this pilot RCT was mainly descriptive in nature. As per the protocol and SAP, school means¹² analysis was carried out as the primary outcome measure of the trial. In this analysis, we used the school averages of the PTE13 and PTM13 scores in two separate regression models. The number of schools in the trial was considered too low for a multi-level model. For the English model, PTE13 school average was the dependent variable with the following covariates:

- an indicator of whether the school is an intervention school, and
- the school's mean prior attainment as measured by Key Stage 2 English point score (KS2_KS2READPS variable).

Similarly, the model for the maths outcome included the mean PTM13 score as a dependent variable regressed on the following covariates:

- an indicator of whether the school is an intervention school, and
- the school's mean prior attainment as measured by the Key Stage 2 maths point score (KS2_KS2MATPS variable).

Secondary outcome analyses

The secondary analyses included attainment and self-confidence models. We analysed these using multilevel models, taking account of clustering and to mimic what might be done in a fully powered trial. These models included all pupils with non-missing values for all the variables in the models (covariates as well as the outcome measures) i.e. these were completers models. For the attainment outcomes, pupils were sampled to take either a maths test or an English test. Self-confidence measures were collected through UCL Institute of Education's pupil surveys. So the self-confidence models included all the pupils with a baseline and an end-point self-confidence measure in a given subject. Details on the secondary outcome analyses are given below.

¹¹https://educationendowmentfoundation.org.uk/public/files/Evaluation/Writing_a_Research_Report/2015_Analysis_for_EEF_evaluations.pdf

¹² School means were calculated using the assessment data for the analysed groups only. e.g. school's prior attainment in English were the mean KS2 English point score for the pupils who also had a valid English outcome measure.

Attainment outcomes

Multilevel models with two levels (school and pupil) were used for the analysis to account for the cluster randomisation and to rehearse what might be done in a fully powered trial. We included all schools with pupil data on the GL Assessment English and maths outcomes in this analysis, irrespective of whether or not the schools had implemented the intervention. There were two separate models, one for each subject¹³. The dependent variable for the mathematics model was the raw total score in mathematics for PTM13 with the following covariates:

1. an indicator of whether the pupil is in an intervention school
2. pupil prior attainment as measured by Key Stage 2 Maths point score (KS2_KS2MATPS variable).

The dependent variable for the English model was the raw total score in English for PTE13 with the following covariates:

- an indicator of whether the pupil is in an intervention school
- pupil prior attainment as measured by Key Stage 2 English point score (KS2_KS2READPS variable).

Self-confidence outcomes

We analysed pupil self-confidence data as the secondary outcome measures and used them as dependent variables in two separate multilevel models. The covariates for these models were similar to the attainment models, insofar as pupil self-confidence measures in the given subject at baseline were one of the covariates, rather than prior attainment measures. Pupils were included in the models if they had both baseline and end-point self-confidence measures in a given subject i.e. these were completers models.

Non-compliance with intervention

The delivery team collected and supplied us with data on the level of school engagement throughout the two-year delivery period, using a number of pre-defined variables. This categorisation yielded four broad measures, which are listed in Table 4. Measures for English are given as an example. Similar information was collected for mathematics. Unfortunately, there is not sufficient data to create combined measures using this data (as planned in the SAP). Hence, descriptive findings are presented in this report. Please see Table 4 for a list of dosage variables.

¹³ Note that pupils taking the English tests were different from those taking the maths tests but they were from the same schools.

Table 4: Dosage variables for English subject

Combined measure	Variable	Level of measurement
Effectiveness of training practices	1. English department represented at each training session	Binary. Did the expected number and type of staff attend each session?
		0 = No
		1 = Yes
	2. Training is cascaded to members of the English department	Binary. Has some form of cascading/internal training taken place?
		0 = No
		1 = Yes (if one or more departmental members concur)
Effectiveness of student allocation	3. Students are allocated to classes so that there is a similar range of KS2 results in each class	Binary – there is a similar range of KS2 results in each class
		0 = No
		1 = Yes
High expectations	4. Teachers have high expectations for all students	Binary
		0 = No
		1 = Yes
Effectiveness of differentiation practices	5. Teachers apply effective differentiation techniques in the classroom	Binary
		0 = No
		1 = Yes

Grouping practices of all schools

The delivery team also collected information from all participating schools about their grouping practices for each subject. This was collected twice during the pilot trial, via teacher interviews in the first academic year and via a final grouping proforma in the second academic year. The grouping practices were defined as completely 'mixed', 'mixed with a top set', 'mixed with a bottom set', 'mixed with a top and bottom set', 'setting', 'streaming' or 'other'. This data is presented in the report as important contextual information as it was never intended to be incorporated into any of the findings.

Subgroup analyses

The main purpose of running the subgroup analyses was to rehearse for a future fully powered trial. We explored the differential effect for different pupil ability levels. We achieved this by adding an interaction term to the multilevel attainment models. The intervention indicator was interacted with pupil ability as measured by the aforementioned prior attainment measures at Key Stage 2.

As per the EEF guidance, there was another interaction model regarding whether a pupil had ever received free school meals (as measured by the EVERFSM_6 variable). This was done using a model identical to the attainment outcome model but including EVERFSM_6 and EVERFSM_6 interacted with the intervention indicator as covariates.

A separate analysis of everFSM pupils was also carried out as per the EEF analysis guidance. These models were similar to the main models of overall effect, but only included everFSM pupils as measured by the EVERFSM_6 variable.

As expected, these models included a small number of pupils and hence were exploratory in nature.

We manipulated the data in SPSS while running the multilevel models in R package nlme.

Cost analysis

The information on which the cost calculations are based was supplied on a termly basis by the delivery team as well as the schools themselves. The delivery team provided information about staff time and other direct, marginal costs, such as staff travel costs and venue hire. Data on teacher travel and supply cover costs was also provided. All intervention schools were asked to submit cost data. However, only five participating schools provided this information in 2015/2016, while two provided it in 2016/2017. This included data on the number of hours of cover that they used, the number of these hours that required paid supply cover, and the cost of this, together with details of other monetary costs, such as travel and subsistence costs. These school costs were for activities delivered in support of the intervention, but not directly associated with attendance at the training delivered by UCL Institute of Education. This information was collected using proformas, first by the delivery team in the summer term 2016, and then by NFER in the summer term 2017.

We estimated a per school cost based on the average costs for one department (English or maths) per year. Where both the English and maths department was participating in a school, the aggregated cost data was halved to provide an average for each participating department. Once a mean cost per department per year had been calculated, we then divided this figure by the average number of Year 7 students in participating schools (using administrative student data collected by NFER), to calculate the average cost per student per year. Cumulative costs are reported over a period of three years. More detail is provided in the analysis section below.

Implementation and process evaluation

A light touch process evaluation ran alongside the pilot trial. Its aims were to:

- assist in decision-making as to whether the approach is amenable to a full trial;
- understand the difficulties secondary schools have in introducing and implementing a mixed attainment approach; and
- identify how these barriers can be overcome, and how a mixed attainment approach can be introduced and implemented effectively in secondary schools.

The process evaluation involved four main strands of activity spread across three academic years, as outlined in Table 5.

Table 5: Overview of research strands associated with the process evaluation (conducted by NFER)

Evaluation activities	Year 1 (2014/15) (Feasibility study)	Year 2 (2015/16)	Year 3 (2016/17)
Exploratory scoping interviews with the developer (KCL/UCL Institute of Education) ¹⁴	✓		
Attend training/briefing event for schools ¹⁵	✓		
School proformas (to collect cost data) ¹⁶		✓	✓
Telephone interviews with Heads of English and Mathematics	✓	✓	✓

In the feasibility study, exploratory scoping interviews were undertaken with three key staff from the developer team to identify relevant management and monitoring information. Interviews were also undertaken with two heads of English and three heads of Mathematics to explore their views on the initiative. During this stage of the project, a Theory of Change was developed by the delivery team, which is included in Appendix C.

The NFER team attended one of the three twilight training sessions in July 2015 in London, prior to the pilot trial starting in the autumn term. This allowed the researchers to gain a better understanding of the intervention and teachers' responses to it.

At the start of the pilot trial, information was collected to establish schools' ability grouping practices prior to randomisation. Following randomisation, a further round of data collection was undertaken to establish what the implementation of best practice in mixed attainment looked like and to ascertain what organisational, curricular and grouping changes had been made. Throughout the intervention, information was collected by UCL Institute of Education—using staff questionnaires and proformas—on:

- whether English and/or mathematics departments were represented at each training session;
- whether training was being cascaded to members of the English and/or mathematics department;
- the extent to which students were allocated to classes according to Key Stage 2 attainment results;
- the extent to which teachers reported having high expectations for all students; and
- whether teachers applied effective differentiation techniques in the classroom.

Cost data was also collected from participating schools, first by UCL Institute of Education in the summer term of 2016, and then again by NFER in the summer term 2017.

¹⁴ It was always the intention to speak to the developer only in Year 1 of the evaluation. However, the protocol mistakenly included a table that indicated the developer would be interviewed in each year of the evaluation.

¹⁵ It was always the intention to attend a briefing event only in Year 1 of the evaluation. However, the protocol mistakenly included a table that indicated that such events would be attended in both years one and two of the evaluation.

¹⁶ Note that UCL Institute of Education collected schools' cost data for the academic year 2015/16

In years two and three of the evaluation, we undertook telephone interviews with a sample of Heads of English and Mathematics, or a suitable alternative spokesperson. Alternatives included an assistant headteacher and another teacher working in the same department who had taught the same student cohort in the intervention schools. The interviews explored schools' reasons for participating in the trial, the extent to which schools were implementing each of the key principles behind the intervention, how staff had responded to the intervention, and respondents' perceptions of its impact. The interviews were undertaken using a semi-structured interview schedule and were intended to take approximately 30 minutes to complete. The target was to achieve interviews with staff in one or both departments from five randomly selected schools in each year of the evaluation. In the first year of the trial, a total of seven interviews were undertaken between March and June 2016, covering seven different departments (four mathematics and three English) from five schools. In the second year of the trial, a total of five interviews were undertaken between March and May 2017, covering six different departments (four mathematics and two English) from four schools (one interviewee was an assistant headteacher, who attended the workshops, and who could talk about both the English and mathematics departments). Table 6 clarifies the roles of interviewees.

Table 6: Process interviews undertaken in each year of the evaluation

School	Year 2 of evaluation (2015/16)		Year 3 of evaluation (2016/17)	
	Role of interviewee(s)	Department(s) represented	Role of interviewee(s)	Department(s) represented
1	Head of Mathematics	Mathematics	-	-
2	Head of English	English	-	-
	Head of Mathematics	Mathematics		
3	Head of Mathematics*	Mathematics	Head of Mathematics*	Mathematics
4	Head of English*	English	Head of English*	English
			Mathematics teacher	Mathematics

School	Year 2 of evaluation (2015/16)		Year 3 of evaluation (2016/17)	
	Role of interviewee(s)	Department(s) represented	Role of interviewee(s)	Department(s) represented
5	Head of English Head of Mathematics	English Mathematics	Senior leader	English Mathematics
6	-	-	Head of Mathematics	Mathematics
Total number of interviews	N=7		N=5	

*Indicates that the same interviewee was interviewed in both 2015/16 and 2016/17.

As the numbers suggest, it was harder to secure interviews with participating teachers in the second year of the trial. This was despite repeated email and phone reminders to schools, and the offer of reducing the interview time. Reasons for this lower than anticipated response rate included a perception, amongst at least some school respondents, that the intervention was over-burdensome or not useful which, as a result, adversely affected schools' willingness to participate in the evaluation. One interviewee, a head of Mathematics, reported that his department had withdrawn from the trial in year 3 of the evaluation due to concerns about the performance of their more able students¹⁷. We explore teachers' views in the 'process evaluation' section of this report, having analysed the qualitative data thematically. It should be noted that the process evaluation findings are based on self-reported data from a relatively small number of participants.

Timeline

The intervention was developed with three schools between September 2014 and July 2015. The intervention delivery for the pilot trial commenced in September 2015. The primary attainment outcomes were captured through testing in summer 2017 and the first draft report was submitted in November 2017.

¹⁷ We were able to collect primary outcome measures for maths and English from this school, and the school was included in the ITT analysis.

Table 7: Timeline¹⁸

Date	Activity
September 2014–July 2015	Feasibility study- developing the intervention with three pilot schools
Jan–March 2015	Write, agree and publish the protocol
April–May 2015	UCL Institute of Education developed the Theory of Change Amended version of the protocol published
May–August 2015	School recruitment for the pilot trial completed School randomisation Trial registration on http://www.controlled-trials.com/ Interviews with the delivery team Interviews with the pilot schools (during the feasibility phase) Intervention delivery commenced—first professional development (PD) event in July 2015 Observed training workshops
September–October 2015	Administration pupil data collection commenced (schools administer parental opt-out) Baseline pupil survey (secondary outcome measure) Second PD event
November 2015	Third PD event
March–June 2016	Telephone interviews with sub-sample of participants
March 2016	Fourth PD event
July 2016	Fifth PD event
September 2016	Cost proforma sent to schools (school costs associated with first academic year)
March 2017	Sixth PD event
March –May 2017	Telephone interviews with sub-sample of participants
June 2017	Administrative pupil data collection ended Cost proforma sent to schools (school costs associated with second academic year) End-point pupil survey administration (secondary outcome measures)
May–August 2017	Write and agree SAP
July–September 2017	GL Assessment's test administration and marking (primary outcome measures)
August 2017–November 2017	Analysis and reporting

¹⁸ The timeline excludes aspects of the data collection that was sole responsibility of the delivery team

Impact evaluation

Participants

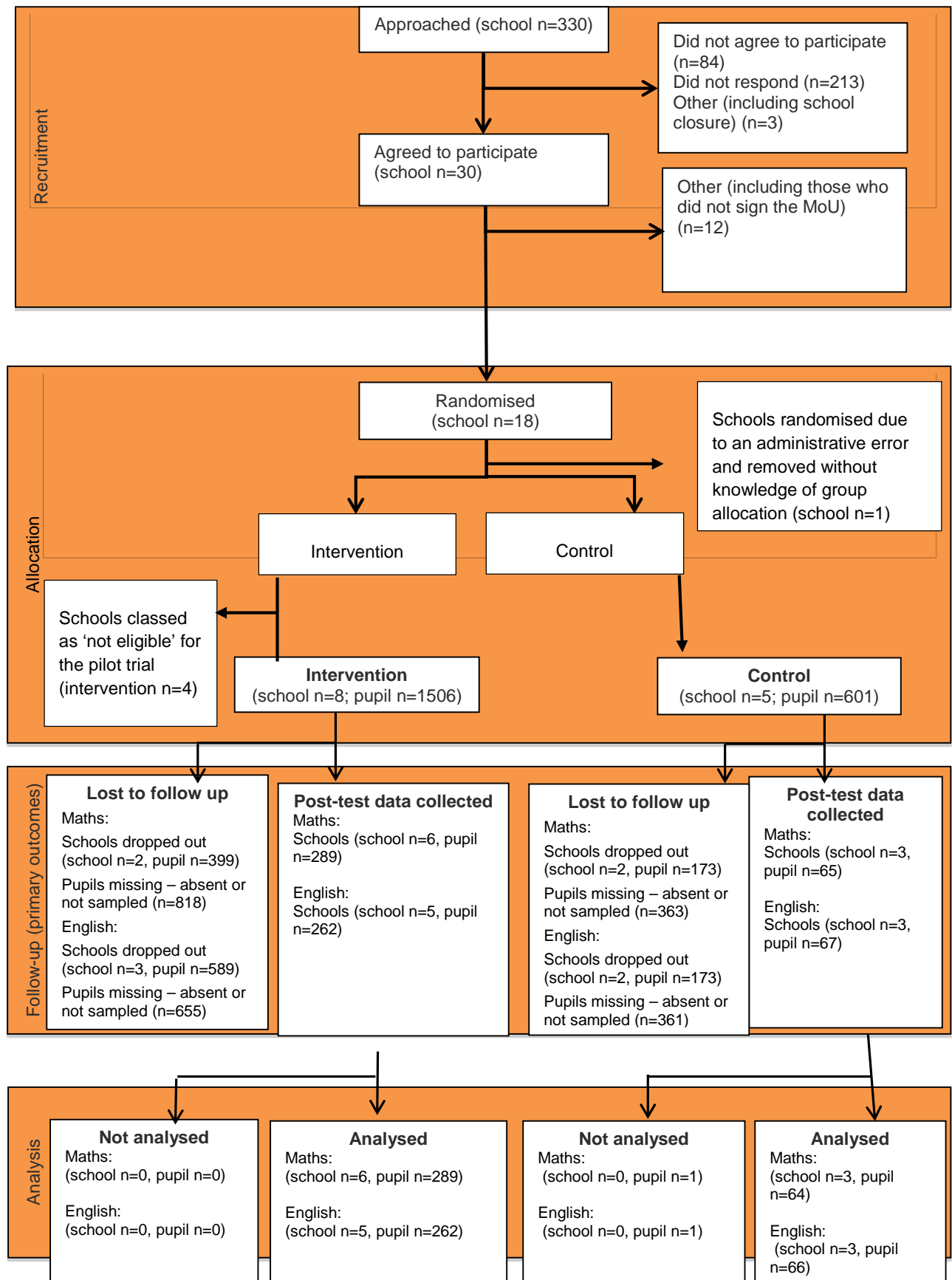
Figure 2 provides details of the participant flow through each stage of the BPMA pilot trial—from recruitment to analysis for the attainment outcome measures. As mentioned previously, it was not possible to recruit 20 schools. Eighteen schools were randomised using an uneven ratio of 2:1 allocation to intervention and control groups. We took this approach so that there were at least ten schools in the intervention group, in line with the developer's preference. Of the 18, one school was randomised due to an administrative error and dropped out without the knowledge of group allocation. This left 12 schools in the intervention group and five in the control group.

Of the 12 intervention schools, four schools withdrew participation from the trial prior to the first professional development session. One school did not realise that it had signed up to a RCT (even after signing a MoU) and three schools approached the delivery team to indicate that they were not planning to operate mixed attainment grouping. In all three of these cases, the schools considered themselves to be operating 'mixed attainment' grouping when they signed up to the trial. Two schools were, in fact, operating with a top and bottom group and a broad 'mixed attainment' in the middle. The third school was operating a novel, flexible, needs-led attainment based grouping, developed from a mixed attainment basic grouping and was not willing to change their practice to completely mixed attainment. As this was a pilot RCT where one of the objectives was to explore whether schools can be recruited to a mixed attainment RCT, it was decided that these four schools would be noted as 'not eligible' for the trial.

In terms of dropout from the intervention (but not necessarily primary outcome measure), two out of eight intervention schools dropped out from the intervention for maths and three out of eight schools dropped out from the intervention for English.

In terms of the primary outcome measures, nine schools took part in GL Assessment's PTM13 and eight schools took part in the PTE13. Sixty pupils were sampled from the entire Year 8 cohort in each school to sit maths and English tests. These pupils were randomly selected from the pupil list provided by the schools which was subsequently matched with the NPD. As FSM eligibility represents an important subgroup, this random selection of pupils was stratified by pupil FSM eligibility. In addition, one school administered the tests to the entire Year 8 cohort, which we permitted to encourage participation in the tests.

Figure 2: Participant flow diagram for BPMA pilot trial



School and pupil characteristics

Table 8 summarises the school and pupil characteristics of the pupils and schools that took part in the primary outcome measurement (GL Assessment tests) in maths¹⁹.

As the primary analysis is school-level regression analysis, we have included a broad range of school characteristics, to facilitate comparison of the final analysed groups. The two groups are similar in terms of most of the background school characteristics we examined. The main noticeable difference between the intervention and control group schools is their Ofsted ratings: all the intervention schools are 'good' compared to one of the control schools being 'outstanding', while the other 'requires improvement'.

Table 8 also presents schools' baseline grouping practices. It is worth noting that the grouping practices of the schools are noticeably different. We carried out simple randomisation as per the protocol rather than minimisation which could have been useful in this instance. As mentioned previously, there were four intervention schools excluded based on eligibility. The baseline grouping practices could have been different if they were retained in the trial.

Table 8: Baseline comparison for analysed groups

Variable	Intervention group		Control group	
	n/N (missing)	Percentage	n/N (missing)	Percentage
School-level (categorical)				
School governance:				
Academy or Free School	3/6 (0)	50%	2/3 (0)	66.7%
Maintained	3/6 (0)	50%	1/3 (0)	33.3%
Ofsted rating:				
Outstanding	0/6 (0)	0%	1/2 (1 ²⁰)	50%
Good	6/6 (0)	100%	0/2 (1)	0%
Requires improvement	0/6 (0)	0%	1/2 (1)	50%
Urban	5/6 (0)	83.3%	2/3 (0)	66.7%
Rural	1/6 (0)	16.7%	1/3 (0)	33.3%
Secondary school type:				
Comprehensive to 16	2/6 (0)	33.3%	2/3 (0)	66.7%
Comprehensive to 18	4/6 (0)	66.7%	0/3 (0)	0%
Other type	0/6 (0)	0%	1/3 (0)	33.3%
School-level (continuous)	n schools (missing)	[Mean]	n schools (missing)	[Mean]
Number of Year 7 pupils in the academic year 2015/2016		185	2 (1)	149

¹⁹ Note that all school characteristics are based on the nine schools that took part in maths primary outcome.

²⁰ This school was converted to an academy in spring term of academic year 2016/17. Prior school and pupil characteristics are not applicable to the new establishment, hence are missing.

	6 (0)			
% pupils eligible for FSM 2015/2016	6 (0)	15.9	2 (1)	13.6
School-level (baseline grouping practices, categorical)	n/N (missing)	Percentage	n/N (missing)	Percentage
Set/stream in Year 7 maths?				
Set and stream				
Set	1/6	16.7%	0/3	0%
Stream	3/6	50.0%	0/3	0%
Neither	2/6	33.3%	0/3	0%
	0/6	0%	3/3	100%
Set/stream in Year 8 maths?				
Set and stream				
Set	1/6	16.7%	0/3	0%
Stream	3/6	50.0%	3/3	100%
Neither	2/6	33.3%	0/3	0%
	0/6	0%	0/3	0%
Set/stream in Year 7 English?				
Set and stream				
Set	0/6	0%	0/3	0%
Stream	2/6	33.3%	0/3	0%
Neither	2/6	33.3%	0/3	0%
	2/6	33.3%	3/3	100%
Set/stream in Year 8 English?				
Set and stream				
Set	0/6	0%	0/3	0%
Stream	2/6	33.3%	1/3	33.3%
Neither	2/6	33.3%	0/3	0%
	2/6	33.3%	2/3	66.7%
Student-level (categorical)	n (missing)	Percentage	n (missing)	Percentage
% eligible for FSM	289 (0)	33.2%	64 (0)	28.1%

(Ever, Spring 2015)				
Student-level (continuous)	n (missing)	[Mean]	n (missing)	[Mean]
Key Stage 2 maths point score	289 (0)	29.6	64 (0)	29.0
Key Stage 2 reading point score ²¹	262 (0)	30.0	66 (0)	30.1

Outcomes analysis

Primary analysis

School level regression was the primary analysis for this trial. We ran two separate regression models, one for maths and the other for English. The maths model included GL Assessment's PTM13 raw scores as the dependent variable, regressed on whether the school was an intervention or a control school and the school's average prior attainment at Key Stage 2²² (as measured by KS2_KS2MATPS variable). We ran a similar model for the English outcome. The outcomes from the models are summarised in Table 9. As seen in the table, there was no evidence of a difference in either of the outcomes between the intervention and control schools. For the maths outcome, there is a non-significant negative association for the intervention group with an effect size of -0.46 (95% CI: -1.39, 0.48, $p=0.277$). For the English outcome, there is a non-significant positive association for the intervention group with an effect size of 0.33 (95% CI: -1.11, 1.77, $p=0.58$).

Table 9: Primary analysis

Outcome	Raw means				Effect size		
	Intervention group		Control group		n in model (intervention; control)	Hedges g (95% CI)	p-value
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)			
Mean Progress in Maths total raw score (PTM13)	6(2)	31.46 (24.10, 38.82)	3(2)	27.73 (8.18, 47.28)	9 (6,3)	-0.46 (-1.39, 0.48)	0.277
Mean Progress in English total raw score (PTE13)	5(3)	32.08 (27.20, 37.0)	3(2)	30.71 (14.18, 47.25)	8 (5,3)	0.33 (-1.11, 1.77)	0.58

As mentioned before, one school used GL Assessment's online progress tests. As these were not the primary outcomes tests (PTM13 and PTE13 paper based tests), we excluded this school from the primary analysis. However, we ran further regression models including this school's outcomes (please see the section on *primary outcome* measures for more details). Outcomes from this sensitivity analysis

²¹ Note this is based on eight schools that took part in the primary outcome measure for English

²² Please note that school's prior attainment was averaged across all the pupils that were included in the final ITT models, i.e. those with a valid outcome measure and a valid prior attainment data.

yielded increased effect size for both the outcomes. The effect size for maths outcome was 0.28 (-1.05, 1.60), $p=0.64$ and the effect size for the English outcome was 0.72 (-0.68, 2.12), $p=0.25$ so they were statistically not significant.

Secondary analysis

As per the SAP, we also ran multilevel models accounting for the cluster randomisation to rehearse them for a full efficacy trial. The multilevel models used similar variables to the primary regression models, where raw scores on PTM13 and PTE13 were regressed on whether the pupil was in the intervention school or the control school and pupil prior attainment at Key Stage 2. Outcomes from this analysis suggested that there is no evidence of impact for either of the outcomes. For the maths model, the raw co-efficient of the intervention group was 1.24 (-7.37, 9.84; $df=7$, $p=0.74$) and for the English model, 1.06 (-5.91, 8.04; $df=6$, $p=0.72$). Effect sizes were not calculated beyond the primary outcome since this was a pilot trial.

Subgroup analysis

As per the SAP, we also explored differential effects using pupil prior attainment (as measured by KS2_KS2MATPS and KS2_KS2READPS variables) and pupil everFSM status (as measured by EVERFSM_6). These models were similar to the multilevel models in the secondary analysis with an additional interaction term included. For example, in the maths model, the additional covariate was the interaction term between the intervention group and prior attainment at Key Stage 2. The outcomes from the maths model suggested that there was no evidence of impact of any interaction. However, the English model had a negative interaction between prior attainment and intervention; this was significant at $p<0.01$). This is implying that the intervention is working in favour of less able pupils (as measured at Key Stage 2) compared to their more able peers. The everFSM interaction model suggested that there was no evidence of impact on either of the outcomes.

As per the EEF guidelines, two further models were run for everFSM pupils only. These were identical to the secondary analysis multilevel models except that they only included the everFSM pupils (for maths model $n=114$ and for English model $n=109$). Results from these models suggest that there was no evidence of impact for this subgroup of pupils on either of the outcomes.

Self-confidence outcomes

As mentioned earlier in the outcome measures section, composite measures on pupil self-confidence in English and maths were created using the baseline pupil survey data. Similar measures were also created from the end-point pupil surveys. Two multilevel models accounting for the cluster randomisation to rehearse them for a full efficacy trial were run using the end point self-confidence measures in English and maths as outcome measures and the baseline measures as covariates. Whether or not the pupil was in an intervention or control school was also included in the models. Outcomes from this analysis suggested that there is no evidence of impact for either of the outcomes. For the maths model, the co-efficient of the intervention group was -0.12 (-0.44, 0.23; $df=6$, $p=0.46$) and for the English model, 0.04 (-0.39, 0.48; $df=6$, $p=0.81$).

Compliance

The delivery team collected data on intervention fidelity. These measures are summarised in Table 4. Findings from this data are summarised in this section. Compliance with the training practices was measured according to whether a school's English or maths department was represented at the six relevant professional development sessions (attendance of expected number and type of staff) and whether the training was cascaded to other members of the department.

Attendance at the professional development events (measure 1, Table 4) varied by school: one attended all six events, and another attended none (this school dropped out of the intervention completely). For English, the lowest attendance for a professional development session was at two schools and the highest was at six schools. The pattern of attendance was similar for the maths departments except that the lowest attendance for a professional development session was at six schools. As cascading the training to other members of the department (measure 2, Table 4) was dependent on the attendance at the training events, this was quite varied too. This ranged from as low as one school cascading the English training to other members of the department after the session to maximum of four doing so for another training session. The cascading for maths training was slightly higher than that of English. This ranged from as low as two schools cascading the training to other members after the event to a maximum of five schools cascading the training from another session.

Compliance with student allocation requirements was one of the fidelity measures (measure 3, Table 4). Student allocation was deemed to be compliant if students were allocated to classes such that there was a similar range of Key Stage 2 results in each class. Data couldn't be collected from two schools that withdrew from the intervention in the first academic year. Of the remaining six schools, two and four schools reported that students were allocated to classes such that there was a similar range of Key Stage 2 results in each class for maths and English respectively although the developer had permitted a number of other arrangements. e.g. nurture groups. Please note that these were not mentioned in the trial protocol and were not part of the intervention handbook.

Whether or not teachers had high expectations of all pupils was asked on two occasions (see measure 4 in Table 4). This data was collected through the teacher questionnaires administered by UCL Institute of Education in November 2015 and June 2017. Six schools provided this information for English and maths on both occasions. For both English and maths, three schools reported that teachers had high expectations of all pupils and three schools reported that teachers did not. This changed slightly in June 2017, where four schools reported that teachers had high expectations of all pupils. This means one additional school reported high expectations of all pupils at follow-up compared to the baseline.

On the same two occasions there was an assessment of whether teachers applied effective differentiation techniques in the classroom (self-reported measures collected via the teacher questionnaires administered by UCL Institute of Education in November 2015 and June 2017, measure 5, Table 4). There was considerable missing data for this measure for English. Of the three schools that provided data in 2015, two reported that effective techniques were applied for English. Three out of the four schools that provided data in 2017 reported that effective techniques were applied for English. More schools provided data for the maths measure. In 2015, four out of six schools that provided data reported that effective techniques were applied for maths. In 2017, one out of five schools that provided data reported that effective techniques were applied for maths.

Grouping practices of all schools

The delivery team collected information from all participating schools about their grouping practices over two academic years. The findings are summarised in this section.

Intervention schools

Six intervention schools provided data on their grouping practices. Half the schools were grouping by completely mixed attainment for English in the first year (when trial pupils were in Year 7) and the other half were employing mixed attainment grouping with a bottom set or a nurture group. In the second year, two schools continued grouping by completely mixed attainment and one school moved from mixed attainment to setting. Three schools that were employing mixed attainment with a bottom set or a nurture group in the first year continued with this practice in the second year. For maths, the picture is similar to that of English in the first year: three schools were grouping by completely mixed attainment

and the other three were employing mixed attainment with a bottom set or a nurture group. Three schools that were grouping by mixed attainment in the first year had changed their practice to *partially or completely grouping by setting*. Three schools that were employing mixed attainment with a bottom set or a nurture group in the first year continued with this practice in the second year.

Control schools

Not all of the five control schools took part in this data collection. Two schools took part in the first year data collection, both of which were grouping by completely mixed attainment for both English and maths. Four schools took part in the data collection in the second year. Of these, three were grouping by mixed attainment in English and one was grouping by mixed attainment with a top and a bottom set in English. For maths, three schools were using setting and one school was grouping by completely mixed attainment in the second year.

Across both the groups, one out of ten schools was grouping by completely mixed attainment in Year 8 maths. In contrast, five out of ten schools were grouping by completely mixed attainment in Year 8 English.

Cost analysis

How much does it cost?

The intervention cost a total of £3488 per department per school or £16.37 per student for the first year of implementation. The main financial costs were the developers' staff and monetary costs (such as staff travel and venue hire) for delivering the training, together with teacher participants' travel and subsistence and printing and photocopying costs. In addition, participating schools reported using an average of three hours of cover per department in the first year, so that staff could take part in the initiative. The annual cost for the second year of implementation was £3962 per department or £18.60 per student per department. As this was a two year intervention, and the training was cascaded to staff across participating departments, we have assumed there would be no additional costs to schools if it was continued into a third year. The cumulative cost per student per year equates to an EEF cost rating of 'very low'.

Cost per pupil per year

The cost of the intervention for this evaluation, presented below, has been estimated on the basis of costs that were part of the project. While it provides a rough estimate of the cost that might be expected, the costs do not necessarily represent what the cost would be to a school if the intervention was scaled-up, for example, if the staff delivering the intervention changed.

This cost evaluation estimates the cost to schools under the assumption that the EEF funding for financial costs, pre-requisite costs and compensation for staff time is not being provided and that schools are paying for their share of the total costs they would otherwise bear. As the intervention requires schools to modify their existing practice, we have worked on the assumption that there were no additional pre-requisite costs.

Financial costs

The single main financial costs of delivering the intervention, borne by EEF as part of this project, were the UCL Institute of Education staff costs associated with delivering training, amounting to £252,262 in 2015/2016, and £260,280 in 2016/2017. Other direct, marginal costs, such as staff travel costs and venue hire, were also recorded, and are presented in Table 10 below. As the same UCL Institute of Education team was working on both the BPMA pilot study and the BPS efficacy trial, this information was aggregated across both interventions. However, it has been possible to split direct marginal costs

across both interventions using the ratio 0.983 (for the BPS trial) and 0.017 (for the BPMA pilot study). These ratios represent the relevant award of funds for the two interventions from the EEF grant, and have been agreed with the developer. Unfortunately, it has not been possible to provide a similar breakdown to the staff costs, as the delivery team have recorded their time as working on one holistic project, rather than as standalone contributions to two separate interventions. The staff costs for delivering the BPMA pilot study are therefore overinflated, as the costs include time required to deliver the BPS intervention. To calculate a more accurate cost per department per year in the cost per student per year calculations, UCL Institute of education staff costs were divided by the total number of departments participating in both the BPS trial and BPMA pilot trial.

Table 10: UCL Institute of Education’s financial costs for delivering the BPMA pilot for the academic years 2015/2016, 2016/2017

	2015/16	2016/17
UCL Institute of Education Staff travel	£38.30	£22.99
UCL Institute of Education Staff hospitality	£12.85	£4.23
Teacher travel and subsistence	£45.58	£5.76
Professional development venue and catering	£114.45	£9.15
Professional development additional delivery support	£10.63	£20.99
Print materials	£85.32	£0.00
Web materials	£26.01	£0.00
Total	£333.14	£63.12

We also asked schools whether they had incurred additional costs as part of their implementation of the intervention. We received responses from five schools to these proforma questions in 2015/16, and responses from two schools in 2016/2017. Schools’ responses indicated that little additional financial expenditure was required, particularly in the second year of implementation. Using closed response options, those that responded cited the costs of:

- printing/photocopy or material costs, which equated to an average of £115 per department in Year 1 and £12.50 in Year 2; and
- travel and subsistence costs (related to CPD, not delivered by UCL), which equated to an average of £65 in Year 1, and £0 in Year 2.

We include these figures in our estimate of total cost, although it is a rough estimate, given the limited information it is based on.

UCL Institute of Education collected information about when the schools dropped out from intervention. This was collected and recorded between the professional development sessions. On these bases, we calculated the number of schools that were still continuing to participate in each intervention in each academic year. Table 11 shows our estimate of the cost per student per department (English or mathematics) over time, the cumulative cost per student over three years, and the average cost per student per year. We derived the cost per student by dividing the average cost per department per school per year by the average number of Year 7 students per school (using student administrative data collected by NFER). This calculation yields a total cost per student over three years of £34.98 and

an average cost per student per year of £11.66. The cumulative cost per student per year equates to an EEF cost rating of 'very low'.

Table 11: Average financial cost per student per year

Cost item	Cost per student	Cumulative cost per student	Average cost per student per year
First year	£16.37	£16.37	£16.37
Second year	£18.60	£34.98	£17.49
Third year	£0.00	£34.98	£11.66

Staff time

The developer's original intention was that the BPMA pilot would involve staff attending three full days of CPD and three twilight sessions. It was expected that two Year 7 English and two Year 7 mathematics teachers would attend each CPD session, and that, where possible, this should include the head of English and the head of Mathematics. However, instead, UCL Institute of Education held four full days and two twilight sessions for the London hub and three full days plus two twilights for the York hub (one full day was cancelled due to industrial action by a teaching union). The reason for moving to an additional full day was because attendance at twilight events was found to be a real struggle for schools. Data supplied by the developer suggested that each full day's PD session lasted for approximately five hours, while each twilight session lasted for approximately two hours. Data was not collected on participants' average travel time to each session. As some of the training was undertaken in the school day, schools were able to claim for supply cover which was paid for by the EEF grant. This totalled £9,094.17 in 2015/2016 (or an average of £758 per department) and £3,400.00 in 2016/2017 (or £309 per department). As per the guidance for EEF cost evaluations, these costs have not been included in our cost per student per year calculations. Excluding participation in the professional development sessions, participating schools also reported using an average of three hours of cover per department in the first year of the intervention so that staff could take part in the intervention. Of this time, approximately 1.5 hours on average per department per year required paid supply cover. No use of cover time was reported in the second year of the intervention. In summary, the overall costs of participating in the intervention incurred by a typical mathematics or English department is summarised below.

Table 12: Average costs directly incurred by participating mathematics and English departments

	Printing/photocopy or material costs (£)	Travel and subsistence costs (related to CPD, not delivered by UCL) (£)	Supply cover (£)
First year (2015/2016)	115.00	65.00	758.00
Second year (2016/2017)	12.50	0.00	309.00

Process evaluation

Implementation

Most interviewees reported that the intervention was different to their usual approaches to student grouping at Key Stage 3.

Respondents from five of the six schools interviewed reported that their departments had previously set by attainment. Only one interviewee, from a mathematics department, reported that prior to the intervention, about half of his Key Stage 3 students were in mixed attainment groups, but this still suggests there was some degree of setting by prior attainment.

Interviewees reported that their schools had decided to participate in the trial for a range of reasons.

Understanding the motivations of schools for becoming involved in the intervention could have implications for the future scale-up of the intervention. The main reasons given by interviewees included:

- concerns about the performance of students in low attainment sets, and the desire to try a different approach to grouping students (three interviewees)
- increasing student numbers and reducing school budgets meaning there was a need to increase student class sizes (teaching students in mixed attainment groups was thought to be more appropriate in this arrangement) (one interviewee)
- a need to free up teacher capacity at Key Stage 3 so that students could be taught in attainment sets at Key Stage 4 (one interviewee)
- interest in mastery approaches to learning²³, particularly in mathematics, and the belief that a mixed attainment grouping model lent itself to this approach; and
- some schools were already planning to move to mixed attainment groups, and thought the intervention provided the support required to help smooth this transition (two interviewees).

Perhaps not surprisingly, in all cases, school senior leaders appeared to make the final decision as to whether the school joined the intervention or not. However, it was not always clear to what extent there had been discussion and agreement with the relevant heads of departments. Findings reported later in this report, suggest that some English and mathematics teachers did not fully support the intervention, perhaps suggesting that more could have been done by senior leaders to persuade them of the potential benefits of mixed attainment grouping, or to have involved them more in the decision to join the trial.

Interviewees had mixed views on the quality of the training sessions.

As noted earlier in the report, the developer's original intention was that the BPMA pilot would involve staff attending three full days of CPD (delivered by the KCL (UCL Institute of Education) delivery team in school time, as part of regional workshops) and three twilight sessions. However, instead, UCL Institute of Education held four full days and two twilight sessions for the London hub and three full days plus two twilights for the York hub (one twilight was cancelled due to industrial action by a teaching union).

While most interviewees appeared to enjoy the first regional workshop, there was a general feeling that later sessions were too repetitive, and did not move the learning on sufficiently.

²³ Approaches whereby teachers spend more time going into depth about a subject as opposed to moving quickly through the things that all children are expected to know.

A senior school leader, commenting on the involvement of the English and mathematics departments that were both participating in the intervention, spoke about their differing responses to the training sessions, which he attributed to teachers' respective prior experiences of working with mixed-attainment groups:

Compared to the maths department, some staff from the English department are not very enthusiastic about the initiative, because they already have experience of working with mixed attainment groups. I think the delivery organisation needs to take some responsibility for this, as they need to provide training and materials that are more compelling for teachers and which more effectively encourage them to take part.

However, the same interviewee thought the training handbook was 'excellent', and was very happy to encourage its circulation to those who could not attend the training session.

Another interviewee, a head of mathematics, reported that she would have liked 'more practical tasks' as part of the training sessions, and gave the following example: 'It would have been useful if everyone had brought in a mixed attainment lesson plan and then for us to have gone through and reviewed them as a group.'

Fidelity²⁴

The process evaluation explored the extent to which the intervention had been delivered as intended by the developers. The following section discusses each fidelity measure in turn, based on the views of interviewees.

Attendance at training

Most interviewees indicated that their departments had been represented at each training session, although the number of staff attending each session varied.

It was expected that two Year 7 English and two Year 7 mathematics teachers would attend each regional professional development session and that, where possible, this should include the heads of each department. Most interviewees indicated that one or two staff had attended each session. This usually involved the head of department, but sometimes another subject specialist would attend and/or a school senior leader.

One interviewee, a head of English, reported that she had not been able to attend one training session because the school had said there were too many staff out of school on the same day. In that instance, the head of mathematics had been able to attend and was able to share the training materials/notes from the event. Another interviewee, a head of English, reported that she had missed one event in London because it was 'too far to get to'. However, for most interviewees, attendance at the training sessions did not appear to be a problem. Overall, there appeared to be a moderate level of fidelity with regard to this measure.

Cascading training

Most interviewees reported cascading the training to the other members of their departments, as required by the intervention, although a minority reported doing this on a limited basis only.

²⁴ Note that findings presented in this section are based on the interviews undertaken by the evaluator. Data collected by the developer is reported in the 'Compliance' section

Staff in five of the six schools that we spoke to confirmed that all of the training and materials had been shared with other colleagues in their departments. The main method of doing this was through department meetings or school twilight CPD sessions. Electronic versions of training materials/handouts were circulated to staff by email.

Interviewees' responses suggested that teachers were engaged by this sharing of information to varying degrees. For example, one head of mathematics reported that every member of staff was actively engaged: 'We got every colleague to produce an activity in the style of the one shown to the teachers in the first meeting'. Other interviewees were more frustrated with the level of engagement of their colleagues, due to logistical or other reasons. For example, one head of English reported that he would have liked more opportunities for staff to observe those who attended the training putting it into action in the classroom, but had found this challenging due to timetabling constraints. Two interviewees, both heads of mathematics, had found some resistance from their staff, as illustrated by this comment: 'The main challenge with cascading the learning stemmed from the fact that there was reluctance from some staff within the department to buy into mixed-attainment grouping. We've never really worked out how to overcome this.'

Overall, there appeared to be a moderate level of fidelity with regard to this measure.

Allocation of students to classes

It was expected that each mixed attainment class would include students with a similar range of Key Stage 2 results.

Most interviewees confirmed that their departments had allocated students to classes so that each had a similar range in terms of student prior attainment, although factors other than prior attainment, were sometimes taken into consideration.

Interviewees from all but one of the six schools we spoke to said that students were allocated to classes to ensure that each had a similar range of students in terms of their Key Stage 2 results. However, the heads of English and mathematics from one school reported that in addition to grouping students on the basis of their Key Stage 2 national curriculum results in English and mathematics, they also took into consideration other factors, such as behaviour and friendship groups.

One head of English, in a department that had implemented the intervention in all Year 7 and 8 classes, described how all Year 7 and 8 students were in classes with a similar mixture of student prior attainment with the exception of a 'nurture group', which included students with 'significant special educational needs and disability (SEND) or who had low scores on CATs²⁵'.

One interviewee, a head of mathematics, reported that instead of grouping by mixed attainment, his department had adopted what he called 'near ability' setting in year two of the trial, with the intention of moving to full mixed attainment in year three:

We've been using attainment setting for so long that to go straight to full mixed attainment in one year would have been too much for the staff. This year we have done 'near ability' setting, with the top two classes mixed, and the bottom two classes mixed. Next year we will move to full mixed attainment.

²⁵ Cognitive abilities tests

Having spoken to the same interviewee in year three of the evaluation, it transpired that the trial cohort, then in Year 8, had continued to be taught in 'near ability' classes, while the new Year 7 cohort were taught in mixed attainment classes.

Overall, there appeared to be a moderate level of fidelity with regard to this measure.

Teachers having high expectations for their students

All interviewees reported that their teachers had high expectations for their students, as expected by the intervention, and that they had systems in place to ensure that these expectations were maintained.

These systems varied, but were often reported to come from a range of measures or features including:

- school culture or learning environment—senior school leaders (as distinct from middle leaders or class teachers) were seen to play a major role in establishing and maintaining the expectations against which staff and students were judged; students were taught that getting something wrong is not a bad thing, and can provide an opportunity to learn; staff enforced a 'can do' attitude in students and praised effort, as distinct from ability and starting point;
- pedagogy—setting aspirational targets for all students, regardless of prior attainment; 'teaching high and differentiating down'; 'the use of mastery approaches; and
- monitoring—through lesson study, learning walks, peer observation and feedback.

While some interviewees acknowledged that some teachers were better at this than others, they all felt that they had processes in place designed to encourage their teachers to have high expectations for all students.

One head of mathematics reported that his students undertook 'peer critique and re-drafting training' which was said to encourage a culture of excellence. However, when asked if all teachers applied these methods, the interviewee said that about half did, with staff 'getting out of the intervention what they put in'.

One interviewee, a head of English, gave the example of using step-by-step tasks and challenge cards to help set high expectations for all students:

We teach the final step first, and then the students do tasks that are suitable for their level. So when we're looking at semi colons, the first step is to identify a semi colon, and the last step is to create a sentence with a semi colon in it. If a student completes the task they can then do the challenge card

A head of mathematics reported that fewer parents had raised concerns about their children's progress since the move to mixed attainment grouping:

At Year 7 parents evening we didn't have anything like the normal number of parental concerns over [student] progress. I think this shows that the students are being challenged [in mixed attainment groups].

Another head of mathematics said he thought the guidance provided by the developers on encouraging a culture of high expectations felt 'a bit abstract' and did not tell them anything that 'we weren't already doing'. However, no other interviewee commented on this aspect of the training.

Overall, there appeared to be a high level of fidelity with regard to this measure.

Use of effective differentiation techniques

Most interviewees reported that their teachers applied effective differentiation techniques in the classroom, as expected by the intervention.

These techniques included:

- giving students choice over the level of difficulty of the tasks they completed in class which was found to be self-motivating;
- the use of seating plans to ensure all class tables were mixed attainment, which encourages high ability students to support their lower ability peers;
- use of colour coding systems to delineate the level of difficulty for students (all students were encouraged to aim for the hardest level of difficulty); and
- differentiating homework by task²⁶, with a mastery of skills focus.

Interviewees reported that they monitored this work through peer observation, learning walks and book scrutiny. Staff were supported by their heads of department, and time was set aside in staff meetings to discuss and reflect on the differentiation techniques being used. Some staff also had the opportunity to attend externally run CPD that the school had bought into²⁷ (that is, in addition to the training offered by the developer).

A small number of interviewees identified barriers or challenges associated with differentiation. One of these barriers was staff capacity, with one head of mathematics reporting that she did not have a deputy head of department at the time of the interview, which meant that she had not been able to support her colleagues with differentiation at the level she would have liked. Two interviewees, both heads of mathematics, reported that some staff had failed to 'buy in' to the initiative, and that this meant they were less likely to apply effective differentiation techniques in the classroom. Another interviewee, a school senior leader, reported he would have liked 'more fine grained detail' from the developers on effective teaching for differentiation, which moved beyond 'broad principles'.

Although not reported as a challenge, one head of mathematics reported that a colleague had grouped her students in such a way that she had one high ability table, and one low ability table in her class. This approach was described as helping to support differentiation, but it is possible that this is also moving closer to setting by attainment group rather than mixed attainment.

Overall, there appeared to be a moderate level of fidelity with regard to this measure.

²⁶ It should be noted that differentiation by task was actively discouraged by the delivery team.

²⁷ Information about the subject of the CPD was not explored further in the process interviews.

Outcomes

Perceived impact on teachers

As reported earlier, interviewees' responses suggested that both English and mathematics teachers had responded to the intervention with mixed reactions, with some enjoying it, and others really struggling to teach in mixed attainment sets.

One head of English reported that the intervention appeared to be beneficial for both newly qualified teachers, and staff who had been teaching for longer:

For those [teachers] fresh out of training, we have noticed an improvement in their teaching through lesson observations. It has also improved the subject knowledge of some other teachers as they have had to plan to differentiate so widely that they have had to go into more depth.

Another interviewee, a head of mathematics, raised a question about the prerequisite skills and the disposition of teachers required to effectively teach students in mixed attainment groups: 'For me, you can only deliver the intervention effectively if you are a good teacher. I think it makes you a better teacher, but only if you fully support it'.

A school senior leader, interviewed in year three of the evaluation, reported that he had observed both English and mathematics teachers reflecting more on their own practice, which he thought could 'only be a good thing'.

Perceived impact on students

Interviewees were asked, what impact, if any, they thought the intervention had on their students. Whilst acknowledging that teachers had mixed views on the intervention (see section above), most interviewees reported that the intervention was having a positive impact on their students. This was the case for both English and mathematics teachers. Specific benefits identified by interviewees included:

- the same high-level of support was now being given to all students, regardless of their prior academic attainment;
- students were no longer being 'pigeonholed' for being in a lower set;
- more able students were being stretched further by deepening their learning rather than moving them on to new topics (that is, adopting a 'mastery approach', particularly in mathematics); and
- more students with low prior academic attainment now believe they can achieve as a result of moving away from setting by attainment.

Most interviewees appeared to think that the lowest attaining students had particularly benefitted from the intervention, with one interviewee reporting: 'they feel comfortable, they feel they can make comments and they don't worry about being wrong'. Another interviewee extended the benefits to 'middle to lower achieving boys'. *'I think probably the middle to lower boys are benefitting the most as they are being stretched by being teamed up with some particularly bright girls.'* It should be noted that, as outlined in the delivery team's 'Theory of Change' (see Appendix C), teachers were expected to cultivate a growth mind-set in students, which in turn was expected to result in students investing more effort in their learning. While this approach is perhaps not reflected in the language used to describe students in the quote above, attempts were made by the developers to encourage teachers not to label students, as this can result in a 'self-fulfilling prophecy' whereby initially erroneous beliefs about a student's potential can result in their fulfilment. Should the intervention go to trial, the developer's might consider including a measure of teacher beliefs about pupil ability.

One interviewee exclaimed that, 'When it's done right, this work can make a real difference to learners', while another noted that since introducing mixed attainment teaching he had been able to enter more

students into a national mathematics competition, because they all felt confident to take part. In previous years, only a few children had been entered because only a few felt confident enough. This, the interviewee thought, was testament to the fact that the new approach was working.

Views on scale-up

Interviewees were asked whether, in its current form, the BPMA intervention could be used by other schools, and was therefore suitable for scale-up. Six of the nine interviewees reported that they thought the intervention was scalable, while one said no, and two were unsure.

One school senior leader commented that the intervention worked well with the ‘mastery approach’²⁸ to learning, but that teachers needed to support it from the outset, and should not rely on it alone to change the views of teachers who were more resistant to mixed-ability teaching:

The initiative links nicely into maths mastery. It feels timely and I think it would be popular with other schools. However, the intervention will only work with schools that want to change. The training is not sufficiently compelling to persuade teachers that don't want to buy into it and schools would be wise to select it on this basis.

Some interviewees commented on the additional workload and challenges of differentiation that they associated with the move from setting by attainment to grouping by mixed-ability: ‘The intervention has caused a bit of discomfort and stress to teachers. If we were offered to go back to setting by attainment, we would probably take it.’

Formative findings

Suggestions for improvement

A number of suggestions were made for how the intervention could be improved (from both those who thought the intervention was ready for scale-up, and those that did not or who were unsure). These included (in no particular order):

- reducing the travel time for participants by delivering all the training in school;
- removing the repetition from later training sessions by spending less time recapping what had already been taught and introducing more new content; and
- reducing the workload on participating teachers by providing more practical support and resources on how to teach mixed-ability groups, such as an exemplar scheme of work.

Feasibility

This section explores how attractive the intervention was to teacher participants and the extent to which they judged it to be practical and feasible.

Interviewees reported mixed reactions from staff.

Of the nine departments represented in our telephone interviews, five department heads felt described how their staff were mostly positive about it, two were mostly negative, and two had mixed views. This range of responses was illustrated by the quotation from the following head of mathematics, who suggested that staff concerns about the intervention could sometimes be of a personal nature: ‘We have the full range of opinions, from those that think this is great, to those who really struggle with it,

²⁸ Approaches whereby teachers spend more time going into depth about a subject as opposed to moving quickly through the things that all children are expected to know.

and everything in-between. Sometimes the concerns can be very personal, with some staff asking themselves: ‘Can I teach this?’”

Of the five departments who said their staff were mostly positive about the intervention, four were from the same two schools (both their English and mathematics departments were participating in the trial). Similarly, the two departments who were negative about mixed-attainment grouping were also from the same school (again, both their English and mathematics departments were participating in the trial). These findings suggest that teacher receptiveness to the trial could be school, rather than department, led.

One head of English attributed his department’s lack of buy-in to the intervention to the fact that most of his staff had only ever taught in attainment sets, and thought this was the best and only way to teach. Another suggested that the age of his teachers appeared to affect their views of the intervention: ‘It’s almost like an age thing, with the under 40s quite receptive, and the over 40s not so [receptive]’.

Supporting factors

Interviewees identified a number of supporting factors associated with their approach to implementation, that they believed helped the intervention to run smoothly. These included:

- sharing the planning work amongst staff;
- staff being open-minded about the intervention;
- using student self- and peer-assessment to promote shared standards and expectations;
- supporting staff by providing opportunities to discuss and share their experiences of teaching mixed attainment classes; and
- having a whole-year approach so that all subjects in Years 7 and 8 are taught in mixed-ability classes (not just English and mathematics).

One interviewee, a head of mathematics, reported that she had shared the task of lesson planning, required as part of the move to mixed attainment grouping, amongst her department. This distributed approach to planning made the task more manageable, but the interviewee still described it as a ‘massive undertaking’. Similarly, another interviewee, also a head of mathematics, was aware that other schools were sharing the workload in this way, and thought this had helped them to secure staff buy-in. He had not been able to secure the same support from his own department, and would have welcomed greater guidance from the developers on how to have planned for these changes or, to have been provided with an exemplar scheme of work.

Barriers and challenges

In response to a direct question about the challenges associated with implementation, a number of interviewees identified the following issues:

- poorly stratified classes with an unequal distribution of students with high and/or low prior attainment;
- lack of teaching assistants to work with students who would benefit from additional one-to-one support;
- staff not buying into the intervention/giving it a chance;
- the additional workload associated with the move from setting by attainment to grouping students by mixed attainment.

One interviewee, a head of English, reported that the increased differentiation required to teach mixed attainment groups was both a success and a challenge, as it had improved the quality of teaching but had also created a lot more work for teachers.

Another head of English reported that his students were regularly interviewed about how they were getting on, and that this had revealed some issues with the most able students, who felt they were not being sufficiently challenged in mixed attainment groups. As a result, the English department had focused more effort on stretching these students, while additional activities, such as school trips, were also being run to help better cater for their needs.

Other challenges associated with implementation of the intervention surfaced during the interviews and are discussed in the sections above.

Conclusion

Key conclusions

1. It was challenging to recruit and retain schools to the pilot trial. Despite directly contacting 158 schools and widely advertising the opportunity to another 330, only 18 agreed to be randomly allocated to either receive the training in mixed attainment grouping or to be a control group.
2. Staff had mixed experiences of the intervention; some enjoyed it, whereas others struggled, particularly with differentiation in mixed attainment groups. Schools that continued with the intervention generally adhered to the programme: allocating pupils to mixed attainment classes, applying differentiation techniques in the classroom, and communicating high expectations for all pupils.
3. Most interviewees felt that the intervention had a positive effect on pupil outcomes and that those with low prior attainment particularly benefitted.
4. The pilot RCT was small and designed to test whether a trial was possible rather than to measure the impact of the intervention. The outcome data that was collected did not show a difference in overall maths and English scores between intervention and control schools.
5. Should a future efficacy trial be considered, particular attention must be paid to eligibility criteria, clarity of expectations at recruitment and the teacher workload associated with implementing mixed attainment teaching.

Interpretation

The statistical analysis conducted during this pilot RCT was carried out in rehearsal for a future trial, not with the intention of deriving conclusions about the impact of the intervention. The trial was based on only nine schools for the maths results and eight for English. As anticipated at the design stage, we cannot conclude anything from the impact results. Whilst the significant prior attainment interaction term in the English model was indicative of helping the less able at the expense of the more able, this must be interpreted in the context of carrying out several other statistical tests that were done across the primary, secondary and subgroup analyses. It is unlikely to mean anything other than a possible avenue to explore were a future larger-scale evaluation to be planned. There were also limitations to the data collected from the process evaluation, as it was difficult to persuade school staff to agree to interviews, particularly as the trial progressed. That said, this element contains the most useful outcomes from this study and should be reviewed carefully by anyone considering influencing the mixed attainment teaching practices of secondary schools in the future.

Formative findings

Of the eight intervention schools, five continued with the English intervention and six with the maths intervention. The fidelity findings are based on schools that continued with the intervention and agreed to be interviewed. The findings concerning how well intervention schools adhered to their programme of mixed attainment teaching are broadly positive, suggesting the delivery team worked effectively with the schools during implementation.

The intervention could be strengthened further, particularly by providing more practical resources for schools, such as lesson plans and exemplar schemes of work, which might help reduce the impacts of the intervention on teacher workload. The other major factor is the need to ensure that the staff in participating English and mathematics departments are 'signed-up' to the intervention. Interviewees who reported their colleagues were 'ready to give it a go', and who shared the work of developing lesson plans for mixed attainment groups, were more likely to report that the intervention was both practical and feasible.

Readiness for trial

This pilot trial demonstrates that, with extensive reminder activities, secondary schools can be recruited to a mixed attainment RCT. However, four dropped out immediately when they fully understood what they had signed up to and only six (five for English) out of the original twelve schools randomised to intervention carried out their post-test. It is interesting to note that there were control schools (that stayed in the trial) with similar grouping practices to those of the four intervention schools that withdrew immediately after randomisation, that is largely employing mixed attainment grouping but with a single top and bottom set. This emphasises the key conclusion that eligibility criteria are important. Had this been an efficacy trial, such withdrawal would have introduced bias since equivalent control schools did not withdraw.

This trial was originally conceived as an opportunity for schools to embrace mixed attainment teaching as an alternative to setting or streaming. By the time the protocol was written and the feasibility study completed, both the developers and evaluators realised that it would not normally be possible to change the grouping practices of a secondary school in such a radical fashion. Instead, it was acknowledged that the kind of school that would be interested in being involved in this research might already be doing mixed attainment teaching. Therefore, no eligibility criteria were applied at recruitment except that the school should be willing to implement grouping by mixed attainment if randomised to the intervention. The data collected at baseline for schools signing up to the trial indicated that schools were using a variety of grouping practices—there was no consistent picture. Furthermore, data collected from controls during the trial itself by UCL Institute of Education also revealed a variety of different grouping practices.

The trial therefore became a comparison of the enhanced mixed attainment teaching practices of the intervention against a control group that exhibited a reasonable amount of mixed attainment teaching itself, particularly in the case of English departments. Furthermore, even with these relaxed eligibility criteria, it was not possible to recruit the target of 20 schools and only 18 (17 discounting the administrative error) were randomised. This pilot RCT has demonstrated that despite the research evidence in its favour (EEF, n.d.a), schools are largely reluctant to participate in an evaluation of mixed attainment teaching; as seen in Figure 2, the difference between the numbers originally planned and the ones analysed demonstrates this. Despite combined efforts of UCL Institute of Education and NFER, our recruitment target of 20 secondary schools was not achieved. Of the 18 secondary schools that were recruited, four intervention schools dropped out on the basis of eligibility and a further four schools dropped out of the maths outcome and five schools dropped out of the English outcome. Once recruited to the trial, many of these enthusiastic schools stay the course in terms of intervention delivery but do not necessarily agree to follow-up testing. We suggest that should an efficacy trial of this intervention be considered, particular attention is paid to eligibility criteria (including different setting practices in maths and English), clarity of expectations at recruitment and the teacher workload associated with mixed attainment teaching. We believe the difficulties experienced during this pilot are surmountable in future evaluation work and that such work is important given the common occurrence of setting in English secondary schools.

References

- Cooper, D. (2011). *Redefining Fair: How To Plan, Assess, And Grade For Excellence In Mixed-Ability Classrooms*. Bloomington, Indiana: Solution Tree.
- Education Endowment Foundation (n.d.a). 'Setting or Streaming'. In: *Teaching and Learning Toolkit* [online]. Available: <https://educationendowmentfoundation.org.uk/resources/teaching-learning-toolkit/setting-or-streaming> [15 November, 2017].
- Education Endowment Foundation (n.d.b). *Consent and the Data Protection Act: Advice for Evaluators* [online]. Available: https://v1.educationendowmentfoundation.org.uk/uploads/pdf/EEF_guidance_for_evaluators_on_consent_and_the_Data_Protection_Act_FINAL.pdf
- Education Endowment Foundation (2017). *Statistical Analysis Plan for Best Practice in Setting. NFER* [online]. Available: https://educationendowmentfoundation.org.uk/public/files/Projects/Round_6_-_Ability_grouping_BPS_SAP.pdf [30 November, 2017].
- Husbands, C. (2014). 'Setting by ability: What is the evidence?' *IOE London Blog*, 3 September, 2014 [online]. Available: <https://ioelondonblog.wordpress.com/2014/09/04/setting-by-ability-what-is-the-evidence/> [15 November, 2017].
- International Association for the Evaluation of Educational Achievement (IEA). (2011). *TIMSS 2011 Student Questionnaire*. IEA: Boston.
- Kutnick, P., Sebba, J., Blatchford, P., Galton, M., Thorp, J., MacIntyre, H., & Berdondini, L. (2005). *The Effects of Pupil Grouping: Literature Review* (RR688). London: DFES [online]. Available: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.127.9357&rep=rep1&type=pdf> [26 October, 2017].
- Marsh, H.W. (1990). 'The structure of academic self-concept: The Marsh/Shavelson model'. *Journal of Educational Psychology*, **82**, 4, 623–636 [online]. Available: <http://www.grajfoner.com/Clanki/Marsh1990JoEP%20Academic%20Self%20Concept.pdf> [15 November, 2017].
- Marsh, H. W. (1990). *The Self-Description Questionnaire II Manual*. Australia: University of Western Sydney: MacArthur.
- Organisation for Economic Co-operation and Development (2012). *OECD Program for International Student Assessment 2012. Student Questionnaire*. OECD Publishing, Paris.
- Taylor, B., Travers, M., Francis, B., Hodgen, J. and Sumner, C. (2015). *Best Practice in Mixed-Attainment Grouping*. London: Education Endowment Foundation/King's College London.
- The Office for Standards in Education, Children's Services and Skills (Ofsted) (2013). *The most able student students: Are they doing as well as they should in our non-selective secondary schools?* [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/405518/The_most_able_students.pdf [12 January, 2018].
- Wilshaw, M. (2016). Speech at the 'FASNA Autumn Conference', 2 November [online]. Available: <https://www.gov.uk/government/speeches/sir-michael-wilshaws-speech-at-the-fasna-autumn-conference> [15 November, 2017].

Appendix A: Recruitment material

BPMA Frequently Asked Questions (on NFER website)

Evaluation of best practice in grouping students Mixed attainment initiative

Purpose

Why are you doing this research?

'Ability grouping' has been a long-standing topic of contention. Evidence suggests that i) students from disadvantaged backgrounds are over-represented in low sets and streams, and ii) that students in these low groups make less progress than students in higher sets/streams, and than their counterparts in mixed attainment classes. However, the research is less clear on why this is the case, and whether it can be addressed by better practice. The purpose of the study is to evaluate the impact of different, research-informed approaches to grouping students, on students' learning outcomes. It will also explore how outcomes vary between different students according to factors such as student background, prior attainment, student gender and ethnicity.

Who will benefit from this research?

We hope that students in schools across the country will benefit from this research as it will inform education policies and practices in England. We also hope that your school will benefit, via access to high quality CPD and research-informed practice designed to support progress and achievement for all students.

Who is carrying out the research?

The research is being carried out by a team from King's College London (KCL), led by Professor Becky Francis. It is funded by the Education Endowment Foundation (EEF) and is being independently evaluated by the National Foundation for Education Research (NFER). NFER are also helping to recruit schools to the study.

Roles

What is NFER's role in the study?

NFER are evaluating the project. They will randomly allocate schools to the intervention or control group. NFER will conduct telephone interviews with some teachers. NFER will administer tests developed by GL Assessment in English ('Progress Test in English') and maths ('Progress Test in Maths') to a sample of students at the end of Year 8. NFER will supply additional copies of the tests to schools on request. NFER will use the data from the tests and staff questionnaires to evaluate the intervention and will disseminate the headline findings of the intervention. NFER are also helping to recruit schools to the study.

What is the Education Endowment Foundation's role in the project?

EEF are funding the project and will help to disseminate the findings.

What is KCL's role in the project?

KCL has developed and will be delivering the intervention to schools. They will also be undertaking their own research, although this is separate to the evaluation NFER will be undertaking.

What does the school need to do as part of the project?

What the school does depends on whether it is allocated to the intervention or control group.

All schools will need to provide the research teams at KCL and NFER with data about their students which will enable NFER to administer the GL Assessment progress tests to a sample of Year 8 students in summer 2017. Schools will need to facilitate students and teachers completing questionnaire surveys and may need to make students and staff available for interviews.

Schools will need to follow instructions for organising and teaching students in mixed-attainment groups according to evidence-based best practice for teaching students in mixed-attainment groups. Students will be organised into mixed-attainment groups on the basis of their KS2 results. Schools will need to send two English teachers and two maths teachers to each of the twilight professional development sessions (three full days and three twilight sessions). Teachers will then implement the principles for best practice teaching in mixed-attainment groups in their own lessons.

Schools in the control group continue with their usual teaching practices.

Eligibility

We currently stream students – are we eligible to participate?

If you would be willing to change your practice to mixed-attainment grouping, you would be able to participate. Participant schools are not allowed to apply streaming.

We divide our students into two (or more) bands by ability and group within these – are we eligible to participate?

If you would be willing to change your practice to mixed-attainment grouping, you would be able to participate. Participant schools are not allowed to apply banding.

Timeline

When is the deadline for committing to the project?

The deadline for committing to the project is Friday 12th June 2015 but we would like schools to commit to the project as soon as possible.

When does the study start?

Schools will be randomly allocated to the intervention and control groups in June 2015. The first professional development events for the intervention will be in July 2015. The intervention applies to students who will start in Year 7 in September 2015.

When does the study end?

The study will end in July 2017, when the participating students have reached the end of Year 8. The final report will be published on the EEF website in early 2018.

When will we know if we have been allocated to the control or intervention group?

Randomised allocation to intervention and control groups will take place in June 2015. You will be informed soon after.

Professional development

When will the professional development events be?

Professional development events will begin in July 2015. Schools allocated to the intervention group will be given a calendar of events and venues in June 2015.

Where will the professional development events take place?

Professional development events will take place at regional venues. These will be confirmed in June 2015.

How many professional development events will there be?

The intervention consists of three full days and three twilight professional development sessions over two years.

Data collection

What tests will you be using to assess students' progress in English and maths?

We will be using GL Assessment's new Progress Test in Maths and Progress Test in English. These are designed for Year 8 students and correspond to the new National Curriculum.

What data will be collected?

Background data about students will be collected through the National Pupil Database. To enable us to undertake the analysis we will require year 7 pupil names, DOBs and UPNs. NFER will contact participating control and intervention schools to collect this information in September 2015.

KCL will ask schools for students' Key Stage 2 results. KCL will collect their own data about student and staff attitudes to grouping and to learning through questionnaire surveys and interviews. Attainment data for English and maths will be collected from a sample of year 8 students in summer 2017, using GL Assessment's Progress Test in English and Progress Test in Maths.

How will you handle data?

All those involved in the project treat all personal data in the strictest confidence and no individual school, student or teacher is identified in any report arising from the trial. Full ethical approval for the study has been obtained from the KCL ethics board.

When will the report be available?

The report will be available in early 2018.

Where will the report be published?

The report will be published on the EEF website and disseminated by EEF, KCL and NFER.

Do you need to visit the school?

We will need to visit the school only if your school is selected for student and staff interviews.

Do the research team have current Disclosure and Barring Service checks?

All members of the research team who may visit schools have current DBS checks.

Who needs to give consent for participation in this study?

The Headteacher gives consent to take part in the trial on behalf of the school. In addition, parents/carers and students themselves will be fully informed of the data we will be collecting (e.g. from questionnaires, tests and interviews) and may choose to withdraw their data from the study. They can do this by returning an opt-out consent form which we will send to schools for circulation to parents/carers. Students can withdraw their data from the study at any time up until the end of the intervention. Although we would encourage them to take part, teachers may also withdraw their data from the study and choose not to participate in research activities.

Who has given ethical approval for this study to take place?

The study has ethical approval from KCL.

Research methods

What is a randomised controlled trial (RCT)?

A randomised controlled trial is a type of study in which the people being studied are allocated randomly either to receive an intervention or to be in a control group that does not receive the intervention, so that any effect of the intervention can be assessed by comparing outcomes for the two groups.

Can I choose whether I am in the intervention or control group?

No. In a randomised controlled trial, participants are allocated at random to the intervention or control group.

Incentives

What are the incentives for participation?

Schools in the control group will receive a one-off payment of £1000 at the end of the study; the opportunity to access free GL Assessment maths and English tests, the chance to contribute to the evidence-base for raising attainment for disadvantaged students, and be among the first to receive the report of the study findings in 2018.

The school

Who from the school will need to be involved?

Students in Year 7 (2015-16) and Year 8 (2016-17) will participate in the study, along with their English and maths teachers. Two English teachers and two maths teachers will need to attend professional development events if the school is in the intervention group. The Headteacher must give consent for the school to participate and senior leadership team support will be needed throughout the study.

What expenses will be paid?

KCL will pay standard class travel expenses to and from the professional development events.

How can we get involved in the study?

You can get involved by responding to your school's approach letter which will include a form to complete and return which commits you to the project. Alternatively you can contact Asma Ullah on 01753 637432 or by email (a.ullah@nfer.ac.uk) if you have any queries.

Invitation letter to schools

RPO/EEFK/41636/2

NFER No: _____

Date

Dear Headteacher

'Best Practice in Mixed Attainment' initiative

What are the best ways of grouping students to raise achievement?

Can better approaches to grouping students benefit students from disadvantaged backgrounds?

Take part in research with King's College London and the National Foundation for Educational Research (NFER) to help find out.

We are looking for schools interested in organising their students into mixed attainment teaching groups in Key Stage 3 English and maths. By taking part in this Education Endowment Foundation-funded research project, you will get:

Schools in the intervention group

- **Free, high quality professional development**
- **Access to research on best practice**
- **Practical strategies for raising attainment**
- **Evidence-based guidance on differentiation**
- **Direct engagement in research**
- **Access to free GL Assessment tests to assess your students' progress**

Schools in the control group

- **£1000 at the end of the study**
- **Direct engagement in research**
- **Access to free GL Assessment tests to assess your students' progress**

Interested?

Please continue reading to find out more about what's involved and how to contact us.

What is the study about and who is involved?

The 'Best Practice in Grouping Students' study evaluates the impact on students' educational progress and attainment of different approaches to grouping students. It will explore the use of mixed attainment teaching in secondary schools. The study has a particular focus on low-attaining young people - often from disadvantaged backgrounds - but seeks to promote the progress and raise the attainment of all young people. The trial is funded by the Education Endowment Foundation (EEF). The research team for the trial consists of staff at King's College London (KCL), the University of Nottingham (UoN), and Queen's University Belfast (QUB). The evaluation team for the trial consists of staff at NFER. The trial will measure the performance of pupils in years 7 and 8. The evaluation should equip school staff with the teaching practices to enhance students' learning through mixed attainment arrangements. As part of that process we are asking a number of schools to participate in the trial. Your school is one of a representative sample of schools randomly selected for this purpose and I would like to invite your help with this work. If your school is LA maintained we have notified your LA that we will be contacting you.

What does participation involve?

Schools that agree to take part will be randomly selected to receive an intervention or be part of the control group. Schools in the intervention group will receive the intervention led by the expert team from KCL in **2015/16 and 2016/17**. The intervention will involve implementing the evidence-based best practice for teaching students in mixed attainment groups. Two maths teachers and

two English teachers in the intervention group will need to attend three full-day workshops and three twilight sessions. Schools in the control group will receive £1000 on completion of the trial and the opportunity to access free GL Assessment maths and English tests, and be the first to receive the final research report.

All schools are eligible to participate, provided that they are prepared to implement the evidence-based best practice for teaching students in mixed attainment groups in English and/or maths in 2015/16 and 2016/17.

To evaluate the intervention, NFER will randomly select 30 students to take the Progress Test in English (GL Assessment) and a further 30 students to take the Progress Test in Maths (GL Assessment) from the Year 8 students on the school roll in 2016/17. NFER will administer the tests to the students at the school in summer 2017. The school will need to withdraw the students from two lesson periods and provide an appropriate place for them to take the tests. GL Assessment will mark the tests and report the students' results to the school.

All students in both intervention and control schools will have the chance to complete questionnaires about their experiences and engagement with English and mathematics in 2015/16 and once again in 2016/17. English and maths teachers in intervention schools will also be offered the opportunity to complete questionnaires about practices relating to mixed attainment grouping in 2015/16.

All schools will need to provide data on all pupils in Year 7 in September 2015. We will require basic pupil information such as name, UPN and date of birth. Pupils will be sampled from these lists in the second year of the initiative to sit the GL Assessment progress tests.

I am interested in taking part – what should I do next?

Please complete the enclosed reply form and fax it to 01753 790114 or send it to us in the enclosed pre-paid envelope, or by email (groupingstudents@nfer.ac.uk). We would like to receive your reply **as soon as possible**. To help us to communicate efficiently with your school, and also in view of the confidential nature of this trialling, please provide a named contact from your school staff.

For more information, please see the 'Best Practice in Mixed Attainment' FAQs at: www.nfer.ac.uk/EEFK1. I look forward to hearing from you. Please do not hesitate to contact XXX on XXX or by email XXX, if you have any queries.

Yours sincerely

XXX

Research Manager
Research and Product Operations

Memorandum of Understanding

MEMORANDUM OF UNDERSTANDING

Best Practice in Mixed Attainment trial

This is a memorandum of understanding between King's College London, National Foundation for Educational Research and the school:

School name:

NFER No: _____

The trial

The aim of this trial is to evaluate the impact of the Best Practice in Mixed Attainment intervention on student engagement and attainment in English and mathematics, particularly that of disadvantaged students. A sample of 20 schools in England will participate in the trial, which is funded by the Education Endowment Foundation (EEF). The research team for the trial consists of staff at King's College London (KCL) University of Nottingham (UoN), and Queen's University Belfast (QUB). The evaluation team for the trial consists of staff at the National Foundation for Educational Research (NFER). The results of the evaluation will contribute to understanding which teaching practices are effective in supporting pupil progress and outcomes, and will be widely disseminated to schools in England. Ultimately the evaluation should equip school staff with the teaching practices to enhance students' learning through mixed attainment arrangements and associated practices.

For the purposes of evaluating the Best Practice in Mixed Attainment intervention, NFER will randomly allocate the school to either an intervention group or a control group. Random allocation is the best way of separating out the effect of the intervention from the effect of individual school's qualities. To ensure that this allocation is genuinely random, it is important that the school commits to the trial before it takes places. If the school is allocated to the intervention group, KCL will deliver the intervention to the school in 2015/16 and 2016/17. If the school is allocated to the control group, KCL will send the final research report to the school and make a payment of £1000 to the school on completion of the trial. The intervention itself consists of the operation of specified evidenced-based principles in class organisation, and three full-day workshops and three twilight sessions for two English teachers and two mathematics teachers in the school with information and materials to cascade to colleagues in their departments.

To evaluate the intervention, NFER will randomly select 30 students to take the Progress Test in English (GL Assessment) and a further 30 students to take the Progress Test in Maths (GL Assessment) from the Year 8 students on the school roll in 2016/17. NFER will administer the tests to the students at the school in summer 2017. The school will need to withdraw the students from two lesson periods and provide an appropriate place for them to take the tests. GL Assessment will mark the tests and report the students' results to the school. The school has the option of requesting additional copies of the tests for the remainder of the cohort (although these will not be externally administered or marked). NFER will match the test data with the Department for Education's National Pupil Database and share these data with the research team, EEF, EEF's data processor the Fischer Family Trust (FFT) and, once anonymised, the UK data archive. Prior to the intervention (summer 2015) and during the intervention (summer 2016), NFER may conduct a telephone interview with the Head of English and with the Head of Mathematics about their mixed attainment arrangements and associated practice.

To contribute to the evaluation of the intervention, the research team will administer online questionnaires about their engagement with English and mathematics to the cohort of students in the school soon after they start Year 7 in 2015/16 and again shortly before they finish Year 8 in 2016/17. The research team will conduct an online

questionnaire survey of the school's Year 7 English and mathematics teachers about practices relating to mixed attainment early in the 2015/16 academic year. The research team may also conduct interviews with a small number of students and teachers involved in the intervention (if your school is involved in this element, it will only mean one visit by researchers).

KCL, QUB, UoN, NFER, EEF and FFT treat all personal data in the strictest confidence and no individual school, student or teacher is identified in any report arising from the trial. Full ethical approval for the study has been obtained from the KCL ethics board.

Responsibilities

The research team:

- Acts as the first point of contact with the school for the intervention and the wider trial
- Delivers three full-day workshops and three twilight sessions to two English teachers and two mathematics teachers from the school in a regional location
- Reimburses the teachers/school for standard class travel to and from the meetings
- Provides the four teachers from the school with information and materials for the intervention
- Ensures all staff from the research team visiting the school have received Disclosure and Barring Service clearance
- Conducts the questionnaire surveys of teachers and students
- May conduct interviews with teachers and students involved in the intervention
- Analyses and reports on data from the questionnaire surveys and the interviews
- Analyses and reports on data from the English and mathematics tests
- Administers a £1000 payment to the school, if it is allocated to the control group, once the Year 8 students have completed the tests.

The evaluation team:

- Randomises the allocation of the school to the intervention or control group
- May conduct telephone interviews with some English and mathematics teachers in the school
- Randomly selects 60 students from the Year 8 school roll and administers the English and mathematics tests to them in summer 2017
- If requested, supplies copies of the test to the school for up to/as many as the remainder of the Year 8 cohort at no cost
- Liaises with GL Assessment to ensure the school receives the results of the tests
- Matches the test data with the National Pupil Database for the purposes of the trial, and shares this information with the research team
- Uses data from the tests, questionnaires and interviews to evaluate the intervention
- Disseminates the findings of the evaluation.

The school consents to random allocation to the intervention or control group and commits to participate in the trial whether allocated to the intervention or control group.

The school also agrees to the following instructions:

Organising classes

1. Schools should allocate students to Year 7 classes primarily on the basis of their Key Stage 2 National Curriculum results in English and mathematics available in July, to ensure a broad range of attainment in each class.²⁹
2. Schools should help the research team by:

²⁹ Likewise, students who join Year 7 later than the beginning of the academic year should be allocated to classes according to their National Curriculum Key Stage 2 results in English and mathematics if they are available, or according to their results in other tests, gathered at the earliest opportunity.

- i) Facilitating a questionnaire survey of all Year 7 students in autumn 2015 and all Year 8 students in summer 2017 (questionnaire provided by the research team)
 - ii) If requested, facilitating the researchers to interview a small sample of Year 7 and Year 8 students
 - iii) Facilitate NFER's administration of English and Mathematics tests to Year 8 students in summer 2017.
 - iv) Completing the proforma about existing setting/mixed attainment practices and supplying a list of future Year 7 student names, dates of birth and Unique Pupil Numbers to NFER, in Summer 2015.
3. Two Year 7 English and two Year 7 mathematics teachers should attend regional professional development sessionworkshops³⁰ on best practice in mixed attainment (provided by the research team). These sessions will include elaboration of principles relating to high expectations, differentiation and within-class grouping, which teachers will be expected to apply (and supported to do so).

³⁰ The project will provide three full days of CPD and three twilight sessions.

Reply form



National Foundation for Educational Research

RPO, The Mere, Upton Park

Slough, Berkshire, SL1 2DO

Telephone 01753 637433

Fax: 01753 790111

Email address:
groupingstudents@nfer.ac.uk

RPO/EEFK/41636/2a

NFER No: _____

'Best Practice in Mixed Attainment' initiative in Key Stage 3 English and maths

	Are your details correct?	Please amend
School Name		
Headteacher		
Tel. No:		
Fax No:		
Email:		

Please complete the following:

I have read the attached memorandum of understanding.

My school **can/cannot** take part in this initiative of 'best practice in mixed attainment' grouping in Key Stage 3 maths *(please delete as necessary)*

My school **can/cannot** take part in this initiative of 'best practice in mixed attainment grouping' in Key Stage 3 English *(please delete as necessary)*

Headteacher/SMT signature:

Name of contact in the school:

Mr/Mrs/Miss/Ms/Dr.....

Contact phone number: **Contact job title:**

Contact email address:

Please tick below if your school currently sets and/or streams for years 7 and 8 in English and maths.

	Yr 7 maths	Yr 7 English	Yr 8 maths	Yr 8 English
Do you set				
Do you stream				

Please turn over

**Please return this completed form in the pre-paid envelope provided
or by fax on 01753 790114 as soon as possible.**

If you cannot help us on this occasion, we would be grateful if you could let us know the reasons why your school cannot take part. This will help us understand the factors schools consider when deciding whether or not to participate.

.....
.....
.....

Thank you for your help.



**Evidence for
Excellence in
Education**



INFORMATION SHEET FOR PARENTS/CARERS

Best practice in grouping students

Invitation

We would like your child to participate in this research study. The study is being led by King's College London with the help of Queen's University Belfast and University of Nottingham, independently evaluated by the National Foundation for Educational Research, and funded by the Education Endowment Foundation.

Participation is voluntary. Choosing not to take part will not disadvantage your child in any way. Before you decide whether you want them to take part, it is important for you to understand why the research is being done and what their participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask the research director using the contact details below if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

The purpose of the study is to evaluate the impact of different groupings on students' experiences and learning outcomes. Schools group pupils in different ways; for example in groups based on their prior attainment, or in mixed attainment groups. We want to find out how the grouping arrangements in your school affect pupils' experiences and attainment. The study seeks to draw conclusions about which types of student grouping best improve students' learning outcomes and experiences.

Why is my child taking part?

The study requires a large sample of schools so that the various types of school in England are represented in the research. This will make it possible for us to generalise from the results for the schools in the sample to all of the schools in England. All Year 7 students in your child's school and the other schools participating in the study are being invited to take part in the research.

Does my child have to take part?

Your child does not have to take part. You should read this information sheet and if you have any questions you should ask the director of the study, whose contact details are given below. If you are happy for your child to take part, please keep this information sheet. If you decide that you do not want your child to take part, please complete the enclosed form and ask your child to return it to their teacher. If the teacher does not receive a completed form from you, we will assume that you are happy for your child to take part.

What will happen to my child if they take part?

The following information will be shared between King's College London, Queen's University Belfast, National Foundation for Educational Research, Education Endowment Foundation, their data contractor FFT Education and in an anonymised form, to the UK Data Archive.

- Information provided by your child's school (including your child's name, date of birth and unique pupil number) will be linked with information about your child from the National Pupil Database (held by the Department for Education).
- National Foundation for Educational Research may administer a GL Assessment standardised test in English and/or mathematics to your child towards the end of Year 8. In this case, your child's name, date of birth and gender will be provided to GL Assessment, who will report results confidential to your child's school for educational purposes.
- King's College London will ask your child to complete a short online questionnaire about their experiences of learning English and mathematics, once in Year 7 and once in Year 8. Their teacher will give them an opportunity to complete the questionnaire during school time. It will take your child about 30 minutes to complete the questionnaire. The online questionnaire ensures that their responses are confidential.

Possible benefits

The study is intended to inform education policies and practices in England – and ultimately to benefit students in schools across the country.

What are the possible risks of taking part?

There are no foreseeable risks of taking part in the study.

Will taking part be kept confidential?

Data from the tests, questionnaires and interviews is regarded as strictly confidential and will be held securely until the research is finished, at which point it will be deleted. Your child's participation is entirely voluntary. If you change your mind, you are free to stop their participation. All data for analysis will be anonymised. In reporting on the research findings, we will not reveal the names of any participants or your school. There will be no possibility of any individual being linked with the data.

The UK Data Protection Act 1998 will apply to all data gathered from the tests, questionnaire and interviews. This data will be held securely within the organisations listed above. No data will be accessed by anyone other than the research team or the evaluation team. It will not be possible to link any data back to any individual participating in the research.

You may withdraw your child's data from the project at any time during the project, specifically until 31 August 2017. If you ask us to withdraw your child's data at any time before then, we will remove all traces from the records.

How is the study being funded?

The Education Endowment Foundation is funding this study. For further information about the organisation, visit: <http://educationendowmentfoundation.org.uk>. The study has been approved by the King's College London Research Ethics Committee.

What will happen to the results of the study?

We plan to make our research findings publicly available through events such as seminars, conferences and meetings, and through publications such as reports, articles and books.

Who should I contact for further information?

If you have any questions or require more information about this study, please contact the director of the study, Professor Becky Francis, by email at Becky.Francis@kcl.ac.uk or by telephone on 020 7848 3095.

If this study has harmed you in any way or if you wish to make a complaint about the conduct of the study you can contact King's College London using the details below for further advice and information:

The Chair, Social Science and Public Policy, Humanities and Law, Research Ethics Subcommittee:
rec@kcl.ac.uk

Thank you for reading this information sheet and for considering taking part in this research.

OPT-OUT FORM FOR PARENTS/CARERS



Best practice in grouping students

*Please note that you need only return this form to your child's school if you do **not** want your child to participate in the research project.*

I do **not** want my child to take part in the King's College London research project:

Best practice in grouping students.

Your Name

Name of child

Signed.....(parent/guardian)

Date

Appendix B: Randomisation syntax

title 'EEFK-Best practice in grouping students: Int B(Mixed)'.
 subtitle 'EEFK - First blocked randomisation (15/6/15)'.
 set printback=on.

GET DATA /TYPE=XLSX

 /FILE='K:\EEFK\Research and Product Operations\Randomisation\41636 - recruited schools.xlsx'

 /SHEET=name '41636 - recruited schools'

 /CELLRANGE=full

 /READNAMES=on

 /ASSUMEDSTRWIDTH=32767.

* Check for duplicates.

freq dfeno/format=notable.

sort cases by dfeno.

match files file=*/first=f/last=l/by dfeno.

cross f by l.

temp.

select if any(0, f, l).

list vars=dfeno description.

set rng=mt, mtindex=16.

compute random=rv.uniform(0,1).

exe.

sort cases by random.

freq random.

compute lineno=\$casenum.

exe.

numeric group.

* Uneven randomisation allocate two thirds of schools to the intervention.

if lineno le 10 group=1.

if lineno gt 10 group=2.

add value labels group 1 'Intervention' 2 'Control'.

freq group.

sort cases by contact_id.

SAVE TRANSLATE OUTFILE='K:\EEFK\CfS\randomisation\files for the portal\Randomisation
B - Mixed attainment.xls'

/TYPE=XLSX

/VERSION=8

/MAP

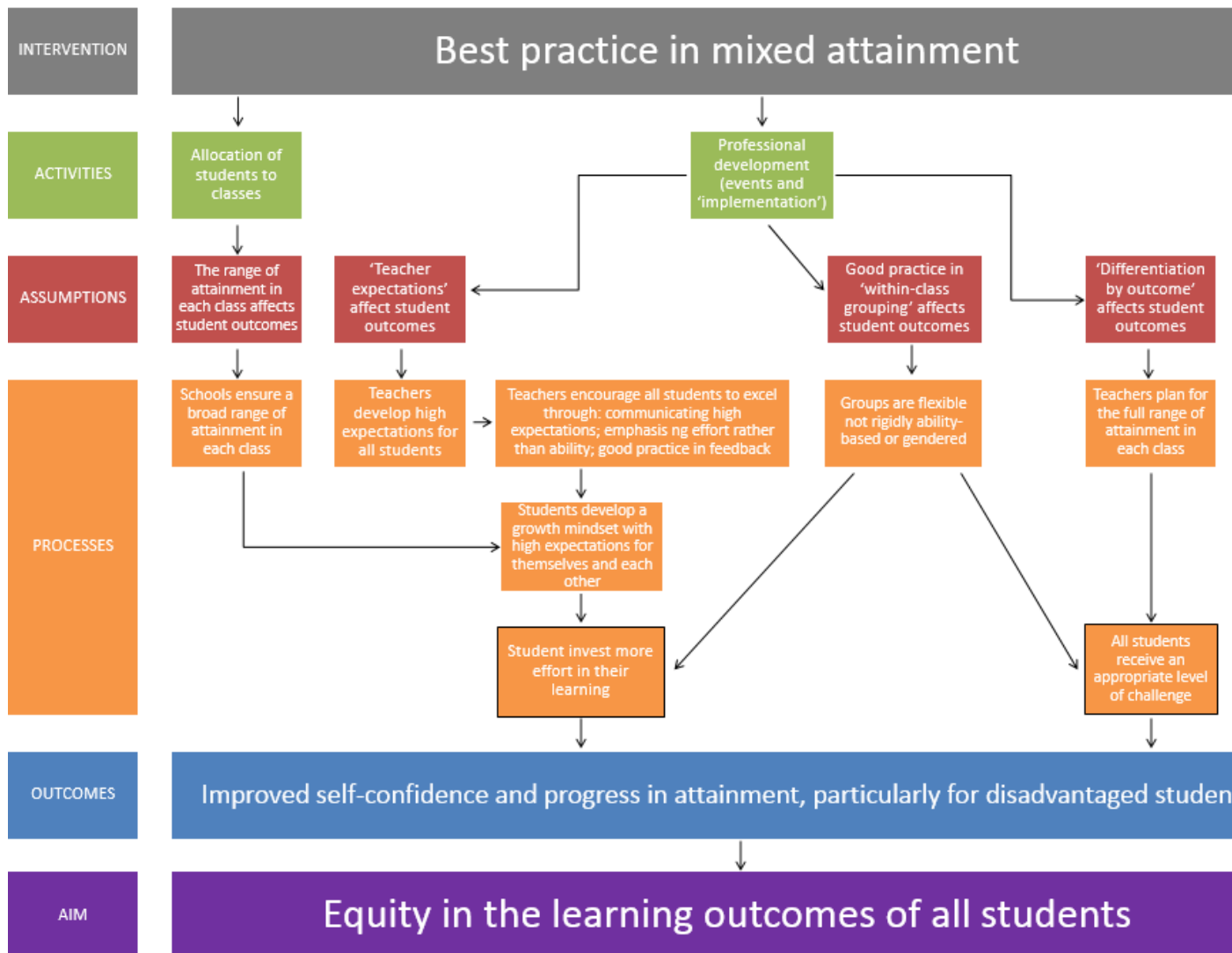
/REPLACE

/FIELDNAMES

/CELLS=LABELS

/DROP=random lineno.

Appendix C: Theory Of Change



You may re-use this document/publication (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0.

OGL This information is licensed under the Open Government Licence v3.0. To view this licence, visit <http://www.nationalarchives.gov.uk/doc/open-government-licence/>

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned. The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

This document is available for download at www.educationendowmentfoundation.org.uk



Education
Endowment
Foundation

The Education Endowment Foundation
9th Floor, Millbank Tower
21-24 Millbank
London
SW1P 4QP
www.educationendowmentfoundation.org.uk