Amendments

Page 6,7- Pupil data collection from schools such as year 7 pupil names, DOBs and UPNs cannot be collected before randomisation as this requires schools to administer parental consent to opt out from the data collection. This will take place after the randomisation.

Page 11 – Table 2 – 4.4.1 - Exploratory scoping interviews. To clarify these are exploratory scoping interviews with the developer (King's College London)

Page 12 – Section 4.4.2 – Attend training/briefing event for schools (autumn term 2015). This should read 'summer term 2015'.

Page 13 – Section 4.4.5 – Telephone interviews with Heads of English and Mathematics (summer terms 2015 and 2016). This should read 'summer terms 2016 and 2017'.

Page 15 – Table 7.1 – Month: May-June 2015 (Telephone interviews with Heads of English and Mathematics in pilot schools). There will be no telephone interviews with Intervention A schools in 2015.

Page 15- Table 7.1 – Month: June 2015 (consent and collection of pupil names, DOBs and UPNs). This will take place in September 2015.

Page 15 – Table 7.1 – Month: September 2015 (Attend launch/training event for intervention schools). This will take place in July 2015.

Page 15- Table 7.1- Months: September 2015 – July 2016 (Obtain key stage 2 results for all randomised pupils) This should read all pupils.

Page 16 – Table 7.2.1 – Evaluation activities for Intervention A (NFER) – Telephone interviews with Heads of English and Mathematics in pilot schools (4.4.5). This activity is not taking place.

Page 16 – Table 7.2.2 - Evaluation activities for Intervention A (NFER) – Administer proformas to schools (4.4.4). This should be omitted.

Page 17 – Table 7.2.3 - Evaluation activities for Intervention A (NFER) – Administer proformas to schools (4.4.4). This should be omitted.

Page 18 – The project will be led on a day-to-day basis by Palak Mehta (not Matt Walker) and she will also oversee the impact evaluation. The process evaluation will be led by Dr. Julie Nelson.

Protocol for the Evaluation of Best Practice in Grouping Students

Intervention A – Best Practice in Setting

Note: This protocol excludes aspects of the evaluation that are the sole responsibility of King's College London and are not requirements of NFER.

1 Introduction

The Education Endowment Foundation (EEF) has commissioned King's College London (KCL) to investigate best practice in grouping students by attainment. The project will consist of two trials. The first trial (Intervention A) will test an intervention which trains schools in a best practice approach to setting. The second trial (Intervention B) is a feasibility study and pilot RCT exploring the use of mixed attainment teaching in secondary schools. This protocol refers to the first trial (Intervention A); the protocol for Intervention B can also be found on the EEF website. The intervention will help schools address poor practices, which include misallocation, low expectations, less demanding curricula and fixed positioning in low groups. The trial will focus on teaching within English and mathematics in years 7 and 8.

The evaluation will run as a randomised controlled trial (RCT), starting in September 2015 and following children through years 7 and 8. The sample for the evaluation will be 120 secondary schools, all of whom currently stream or set for English and mathematics[1], randomised to either receive the intervention or to be part of a control group. The intervention will be piloted with three schools between September 2014 and July 2015.

The National Foundation for Educational Research (NFER) has been commissioned to design and manage the RCT, undertake the process evaluation and assess impact using English and mathematics tests suitable for use at year 8.

[1] Given that many more schools set in mathematics than in English, the trial may be differentially powered for the two outcomes. Ideally all schools selected will already set in English and mathematics, although if there is a struggle to recruit 120 schools, a minimum of 80 that set in both will need to be recruited and then more that set in mathematics only.

2 KCL Project Background

The KCL team explain that there have been various explanations identified for the negative impact of being placed in a low stream or set. Slavin (1990) notes that low achievers "receive a slower pace and lower quality of instruction, have teachers who are less experienced or able and who do not want to teach low-track classes, face low expectations of performance and have few behavioural models" (Slavin, 1990; 473). The KCL team have drawn on the research evidence to identify 7 key factors in poor outcomes for pupils in low sets and streams, of which 5 relate to poor associated practice, as follows:

1) Teacher quality. There is some evidence that teachers perceived as lower quality tend to be placed with lower ability groups (Slavin, 1990; Hallam et al, 2001).
2) Misallocation: There is extensive evidence that certain groups of children (those from lower socio-economic groups, and from particular minority ethnic groups) are over-represented in low sets and streams, and that such allocation does not always match 'ability' as designated by IQ tests (Jackson, 1964) or test scores (Dunne et al, 2007). Teachers and schools tend to underestimate the extent of pupil misallocation (Hallam & Ireson, 2005).
3) Lack of fluidity, or movement between groups: Once placed in an 'ability group' pupils tend to remain there, irrespective of progress/attainment (Flores, 1999; Dunne et al, 2007; Dunne et al, 2010). Teachers and schools tend to overestimate the extent of movement between groups (Hallam and Ireson, 2005).
4) Teacher expectations and related pedagogy. Ireson et al (2005) showed that teachers of high sets convey high expectations through provision of fast-paced and challenging work, whereas pupils in low sets receive slow-paced teaching that covers less of the curriculum. Boaler et al (2000) found teachers adopted much more limited teaching styles when teaching setted groups (see also Hallam & Ireson, 2005) – and Boaler reports these styles were associated with disaffection amongst students.
5) Impoverished curriculum and qualifications: Setting can also produce an 'artificial ceiling', wherein pupils in lower sets are excluded from higher tier study and qualification routes (Dunne et al, 2007; Ireson et al, 2005).
6) Pupil engagement and attitudes: The impact of 'ability' grouping on student identities and expectations is also thought to impact on outcomes (EEF Toolkit, Higgens et al, 2013). Ball (1981) observed that attainment grouping created polarization of students into pro- and anti-school factions: those in low attainment groups became anti-school, with consequent detrimental impact on their achievement and aspirations. Recent research by Hallam & Ireson (2007) hints at continuing dissatisfaction among students in low sets, finding that nearly two-thirds (62%) of young people in bottom sets expressed a wish to change set.
7) Self-Fulfilling Prophecy: The combination of these factors, including teacher quality, misallocation, lack of fluidity, impoverished curriculum, restrict lower set pupils' access to the full curriculum, which in turn creates barriers to moving 'up' sets and to higher tier study. Likewise, pupils' self perceptions with regard to 'ability' impact their aspirations, effort and engagement. Hence these various elements result in self-fulfilling prophecy.

While point 6) (and the potential self-fulfilling prophecy) is hardest to address (the factor most associated with the *act* of grouping, rather than with associated practice), the other factors will all be addressed through an intervention designed to improve

provision for lower sets and ensure pupils are not misallocated and that they can move between groups.

3 KCL Intervention

Several of the identified challenges will be addressed via application of independent, external assignment to sets, and organisational principles. The issues specifically relating to teaching practice (points 4 and 5) will be addressed via the qualitative element of the intervention describing pedagogical and curriculum principles. Hence the detrimental factors identified above will be addressed as follows:

1) Teacher quality. Teachers will be randomly assigned to sets by the research team, to ensure that teacher quality does not relate to set level.
2) Allocation: Pupils will be assigned to sets via independent measures, thus ensuring that allocation is not influenced by practitioner perceptions of pupil background. Sets will also be required to be relatively broad (maximum four sets, rather than the 5-8 often applied in schools) in order to dilute hierarchies and to ameliorate the effects of misallocation. Narrow setting restricts the opportunities for more flexible and fluid within-class groupings (as recommended in the EEF Toolkit, Higgins et al, 2013). Broader sets increase the opportunities for more fluidity through strategies such as 'near-ability' grouping (Askew & Wiliam, 1996).
3) Fluidity: Re-assignment of pupils to different sets, strictly on the basis of attainment results/independent measures, will be applied three times across the two-year period, and this practice communicated clearly to pupils.
4) (4&5) Pedagogy, teacher expectations and curriculum access: The developers will provide workshops[2], and high quality, targeted materials and information to treatment schools, with the intention of: challenging teachers' 'fixed ability' practices and encouraging high expectations[3]; and facilitating schools to implement good practice in setting. Workshops will involve inputs, discussion, group work, and reporting on peer-to-peer work in schools. The first workshop will include provision of dedicated time and support for practitioners to develop a strategy for their school to encourage i) high expectations, ii) quality teaching and iii) a rich and aspirational curriculum, for all young people. Heads of (Maths and English) Departments will be responsible to implementing these strategies in their schools.
6) Pupil engagement and attitudes: Pupils' engagement, self-perception and attitude will be assessed at the beginning of the intervention and at the end of the first year and second year; and compared to those of pupils from the control schools.

[2] The 60 treatment schools will be divided into 6 geographical areas, with workshops provided for each area (ideally these will be hosted in one of the area schools, on a rotating basis). Heads of Maths and English departments, and those teachers delivering Maths and English to the Year 7 cohort in treatment schools, will be expected to attend (ie approximately 6-10 staff from each school). Initial workshops will be focused on the KS3 leadership developing a department-wide approach to addressing the factors identified: Increasing fluidity, raising teacher expectations and pedagogy, improving access to the whole curriculum and higher status qualifications, improving pupil engagement and attitudes and tackling the self-fulfilling prophecy. The developers shall present practitioners with the research evidence together with pre-intervention data on set allocation, pupil attitudes and teacher attitudes / practices together with examples of how other schools (particularly the pilots) have addressed this. Practitioners will be asked to develop a department-wide approach to this, and Heads of Department become agents responsible for implementation, monitoring and working with teachers in their schools. Later workshops will provide refreshment on the evidence and import of the intervention, discussion of progress and issues arising, feedback on peer-to-peer work in schools, an arena for monitoring progress, and sharing of best practice. There will be four workshops provided across the two-year intervention. Additionally, 'bespoke' support and advice will be provided to schools by the project team as necessary.

[3] Information and materials here will draw on Dweck's work on 'mindset' and Ames's work on task-orientated rather than ego-oriented goals, along with other research showing how 'ability' is not pre-determined or fixed, and the impact and import of teacher expectations on pupil outcomes.

Hence the criteria to be applied in the intervention will include:

- Where possible, random distribution of teachers across sets (or where not possible, distribution according to principles provided as part of the intervention)
- Pupils are subject to external measures of attainment, and set at the beginning of year 7 accordingly
- A narrower range of sets (maximum four)
- Pupil progress and attainment is assessed twice yearly, and pupils are moved accordingly
- Schools are expected to support staff to attend workshops; and to develop, implement and monitor a school-wide strategy aimed at providing high expectations and a stimulating, challenging curriculum to all pupils
- Headteacher support for all the above.

4 NFER Evaluation plan

4.1 Research questions

The primary research questions are:

1. What is the impact of best practice in setting on pupils' attainment in mathematics?
2. What is the impact of best practice in setting on pupils' attainment in English?

The secondary research questions are:

3. What is the impact of best practice in setting on pupils' self-confidence in mathematics?
4. What is the impact of best practice in setting on pupils' self-confidence in English?

NFER will carry out a light-touch process evaluation that focuses on the training given to schools, the ease of and barriers to implementation and how scalable the intervention is. The process evaluation will ascertain what aspects of practice are different 'on-the-ground' between intervention and control schools.

4.2 Overall design

4.2.1 Introduction

This best practice in setting main trial (Intervention A) will start in September 2015 and will run until November 2017. The trial will be designed, conducted and reported to CONSORT standards (http://www.consort-statement.org/) and registered on http://www.controlled-trials.com/.

4.2.2 Overview of design

The evaluation will start with a pilot in academic year 2014/15. The main intervention and trial will span the following two academic years. We expect a lag between both understanding of best practice and its implementation and then impact on pupils, hence testing is only being proposed in the second year of the main trial (see Table 1 below).

Table 1 Overview

Year 1	Year 2	Year 3
2014/15	2015/16	2016/17
Application of pilot interventions to three schools	Intervention introduced to year 7 pupils in 60 schools	Intervention continues. Year 8 pupils tested (summer 2017) to capture impact of two years' best practice (during years 7 and 8)

4.3 Impact evaluation

4.3.1 Eligibility and recruitment

The population for this trial will be all state-funded English secondary schools. It is envisaged that most will presently employ setting for both key stage 3 English and mathematics. However, to support recruitment any school using ability grouping in year 7 will be eligible to take part regardless of their prior grouping arrangements, including schools that are undertaking streaming.

Since information on schools' grouping arrangements is not routinely available on school-level datasets, schools will need to complete a proforma at recruitment stage indicating the details of their grouping policy. This proforma will be administered by KCL and will inform the randomisation. It may be that not enough schools are willing or able to participate in both English and mathematics. In this case it could be necessary to embark on a trial that is differentially powered for the two primary outcomes.

The three schools selected for the pilot will all employ setting in both mathematics and English. For the main trial, the following eligibility rules in Table 2 will be used to recruit 120 schools. Note that Table 2 assumes schools always prefer setting in mathematics to setting in English. Schools that stream (i.e. allocate their pupils to fixed ability groups across subjects) will be eligible if they are prepared to amend their grouping arrangements to setting.

Table 2 Eligibility criteria

Year 7		Year 8		Eligible?
English	**Mathematics**	**English**	**Mathematics**	
Setting	Setting	Setting	Setting	Yes
Mixed	Mixed	Setting	Setting	No
Mixed	Setting	Mixed	Setting	Yes (for mathematics outcome only)
Mixed	Mixed	Mixed	Setting	No
Mixed	Mixed	Mixed	Mixed	No
Streaming	Streaming	Streaming	Streaming	Yes, if prepared to set

School recruitment will be carried out jointly by NFER and KCL. Schools will need to sign an MoU and complete a proforma about existing setting practices. Schools will also need to supply a list of future year 7 pupil names, DOBs and UPNs to NFER via secure portal after randomisation. The DfE co-signed letter '*Provision of UPNs for research funded by the Education Endowment Foundation'* will be used for this purpose.

4.3.2 Randomisation

As best practice encapsulates a whole-school approach, for example allocation of teachers across sets, school-level randomisation will be necessary. 120 schools is a large enough sample to avoid the need for stratification or minimisation so simple randomisation will be used. Should it not be possible to recruit 120 schools then the minimum threshold would be 80 schools (40 vs. 40). As indicated above, it is likely to be easier to recruit schools that set for maths. It may therefore be necessary to stratify the randomisation by setting practice (English/maths/both) to allow a lower powered analysis to be run on English outcomes. Randomisation will be carried out by a statistician at NFER using a full syntax audit trail.

The timing of random allocation of schools will be critical. Randomise too early and this does not give enough time to recruit schools. Randomise too late and this does not give schools enough time to implement the logistics of best practice; which involve pupil and teacher allocation to sets and rewriting schemes of work/curricula. To complicate things further, some schools ability group on the basis of key stage 2 results at the start of year 7; others implement mixed ability groups initially and then set in year 8. Changing the way pupils are grouped or teachers are allocated during the school year would be very disruptive so it will be important to have changes to

these strategies in place before the school year starts. Randomisation will be in June 2015 to allow time for KCL to allocate pupils to sets using independent measures by the end of term. UPNs, names and dates of birth for pupils in years 7 will be collected from schools via the NFER schools' portal after randomisation.

In order to keep the control schools engaged with the evaluation they will receive £1000 at the end of the trial, once they have completed the year 8 tests. The payment will be administered by KCL.

4.3.3 Outcomes

As the best practice intervention is aimed at pupils in years 7 and 8, testing will be necessary as there is no statutory assessment in these years. Testing will take place in year 8 after two years of the intervention. For a high-level structural intervention such as this, it is important that the chosen tests are as broad as possible and cover the English and mathematics curricula in use at the time. GL Assessment is presently funding the development of their New Progress in English (NPiE) and New Progress in Mathematics (NPiM) tests that will be available for use by summer 2017. Given the curriculum changes being implemented from September 2014, these tests will be used in preference to existing versions. NPiE for year 8 will consist of two components: spelling, punctuation and grammar (20-25 minutes) and reading comprehension (40-50 minutes). NPiM for year 8 will also consist of two components: mathematical skills and concepts (60 minutes) and mental maths (15 minutes). NPiE and NPiM can each be sat in a single session; they do not have to be split.

NFER will take responsibility for collecting and delivering NPiE and NPiM in paper form.

Secondary outcomes including self confidence in English and maths will be measured by KCL using a pupil survey at the start of year 7 in September 2015 and at the end of year 8 in summer 2017.

4.3.4 Sample size

NFER will randomly select 60 pupils from the year 8 school roll from each of the 120 schools. Half of the pupils will sit the mathematics test and half will sit the English test. Sampled pupils will be required to be out of their normal lessons for more than one period but less than two. Test administrators will be used to ensure that all tests are administered the same way in all schools; they will also help with response rates and with minimising the burden placed on schools.

In addition, pilot schools have indicated that receiving mathematics and/or English tests for pupils not participating in the trial would be helpful. Thus, to support the recruitment of schools to the main trial, all participating schools will be given the option of receiving additional mathematics and/or English test papers to be marked by teachers for as many as all pupils in Year 8.

4.3.5 Power calculations

Randomisation will be conducted at the school level. The power curves in Figures 1 and 2 are for three testing options, one of which is the option that will be used for this trial and is described above. They all use the following assumptions: intra-cluster correlation of 0.15 (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 Power curves for 60 versus 60 schools (recruitment target)

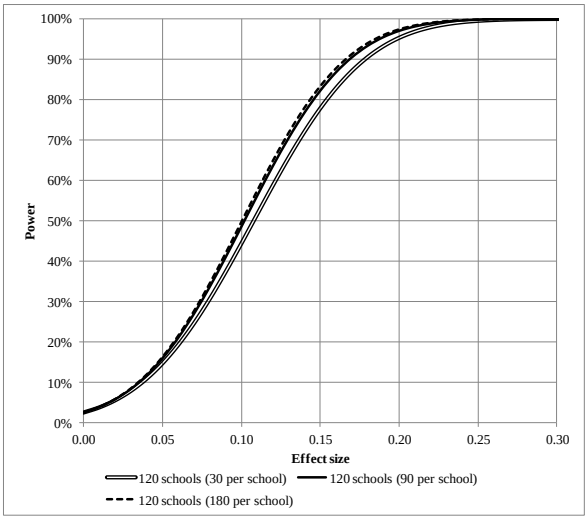
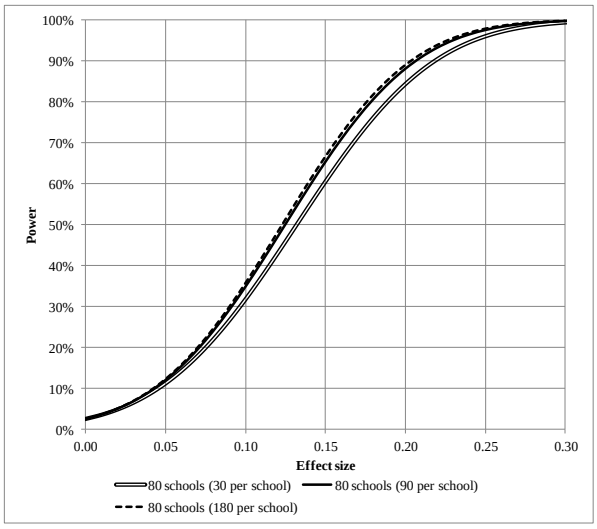
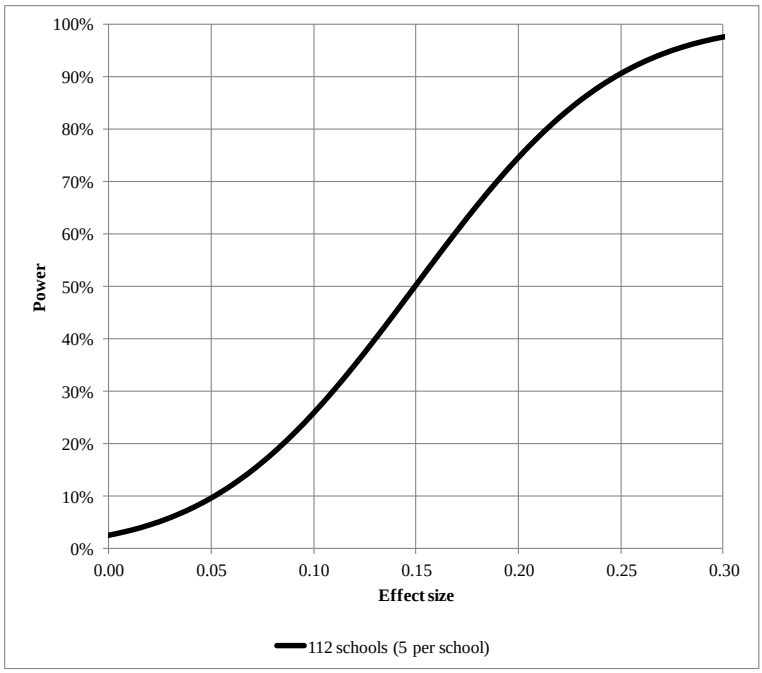


Figure 2 Power curves for 40 versus 40 schools (recruitment minimum)



These power curves clearly illustrate how testing burden per school can be greatly reduced through within-school sampling with minimal impact on power. For such strategies to work, the within school sampling has to be random to ensure unbiased cluster mean estimates. Within-school sampling has an impact on the power of sub-group analysis. As FSM-eligible pupils represent a particularly important subgroup the power of a separate FSM analysis is considered in Figure 3. The average proportion of FSM pupils in years 7 and 8 in eligible schools is 16.4% so we can expect an average of 5 FSM pupils to be sampled respectively in each school cohort. With the chosen testing design, it is critical that within-school sampling is stratified by FSM. With stratification, we can reasonably expect at least one FSM pupil to be sampled in each school in 93% of schools (all schools if we direct the sample to schools with high FSM). As we are just estimating regression coefficients some small cluster sizes will not compromise the multi-level models (Snijders et al, 2005). The minimum detectable effect size with 80% power remains at 0.22 for the chosen design when analysing FSM pupils, if the recruitment target is reached.

Figure 3 Power curves for FSM analysis



4.3.6 Analysis

Primary outcome intention-to-treat analysis of year 8 tests will use multi-level models containing two levels (pupil and school) to account for the cluster randomisation. It will use key stage 2 baseline data as a covariate in the models. Given the evidence from the literature that ability grouping is possibly advantageous for the more able while damaging for the less able, it will be important to establish whether this pattern is attenuated by best practice. As well as testing for an overall effect of best practice, we will therefore be particularly interested in establishing whether best practice is differentially effective for different ability levels. This will be done by testing for the interaction between best practice and ability. A significant negative interaction would imply that best practice is 'closing the gap' between the less and more able. A separate analysis of FSM pupils will also be carried out as per standard EEF practice.

Methods for missing data such as multiple imputation and sensitivity analysis will be carried out if attrition levels dictate.

Secondary outcome analysis will also use multi-level models containing two levels (pupil and school). It will use baseline self-efficacy measures as a covariate in the models.

On-treatment analysis will be based on the fidelity indicators developed by KCL. These will yield a quantitative measure of fidelity that can be used in multi-level models of attainment and self-confidence.

4.4 Process evaluation

The purpose of the process evaluation is to collect information on how far the best practice approach is adopted and implemented by trained schools, and any organisational and attitudinal changes occurring for schools and teachers as a result of the training and best practice approach. The information collected will be important to support consistency and fidelity checks and to determine whether the programme is scalable.

The process evaluation will collect information on:

- the resources involved in the intervention, including the training sessions, teacher time and staffing
- the starting points of schools, in terms of setting in English and mathematics at key stage 3 (including clarification of setting criteria, streaming, etc)
- changes made to school organisation (e.g. staffing allocations, material resources), curriculum and pedagogy (e.g. level of demand/challenge), and grouping criteria (e.g. how schools address misallocation, fixed positioning, movement of pupils between sets)
- brief perceptual feedback on changes in school ethos, and teachers' attitudes, expectations and assumptions around ability grouping

The process evaluation will involve five main strands of activity spread across three academic years, as detailed below.

Table 2: Overview of research strands associated with the process evaluation

	Year 1 (2014/15) (Pilot year)	Year 2 (2015/16)	Year 3 (2016/17)
4.4.1 Exploratory scoping interviews with the developer (King's College London)	✓	✓	✓
4.4.2 Attend training/briefing event for schools	✓	✓	
4.4.3 Collation and review of project documentation	✓	✓	✓
4.4.4 School proformas		✓	✓
4.4.5 Telephone interviews with Heads of English and Mathematics		✓	✓

This combination of methods will give us a good understanding of how and why the intervention has/has not worked including implementation challenges and adaptations, any unexpected outcomes and perceived impacts, views on sustainability and scalability. However, as NFER's process evaluation will be light-touch, KCL will collect much of the information suggested above (e.g. numbers of teachers attending training, schools' current practice in ability grouping, what 'business as usual' looks like).

Further detail on each strand is provided below.

4.4.1 Exploratory scoping interviews (spring term 2015)

Given the complexity and dynamic nature of the intervention it will be important to first establish its scope and the range and availability of administrative data which could support the evaluation. We therefore plan to undertake a number of face-to-face and/or telephone interviews with key stakeholders and identify relevant management and monitoring information. We suggest this could include interviews with up to two key staff at KCL. As part of the interviews we will develop an overall Theory of Change (ToC) or Logic Model setting out foci and aims, inputs and resources, outputs in terms of activities and participation, and desired outcomes (short, medium and longer-term). We will agree this overall ToC/Logic Model with EEF and KCL. This will ensure a common framework for the evaluation and identify common desired outcomes. The ToC/Logic Model will be important for understanding the mechanisms leading to change. It will be reviewed in years 2 and 3 and updated as necessary to reflect any modifications made to the intervention following the pilot.The Models will also support the costs analyses.

4.4.2 Attend training/briefing event for schools (summer term 2015)

We will attend a training/briefing day for schools during the start of the full trial in year 2 (2015/16). Attendance at these sessions will allow us to understand the nature of the best practice approach and how assumptions are challenged.

4.4.3 Collation and review of project documentation (autumn term 2015)

We will collect project documentation from KCL, such as information about the initiative and best practice approach, duration and content of the training sessions, number of teachers attending training sessions from each school, nature of resources and support provided to schools by KCL. We will also need to collect information from KCL about the existing nature of setting in English and mathematics in the 60 treatment schools (e.g. setting, streaming, criteria, number of sets, etc); and in the 60 control schools at recruitment stage.

The data will need to be collected for the schools at the beginning of their respective programmes (e.g. the autumn term2015). We would ask that KCL keep NFER informed of any changes to the intervention or support provided to participating schools.

As part of the process evaluation, we will also review the costs of delivery of the intervention. We will develop a cost information specification. KCL will be required to supply NFER with information on the costs of delivery in summer 2015, and updated in summer 2017. Schools will also provide costs information (e.g. staff time, supply cover costs) through the school proforma (see section 4.4.4).

4.4.4 School proformas (summer 2015 and spring terms 2016 and 2017)

Brief proformas will be administered to the Head of Department/key contact in all 120 schools prior to randomisation. The proforma will collect information on 'business as usual' so that we understand schools' ability grouping practices. Following randomisation a second proforma will be distributed to the 60 intervention schools to capture what the implementation of best practice looks like and to ascertain what organisational, curricular and grouping changes they make. This second proforma will need to be administered in the spring terms of year 2 (2015/16) and year 3 (2016/17).

4.4.5 Telephone interviews with Heads of English and Mathematics (summer terms 2016 and 2017)

We will conduct telephone interviews with the Head of English and Head of Mathematics in year 2 (2015/16). We will interview staff in both roles from a sample of ten randomly selected intervention schools. We will then follow up with the same ten schools a year later to gain a longitudinal perspective of the implementation of the intervention. Should any of the schools need to drop out of the follow-up interviews (e.g. due to staff illness or an Ofsted inspection), replacement schools will be randomly selected from the intervention arm. The telephone interviews will last for about 30 minutes and will explore:

- If and how teachers have responded to the intervention, including evidence of changes in behaviour/activities
- The extent to which participating teachers are following the prescribed guidance
- Respondents' views on the conditions required for success and the extent to which the intervention is scalable
- Any barriers to effectively grouping pupils by ability, and how, if at all, these have been overcome
- Perceived outcomes for learners, and how, if at all, these differ for learners in different ability groupings
- Unintended consequences, positive or negative.

5 Reference

Snijders, Tom A.B. 'Power and Sample Size in Multilevel Linear Models'. In: B.S. Everitt and D.C. Howell (eds.), Encyclopedia of Statistics in Behavioral Science. Volume 3, 1570–1573. Chicester (etc.): Wiley, 2005.

6 Reporting

NFER will provide termly progress reports to EEF. We will prepare a report of the overall evaluation findings to CONSORT standards by November 2017. This will

include findings from the impact evaluation (primary and secondary pupil outcome measures) together with the findings from the process evaluation.

7 Timeline

The pilot intervention will commence in September 2014 and the RCT will commence in September 2015. The primary attainment outcomes will be captured through testing in summer 2017.

7.1 Overall timeline

Month	Activity
May 2014	Inception meeting between EEF, KCL and NFER
May-July 2014	KCL recruit pilot schools; consent
Sept 2014 – July 2015	Pilot schools implement best practice approach, teachers' assumptions/practices change. Target pupils are in year 7
June 2015	NFER and KCL complete recruitment of trial schools; 'business as usual' information collected via proformas; NFER randomise schools
July 2015	Attend launch training/briefing event for intervention schools
September 2015	Consent and collection of pupil names, DOBs and UPNs
September 2015 – July 2016	Trial schools implement best practice approach Target pupils are in year 7 Obtain key stage 2 results for all pupils
September-October 2015	Collation and review of project documentation
Jan–Mar 2016	Administer proformas to 60 trial schools
May-June 2016	Telephone interviews with Heads of English and Mathematics
September – Dec 2016	Continue to implement best practice approach Target pupils are in now year 8
Jan – March 2017	Administer proformas to 60 trial schools
May-June 2017	Telephone interviews with Heads of English and Mathematics
June/July 2017	Administer tests to schools
November 2017	Submit draft report

7.2 Timeline by year overlaid with intervention activities

7.2.1 Year 1 (2014/15)

	2014							2015							
	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A
Intervention A activities – Best Practice in Setting (KCL)															
• Refining pilot design															
• Application of pilot interventions; monitoring and evaluation															
• Design of survey and interview instruments, Piloting and refinement of research instruments. Monitoring and reflection on progress of pilot interventions.															
• Provision of brief report on pilot															

7.2.2 Year 2 (2015/16)

	2015				2016							
	S	O	N	D	J	F	M	A	M	J	J	A
Intervention A activities – Best Practice in Setting (KCL)												
• Implementation of intervention												
• Conduct first survey sweep (pupil attitudes survey)												
• Survey data capture and cleaning												
NFER evaluation activities for Intervention A												
• Collation and review of project documentation (4.4.3)												
• Telephone interviews with Heads of English and Mathematics (4.4.5)												

7.2.3 Year 3 (2016/17)

	2016				2017							
	S	O	N	D	J	F	M	A	M	J	J	A
Intervention A activities – Best Practice in Setting (KCL)												
• Implementation of interventions												
• Qualitative interviews conducted												
• Second survey sweep (pupil attitudes survey)												
• Survey data capture and cleaning												
Evaluation activities for Intervention A (NFER)												
• Telephone interviews with Heads of English and Mathematics (4.4.5)												
• Administer tests to schools												
• Marking and data capture (GL Assessment)												

7.2.4 Year 4 (2017/18)

	2017				2018							
	S	O	N	D	J	F	M	A	M	J	J	A
Evaluation activities for Intervention A (NFER)												
• Data analysis and reporting												

8 Personnel, roles and responsibilities

The project will be directed by Dr. Ben Styles at NFER, and led and managed on a day-to-day basis by Palak Mehta at NFER. She will also oversee the impact evaluation. Michael Neaves in NFER's Research and Product Operations Department will coordinate NFER test administrators and oversee the dispatch and collection of the tests to schools. Dr. Julie Nelson will oversee the process evaluation, supported by an experienced researcher.

NFER's data protection policy is available at: http://www.nfer.ac.uk/nfer/about-nfer/code-of-practice/nfercop.pdf.

In setting out the roles and responsibilities for this trial, the NFER will draw up a Data Sharing Agreement with KCL. This will include a description of the nature of the data being collected by KCL and NFER and how it will be passed between them. In addition, the NFER, EEF and KCL will need an MoU with schools, explaining the nature of the data being requested of schools, how it will be collected, and how it will be passed to and shared with all the organisations involved

9 Risks

Risk	Assessment	Countermeasures and contingencies
Insufficient schools recruited to the study	**Likelihood:** high **Impact:** high	Timescale could be revised.
School attrition	**Likelihood:** moderate **Impact:** moderate	Clear information to schools, explaining the principles of the trial and expectations. Attrition will be monitored and reported according to CONSORT guidelines.
Intervention is not implemented well	**Likelihood:** low **Impact:** low	Clear information to schools, and initial meeting between schools and KCL, explaining the principles of the trial and expectations.
Delays in training sessions and commencing any organisational changes in schools	**Likelihood:** moderate **Impact:** low	Agree a clear timetable, with parameter windows, with the project team up front. Plan evaluation and testing timetable to allow enough time for all schools in the sample to have received training and implemented aspects of the best practice approach.
Control schools adopt similar treatments	**Likelihood:** moderate **Impact:** high	Control schools hear of best practice approaches and adopt these treatments. 'Business as usual' approaches may, or may not, include aspects of best practice approaches. Process evaluation will ascertain ability grouping approaches in the control schools.
Admin data required not available or supplied in incorrect format	**Likelihood:** low **Impact:** low	Data sharing procedures will be agreed in advance with KCL.
Day-to-day trial management required by NFER	**Likelihood:** low **Impact:** low	Experienced KCL team unlikely to require support. Request from EEF that the RCT evaluator role is limited to checking the design, randomising and conducting an independent analysis of the results.
Researchers lost to project due to sickness, absence or staff turnover	**Likelihood:** low/moderate **Impact:** moderate	NFER has a large research department with numerous researchers experienced in evaluation who could be redeployed. Senior staff can stand in if necessary.
Project does not follow correct trial protocols	**Likelihood:** low **Impact:** high	Shared and agreed protocol agreement with KCL and EEF. Provision of clear guidance and protocols for distribution to all schools.