Trial Evaluation Protocol
English Mastery

Evaluator (institution): NatCen
Principal investigator(s): Valdeep Gill, Robert Wishart

Education Endowment Foundation

Template last updated: March 2018

PROJECT TITLE	English Mastery
DEVELOPER (INSTITUTION)	Ark
EVALUATOR (INSTITUTION)	NatCen Social Research
PRINCIPAL INVESTIGATOR(S)	Valdeep Gill, Robert Wishart (formerly Priya Khambhaita, Martina Vojtkova)
PROTOCOL AUTHOR(S)	Rodrigo Torres, Priya Khambhaita, Josep Espasa, Anna Marcinkiewicz, Enes Duysak, Ben Stocker, Valdeep Gill, Robert Wishart
TRIAL DESIGN	Two-arm cluster randomised controlled trial with random allocation at school level
PUPIL AGE RANGE AND KEY STAGE	Years 7-8, Key Stage 3
NUMBER OF SCHOOLS	100
NUMBER OF PUPILS	6000
PRIMARY OUTCOME	GL Pupil achievement in English assessment (Progress Test in English)
SECONDARY OUTCOME	Teacher workload, SPAG and reading comprehension.

Protocol version history

VERSION	DATE	REASON FOR REVISION
1.0	20.12.2019	Original protocol
2.0	06.11.2020	Changes to the intervention and evaluation due to the Covid-19 pandemic,which include: • Changes to the delivery of English Mastery e.g. the movement of training and school visits to online. • The possibility of using a two-level multilevel model for the primary analysis. This will be carried out if the decision rule to implement the preferred three-level multilevel model is not met. • A new Covid-19 optimal compliance measure.

Please note that the contents reflect the study protocol at time of writing (November 2020). Research activities may be subject to change as the Covid-19 pandemic evolves.

Table of contents

The protocol structure

This protocol sets out the plan for the impact and implementation and process evaluation of the English Mastery (EM) programme. Due to the outbreak of the Covid-19 pandemic in March 2020, the intervention and the evaluation have been subject to some changes. The protocol describes the delivery of the intervention/evaluation as intended prior to the pandemic and highlights changes to the intervention/evaluation design and delivery.

Intervention

The aim of the EM programme is to implement a coherent and cumulative approach to curriculum design including bespoke assessments. The aim of the intervention is to improve achievement in the subject of English. This is done by providing teachers with subject-specific training, curriculum materials and ongoing in-school support and coaching. Changes to intervention design and delivery necessitated by the Covid-19 pandemic are described in the table below.

Why: theory/rationale

An Ofsted 2015 report highlighted that pupil progress, in Key Stage 3 English, was too slow and not satisfactory. Weak foundations in English can impede academic performance in not only English but other subjects more broadly[1] [2]. There has consequently been increased focus on changing secondary school curricula, making them broader and deeper in order to equip pupils with the knowledge and skills needed for GCSE and further education. The EM programme is aligned with this approach.

EM is a comprehensive programme with two pathways: the traditional curriculum for children reaching age-related expectations, and the foundation curriculum for those working below expectations. All pupils receive the same dosage and study the same topics, but the foundation curriculum is adapted to be accessible to lower attaining pupils (for example, using abridged texts). If pupils meet age-related expectations in two successive assessments then it is recommended that they graduate from foundation to traditional.

The curriculum revolves around four pedagogical pillars:

1) emphasising accumulation of knowledge,
2) discrete grammar teaching,
3) systematic instruction of Tier 2 vocabulary, and
4) use of standardised, norm-referenced pupil work.

Overarching themes include an emphasis on explicit instruction, a purposeful and cumulative knowledge-rich curriculum, and developing pupils' vocabulary. The project is also heavily influenced by findings from cognitive science like cognitive load and interleaving.

The expected outcomes from this intervention at the teacher level are:

- An increase on teachers' effectiveness to teach canonical texts, grammatical concepts and vocabulary;
- Teachers' better understanding of how to use assessment to drive planning and delivery and give feedback to pupils;
- Reduced teacher workload and improved confidence.

These outcomes at teacher level may result in improved pupil outcomes, primarily these are:

[1] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/459830/Key_Stage_3_the_wasted_years.pdf
[2] https://www.gov.uk/government/news/more-pupils-in-england-reach-expected-standard-at-key-stage-2

- Improved pupils' reading and writing skills;
- Improved pupils' Tier 2 vocabulary and knowledge.

This programme is expected to help pupils possess the key competences required to succeed at GCSE, A' level, University and beyond. The pilot study evaluation of the EM intervention found significant positive effect on pupils' English scores, regardless of pupils' prior attainment and demographic characteristics. The effect size of the intervention from the pilot study of the evaluation showed the equivalent to four months' additional progress (d'=0.28). It was particularly beneficial for pupils with lower prior attainment, and marginally more effective for non-Free School Meals (FSM) pupils[3]. NatCen are building on the work conducted in the pilot study to evaluate the EM intervention on a larger scale, and with a greater scope. The logic model underpinning the programme can be found in Appendix 1.

What: Physical or informational materials used in the intervention

Teachers are provided with lesson plans and all pupil-facing materials. The material is divided into three strands: Literary Heritage, Mastery Writing and Reading for Pleasure. The Literary Heritage strand includes lesson slides, handouts, re-teach tasks and suggested homework tasks. The Mastery Writing strand also includes lesson slides, pupil handouts and a teacher handbook. The reading for pleasure strand includes suggested activities to develop pupil vocabulary.

Teachers have access to all the units for the year within a given strand. Teachers receive a fully resourced lesson by lesson Key Stage 3 curriculum that helps pupils of all abilities work through classic stories like Jane Eyre. Weekly departmental co-planning materials are also provided that help teachers adapt the English Mastery lessons to meet the specific needs of the pupils in their classrooms[4]. Teachers are also offered training and support through induction training, termly school coaching visits, termly Assessing for Mastery days and optional webinar sessions.

What: Procedures, activities and/or processes used in the intervention

There are 8 days of training provided by the English Mastery team, plus three in-school visits. The various training days are aimed at different members of staff depending on their role in delivering the intervention (e.g SLT Champions are only required to attend one induction training day). The training model is designed to build subject-specific pedagogy, curriculum content knowledge and effective delivery methods. The training is sequential and there is coherence between the different elements.

Schools are required to nominate a member of their English department tobecome the English Mastery Ambassador to lead the programme in their school. The induction training for all teaching staff involved in EM delivery at their schools[5] builds subject pedagogy and provides opportunities for modelling and deliberate practice. The assessment days allow teachers to review impact and refine planning and delivery accordingly. This sequence is repeated termly. The full training package covers year 1 and year 2 of the trial.

In addition, there are webinar twilights that EM Ambassadors and teachers delivering EM can join. These are optional and they are designed to build curriculum content knowledge and focus on effective delivery methods.

EM will develop series of videos aiming to help EM Ambassadors implement the programme successfully in schools. The videos will focus on demonstrations of best practice in the classroom and interviews with EM Ambassadors. The videos will be a mixture of stand alone

[3] https://www.englishmastery.org/wp-content/uploads/2018/05/English-Mastery-Pilot-Impact-Evaluation-.pdf
[4] https://www.englishmastery.org/programme/
[5] Senior Leadership Team representative, English Mastery Ambassador, and English Teachers.

resources and footage that can be used to complement training in webinars and on training days.

In-school, the weekly co-planning sessions, led by the EM Ambassador, support teachers to plan together. Schools are provided with co-planning materials, scripted examples and video examples to support this process.

Schools also receive specialist support via three in-school visits. The first part of the visit is dedicated to instructional coaching via lesson observations and the second half is given over to bespoke Continual Professional Development (CPD) for the English department. Beyond the default model, extra visits are available for schools that need more support.

Who: Intervention providers/implementers

The intervention provider is English Mastery which is an Ark venture that is part of the education charity, Ark. English Mastery is designed to be delivered by secondary English teachers. Every school participating in the English Mastery programme is assigned a link person from English Mastery who visits the school at the start of the intervention to provide bespoke training and supports the English Mastery Ambassador with programme implementation. The English Mastery link visits the school two more times across the year to build on the work from the first visit and deliver bespoke support and training if requested by the school.

How: Mode of delivery

All English teachers delivering the programme receive an induction training session before the start of the school year. In addition, a member of their senior leadership team and the English Mastery Ambassador receives an additional day of training that focuses on the leadership and implementation of the programme. This training is followed by termly subject mastery webinar sessions, termly in-school visits (which include coaching and bespoke CPD sessions), and termly Assessing for Mastery days.

In parallel to the training, teachers receive lesson plans and resources. This is done by granting English departments access to the English Mastery Library, where co-planning guidance, lesson-by-lesson resources and units of work can be downloaded.

The English Mastery link visits a school a minimum of three times a year to ensure the programme is implemented effectively. Schools that require further support to embed the programme may receive additional visits. This is done through teaching observations and discussions with members of staff that are at different levels of hierarchy within the English department and the wider school that are involved in delivering the intervention.

Where: Location of the intervention

The English Mastery curriculum is implemented in regular classrooms of participating schools in England.

When and how much: Duration and dosage of the intervention

English Mastery lessons are designed to be delivered six times a week in a way that fits the school timetable. The units run across each school term. There is an abridged version of the curriculum for schools that only have 4 or 5 timetabled hours of English. Guidance is available for these schools on how to adapt the curriculum to fit their timetable.

Tailoring

English Mastery is an evidence-informed intervention and optimal treatment fidelity is

emphasised. Schools are meant to deliver the curriculum as intended, including the three components of the programme (Literary Heritage, Mastery Writing and Reading for Pleasure). Nonetheless, implementers are encouraged to make intelligent adaptations to facilitate a sense of ownership and to respond to the children's mastery of a given concept. All teachers delivering the programme are trained in the subject-specific pedagogy. The training provides teachers with opportunities for co-planning with others, exposition, modelling and deliberate practice of lesson delivery, and instructional coaching of teachers' own classroom delivery.

Changes to delivery of English Masterty since Covid-19

The table below presents some of the changes to the English Mastery programme due to the outbreak of Covid-19 and resulting changes to schoolling.

Activity	Before March 2020/ Covid-19	Since March 2020/Covid-19
Induction training	2 days in person. Day one is attended by the EM Lead and SLT Champion. Day two is attended by the EM Lead and EM teachers	Replaced it with a fully digital offer, where teachers access some sessions as live Zoom sessions and the majority of the content as non-live, online modules in the form of PowerPoint videos. The training retained as much of the original content as possible. One significant change was that there was not a separate live session for the senior leadership team (SLT) Champions. Instead, they had additional non-live content to support their knowledge and understanding of both the programme and their role in its implementation. The induction was required for new teachers and offered for current teachers as a refresher.
School visits	One per term	2 x online development sessions (90 minutes each) per term.
Assessing for Mastery	One day of face to face training at the end of each term attended by the Mastery Lead comprising: Assessment Theory (90 mins) Standardisation (2 hours) Getting Ready to Teach (2 hours)	The day has been replaced by a combination of online sessions and asynchronous training: • Assessment Theory: Now taught through asynchronous training on the MyMastery platform[6]. • Standardisation: Now a 90 minute zoom session. • 'Getting Ready to Teach': Will be offered as an optional part of Spring visits. Delivery leads will monitor completion of the training through the online platform and online development sessions.
Recommended teaching	Literary Heritage = At least 100 mins+ per week.	Literary Heritage = 100 mins+ per week Mastery Writing = 50 mins a fortnight

[6] MyMastery is an online digital platform offering access to the teaching content: programme of study, planning and lesson resources, videos and tutorials and assessment tools.

	Mastery Writing = At least 50 mins per week Reading for Pleasure = At least 50 mins per week An assessment under standardised conditions at the end of each term.	Reading for Pleasure = 50 mins a fortnight from Autumn 2 Autumn 1 to be spent teaching transition modules provided by EM to accommodate the loss of learning caused by school closures. An additional option of a quiz, if an assessment under standardised conditions is not possible. These assessments are accompanied by reteach tasks to help students consolidate the essential knowledge.

All of the above measures will stay in place for the period of disruption caused by Covid-19, starting in March 2020. Currently this is expected to continue for the spring term 2021. The feasibility of conducting in-person visits and training will be reviewed ahead of the 2021 summer term.

Study rationale and background

The study builds on the pilot evaluation of English Mastery was conducted by The Brilliant Club which involved a total of 1700 pupils were involved. Of these, 851 pupils in seven schools received the intervention and completed GL's Progress Test in English (PTE) once a year over the duration of the intervention in Years 7, 8 and 9. In parallel, a group of 849 pupils in seven other schools made up the control group. The effect of the EM intervention was estimated at an additional four months of progress. The programme had similar effects for girls and boys but was particularly beneficial for pupils with lower prior attainment, and marginally more effective for non-Free School Meals (FSM) pupils[7]. This study builds on the pilot evaluation and is an randomised controlled trial (RCT) to evaluate the effectiveness of EM when delivered to Year 7 and Year 8 pupils.

Impact Evaluation

Research questions

The evaluation of English Mastery aims to answer the following research questions:

Primary Research Question

- RQ1. What is the impact of a two-year EM programme on the overall English attainment of participating Year 8 pupils in England?

Secondary Research Questions

- RQ2. How does the impact of a two-year EM programme differ by FSM status?
- RQ3. What is the impact of a two-year EM programme on the English Spelling, Punctuation and Grammar (SPAG) and reading comprehension attainment of participating Year 8 pupils in England?
- RQ4. What are the effects of the programme on teacher workload?

[7] Cheung, C. & Sirbu, I. The English Mastery Programme Pilot Study Impact Evaluation

- RQ5. What are the longer-term effects of the programme on GCSE English attainment and Attainment 8[8] of the participating pupils?

Exploratory Questions

- RQ6. How does the impact of a two-year English Mastery programme differ by number of hours of English lessons?
- RQ7: How does the impact of a two-year English Mastery programme differ by Key Stage 2 prior attainment in English?
- RQ8. How does the impact of a two-year English Mastery programme differ by gender?

Design

The evaluation is a two-armed cluster randomised controlled trial with 100 schools, half of them under the English Mastery intervention and half of them in a control group[9]. We aimed to target 110 schools, with an expectation that around 10% of schools will likely dropout before and after school minimisation, leaving around 100 at analysis stage. All pupils in the treatment group enrolled in Year 7 in the academic year 2019-20 are expected to receive the intervention for a period of two years.

The trial will incorporate Key Stage 2 English SATs results[10] as a baseline measure of academic attainment. To track progress in learning as a primary outcome we will measure participating pupils' English attainment in both treatment and control groups at the end of the intervention by administering the GL Progress Test in English[11] (PTE). Subscales of this same instrument will be incorporated as secondary outcomes in the present evaluation, to track progress in spelling, punctuation and grammar (SPAG), and reading and comprehension. We expect around 60 pupils per school to be tested in both treatment and control groups. Pupils taking the assessment will be randomly selected from all pupils in their year group in each school.

In addition to this, a school survey including questions on teacher workload will be implemented at different stages of the intervention in both groups (for details see pp.18-19 of the protocol). The workload questions in the survey will be asked of all Year 7 English teachers, including English Mastery Ambassadors in intervention schools. Measurements for this indicator will be reported directly by teachers as an estimate of number of hours worked in tasks involving Year 7 English. Results for this will also be included as a secondary outcome.

Furthermore, the trial will also explore the longer-term impacts of the programme on the GCSE English attainment of the participating pupils. To evaluate these, the trial will involve analysing GCSE English Language and English Literature data, as well as Attainment 8 results for subjects not related to English available from the National Pupil Database (NPD) in 2023. This last analysis will assess whether the programme has had a positive or negative impact on other subjects.

Schools assigned to the control group will implement a business-as-usual approach to teaching English to Year 7 pupils and may participate in English Mastery only after the end of the intervention two years later. An incentive of £1,500 is being offered to all control schools

[8] Attainment 8 measures the achievement of a pupil across 8 Key Stage 4 qualifications including mathematics, English, 3 further qualifications that count in the English Baccalaureate (EBacc) measure and 3 further qualifications that can be GCSE qualifications (including EBacc subjects) or any other non-GCSE qualifications on the DfE approved list. For more detail on how these are calculated see: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/561021/Progress_8_and_Attainment_8_how_measures_are_calculated.pdf

[9] As opposed to class level randomization, this design reduces the chances of contamination between treatment and control groups and the possibility of spill-over effects.

[10] This will be a composite score incorporating all English-related dimensions assessed in this test, including: grammar, spelling, punctuation, language strategies, reading and comprehension.

[11] For more information about the test see the following document: https://www.gl-assessment.co.uk/sites/gl/files/images/PT%20Series%20presentation%20-%2021%20April.pdf

to participate in the trial. The incentive is intended to mitigate the risk that these schools drop out from the trial.

Trial type and number of arms		Two-arm cluster randomised controlled trial with random allocation at school level (using minimisation)
Unit of randomisation		School
Minimisation variables		-School attainment (Mean proportion of pupils achieving strong 9 to 5 in both English and Mathematics GCSEs)[12] -School FSM (Proportion of pupils ever eligible for Free School Meals)[13] - School size (Total number of pupils in school)[14]
Primary outcome	variable	KS3 English attainment
	measure (instrument, scale)	GL Progress Test in English (PTE) overall score
Secondary outcome(s)	variable(s)	- Spelling, grammar and punctuation attainment - Reading comprehension attainment - Teacher workload - Long-term outcome: KS4 (GCSE) outcomes for the subject of English, Attainment 8 measure
	measure(s) (instrument, scale)	- GL Progress Test in English (PTE) SPAG subscale. - GL Progress Test in English (PTE) reading comprehension subscale. - KS4 (GCSE) English attainment Attainment 8 measure - Teacher workload survey

Randomisation

To facilitate timely programme training delivery among treatment group schools, schools needed to be assigned to treatment and control groups on a rolling basis before the beginning of the new academic year[15] (2019-20). Consequently, schools were allocated by lots utilizing a minimisation[16] process instead of simple randomisation. Minimisation was chosen as English Mastery is an intensive programme that has implications on staffing and timetables; therefore it was important to let schools know of their allocation as soon as they signed up for the trial, to facilitate the delivery of the programme during the school year.

Minimisation aims to balance treatment and control samples according to a pre-defined set of school observed characteristics (minimisation variables). Although the allocation of schools is not completely at random, this technique has the flexibility of allowing for the allocation of all schools to take place in several stages. Schools were allocated to treatment and control

[12] Proportion of pupils achieving strong 9 to 5 in both English and Mathematics GCSEs (schools' average between academic years 17/18 and 16/17).
[13] Schools' percentage of pupils eligible for FSM at any time during the past 6 years (academic year 17/18).
[14] Schools' total number of pupils in the latest academic year (academic year 17/18).
[15] In order to plan teacher training activities taking place before the beginning of the school year.
[16] The algorithm was first defined by Pocock, SJ. & Simon, R. (1975) Sequential Treatment Assignment with Balancing for Prognostic Factors in the Controlled Clinical Trial

groups utilizing minimisation between the end of March and July 2019. Four rounds of minimisation took place during that period.

Schools agreeing to participate in the trial were allocated to one of the two groups utilizing a 'minimisation' algorithm with three prognostic factors[17] and a 50:50 ratio between treatment and control groups. The minimisation procedure was undertaken with the aim of achieving balance across groups on the following factors:

- Mean proportion of pupils achieving strong 9 to 5 in both English and Mathematics GCSEs (schools' average between academic years 17/18 and 16/17)
- Proportion of pupils ever eligible for Free School Meals at any time during the past 6 years (academic year 17/18)
- Total number of pupils in school in the latest academic year (academic year 17/18)

The main reason why these factors were chosen is that, according to previous research, each one of them has explanatory value for the outcome of interest. When comparing units in treatment and control groups one would like to have balance between these groups in those school level characteristics that are likely to affect the outcome. This naturally occurs when a large number of units are randomised at once, but does not necessarily occur for a limited number of units. By implementing minimisation we ensure that samples in treatment and control groups were relatively balanced on those school characteristics.

Minimisation was undertaken using the *Minirand* package in R and both the *R* files (script) and output files were used to record the minimisation process. At time of minimisation, analysts were blinded to school identity. School identifiers were then merged with group allocation data after minimisation.

The minimisation process allowed NatCen to assess balance in the minimisation variables across treatment and control groups. If after the first two rounds of minimisation balance had not been achieved[18], then changes were made to the parameters of the minimisation process to ensure that balance across schools reaches acceptable levels for the most relevant variable (previous attainment), this was recorded and properly discussed in the statistical analysis plan.

Participants

All non-academically selective, state secondary schools in England were eligible for the trial, except for:

- Schools with less than 4 hours of English per week in Year 7 or 8;
- Schools classified as grade 4 according to OFSTED school classification;
- Schools that had previously partnered with English Mastery.

The justification for establishing these eligibility criteria is that, according to English Mastery, a minimum of 4 hours of English per week was needed to implement the programme. Similarly, it was established that low performing schools according to OFSTED classification[19] were deemed not to be in conditions to successfully implement the programme.

The English Mastery team identified and recruited eligible schools, prioritising recruiting schools with high numbers of FSM pupils. Every month, and as soon as a minimum number

[17] A prognostic factor in this context correponds to the variables utilized by the minimisation algorithm to allocate new units to treatment and control groups.
[18] We defined as an acceptable level of imbalance up to 0.1 standard deviations in the variable of interest.
[19] Classified in OFSTED category 4.

of new schools is recruited, minimisation wasperformed[20]. Headteachers were then informed about the allocation of their school to treatment or control groups. Settings were recruited by June 2019 signed an MOU confirming their commitment to delivering the programme as required and taking part in evaluation activities.

All Year 7 pupils in recruited schools were eligible for the trial. Participating schools were asked to enumerate all prospective Year 7 classes and share the trial information leaflet and privacy notice with all Year 7 pupils and their parents/carers. They were given two weeks to withdraw from the trial (and from data processing). After this period, schools were asked to share pupil level information for all pupils in the year group who did not withdraw. This group of pupils corresponds to the trial participants.

Schools were required to provide background information for all Year 7 pupils (who are trial participants) in academic year 2019-20. This included the Unique Pupil Number (UPN), School ID, date of birth, first name and surname. This pupil information will be collected in an Excel spreadsheet template and uploaded by schools using a secure NatCen website.

From the sample of trial participants (in both treatment and control schools), 60 individuals will be randomly chosen in each school to take part in testing at the end of the intervention. This group corresponds to the sample of evaluated participants. A limited sample of individuals per year group was chosen to reduce burden on schools, also looking to avoid unnecessary additional assessment costs[21].

Sample size calculations

		OVERALL	FSM
MDES		0.19	0.21
Pre-test/ post-test correlations[22]	level 1 (pupil)	0.69*[23]	0.62*
	level 2 (class)		
	level 3 (school)	0.30[24]	0.30
Intracluster correlations (ICCs)	level 2 (class)	0.15**	0.15**
	level 3 (school)	0.10**	0.10**
Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		2	2
Average cluster size		60	15***
Number of schools	Intervention	50	50
	Control	50	50
	Total	100	100

[20] In practice minimisation will be performed on a monthly basis on the assumption that at least 5 schools were recruited in that period.
[21] The trial was sufficiently powered for a sample of 60 pupils per school.
[22] Our sample size calculations include estimates of the proportion of variance explained through the included covariates at each of these levels, also known as R-squared. To our knowledge, there is no straightforward way to translate expected pre-test/post-test correlation into the expected R-squared for multi-level models so the R-squared values we use are approximations only. We report these in footnotes separately for each level.
[23] For our sample size calculations, we estimated the R-squared at level 1 to be 0.35 for both the overall sample and the FSM sub-group.
[24] For our sample size calculations, we estimated the R-squared at level 3 to be 0.10.

Number of pupils	Intervention	3,000	750***
	Control	3,000	750***
	Total	6,000	1,500***

* Based on figures for GL PTE and KS2 scores for the subject of English by Allen et al. 2018[25]

** Intraclass correlation coefficients assumptions based on qualitative assessment of previous EEF studies in KS3 summarised in Demack (2019)[26]

*** Recruitment of this project prioritised schools with a higher proportion of FSM pupils, therefore the estimated number of FSM pupils is at 25% of the total enrolment for this age group, higher than the national average of 14% of the total enrolment for this age group[27] in state funded schools. This figure is taken from the percentage of Year 7 pupils in the treatment group classified as 'FSM' at baseline in the 'Multiplicative reasoning professional development programme' trial[28]

As programme participation is voluntary and expected to last two years, some school attrition is likely before and after minimisation. Originally, the English Mastery trial was intended to recruit 110 schools and have an available sample size of 100 schools after accounting for attrition. We assume 60 pupils per recruited school will be randomly selected for testing at the end of Year 8.

School-level intra-cluster correlations (ICCs) are assumed to be close to those for Key Stage 3 English test scores for previous cohorts, while class-level ICCs are expected to be larger, as found on previous EEF three-level model trial.

The calculations were undertaken using 'PowerUp!'[29] and indicate that this study is powered to detect an effect of 0.19 standard deviations for the primary analysis (all pupils), and 0.21 standard deviations for the FSM subgroup analysis, based on the above assumptions.

Outcome measures

Pupil outcomes

The primary outcome is a standardised measure of pupils' English skills and reading comprehension, the GL Progress Test in English (GL PTE)[30]. Both the GL PTE SPAG subscale, which measures spelling, grammar and punctuation, and the 'reading comprehension' subscales will be included in the analysis as secondary outcomes. Pupils will be assessed in June 2021, at the end of Year 8 (second year of the intervention).

Schools have been asked in the school instructions pack that they receive ahead of the intervention to contact NatCen if their school uses the GL PTE test at any point in Year 7 or Year 8. Schools will also be asked about their plans to use GL PTE tests at any point in Year 7 or Year 8 in the first teacher survey in the Autumn term of 2019/2020. This is to record and monitor which schools use the test in Year 7 and 8 and at which timepoint. The expectation is that there will not be many schools that do use the test already and that these schools will be roughly evenly spread across control and treatment groups. Once we are able to

[25] Allen, R. et al (2018) Properties of commercial tests in the EEF database. Education Endowment Foundation Research Report.

[26] Demack, S. (2019) Properties of commercial tests in the EEF database. Education Endowment Foundation Research Report.

[27] Department for Education, Schools, Pupils and their Characteristics: January 2018 - National Tables

[28] Boylan, . et al (2015) Multiplicative reasoning professional development programme: evaluation Technical report & detail of evaluation findings. Pg 57. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/436833/RR406B_-_Multiplicative_Reasoning_Professional_Development_Programme.pdf

[29] Maynard, R. A. & Dong, N. (2013) PowerUP!: A Tool for Calculating Minimum Detectable Effect Sizes and Minimum Required Sample Sizes for Experimental and Quasi-Experimental Design Studies. Available from: https://repository.upenn.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1265&context=gse_pubs

[30] https://www.gl-education.com/products/progress-test-in-english-pte/

ascertain the number and spread of schools, NatCen can consider whether this will need to be controlled for statistically. If this is the case, details of this will be set out in the statistical analysis plan. We will ask these schools not to carry out their usual test in Year 8 and wait for the independent test. NatCen will then share the results with the schools.

NPD-derived Key Stage 2 English attainment will be used as a baseline measure of pupils' prior attainment. This data, together with gender and FSM eligibility information for the full trial sample will be obtained shortly after the beginning of the intervention for pupils in the treatment and control groups[31].

To evaluate the longer-term effects of the programme on participating pupils, we plan to obtain pupils' Key Stage 4 GCSE results for the subjects of English (English Language, English Literature) and an overall attainment across subjects (Attainment 8) from the NPD in Autumn 2023. NPD data will be linked to pupil data obtained from participating schools.

Teacher outcomes

A secondary outcome measure reflecting teachers' workload will be collected to better understand how EM affects teachers' daily practice. Data will be gathered from an online survey of Year 7 English teachers at different time points during the intervention, for treatment and control schools. The time points are shown in Table 2 below. The survey will include questions adapted from the teacher workload survey[32] to assess changes in the time spent preparing and teaching English for Year 7 pupils due to the implementation of the programme.

Analysis plan

Primary Analysis

The primary analysis will estimate the intervention's impact on English language skills[33], as measured by GL's Progress Test in English (raw scores, in line with EEF analysis guidance[34]), using an intention-to-treat approach. The analysis will use a three-level multilevel model with treatment assignment at the school level (level 3). This model will include school and class-level random effects and will account for baseline Key Stage 2 English (raw scores) and the variables used for the school minimisation process.

The basic form of the model is,

$$PTE_{ijk} = \beta_0 + \beta_1 Baseline_{ijk} + \beta_2 Intervention_k + \beta_3 Strata_k + u_{jk} + v_k + e_{ijk}$$

Where pupils (i) are clustered in classes (j) within schools (k). The intervention effect is estimated by β_2. $Strata_k$ represents a vector containing the variables (strata) utilized at minimisation. The term v_k is a school-level random effect, u_{jk} a class-level random effect, and e_{ijk} the error term, assumed to be normally distributed and uncorrelated with all the covariates included in the model. In line with the EEF analysis guidance, other additional covariates will not be considered at this stage. The three-level multilevel model will be followed if class level data is missing for less than 5% of pupils with outcome data. The analysis will be implemented in Stata 16 SE-64.

If the decision rule to implement the preferred three-level multilevel model is not met, the primary analysis will then estimate the intervention's impact on English language skills[35], as measured by GL's Progress Test in English (raw scores, in line with EEF analysis guidance[36]), using an intention-to-treat approach and use a two-level multilevel model with

[31] Autumn of the academic year 2019/20.
[32] https://www.gov.uk/government/publications/teacher-workload-survey-2016
[33] Spelling, grammar, punctuation and reading comprehension.
[34] EEF, 2018. Statistical analysis guidance for EEF evaluations
[35] Spelling, grammar, punctuation and reading comprehension.
[36] EEF, 2018. Statistical analysis guidance for EEF evaluations

treatment assignment at the school level (level 2). This model will assume school-level random effects and will account for baseline Key Stage 2 English (raw scores) and the variables used for the school minimisation process.

The basic form of the model is,

$$PTE_{ij} = \beta_0 + \beta_1 Baseline_{ij} + \beta_2 Intervention_j + \beta_3 Strata_j + u_j + e_{ij}$$

Where pupils (i) are clustered within schools (j). The intervention effect is estimated by β_2. $Strata_j$ represents a vector containing the variables (strata) utilized at minimisation. The term u_j is a school-level random effect, and e_{ij} the error term, assumed to be normally distributed and uncorrelated with all the covariates included in the model. In line with the EEF analysis guidance, other additional covariates will not be considered at this stage. The analysis will be implemented in Stata 16 SE-64.

It is important to note that using a two-level model, rather than a three-level model, may have possible implications for this evaluation. In a review of existing evidence, Demack (2019) showed that ignoring the class-level matters if class-level clustering is present. This is especially apparent in English secondary schools, where within-school pupil segregation is a common feature. Failing to account for this class-level clustering may mean we underestimate our standard errors, bring hidden bias into the design, and make a false-positive finding more likely.

In order to allow for comparability with other studies when reporting the effect size of the impact of the intervention, we will be using standardised scores (z-scores) for both the KS2 scores and the GL PTE assessment corresponding to baseline attainment and post-intervention attainment measures.

The impact of the intervention will be expressed as a standardised effect size using Hedge's g with 95% confidence intervals. Following EEF guidelines, the unconditional variance in the primary outcome for the pooled sample will be used when computing the Hedge's g statistic. If differential loss to follow-up creates an imbalance between trial groups or if attrition is high, the sensitivity of the estimated effect will be assessed by approximating missing outcomes using multiple imputation.

As detailed below, a second sensitivity analysis will include a full model, including a wider range of explanatory variables to increase statistical power.

Pupil FSM status

To assess whether the treatment effect varies between disadvantaged and non-disadvantaged children, we will be estimating the impact according to Pupil FSM status[37]. An appropriate statistical test of interaction will be incorporated to the analysis.

Sensitivity analyses

As part of the sensitivity analysis, a range of alternative models will be estimated to assess whether the findings for the primary analysis are robust to different model specifications[38]. The following five analyses will be carried out:

- An unadjusted analysis that will not include baseline covariates.

[37] Using the variable EVERFSM from the NPD.
[38] These analyses may include using multiply imputed data sets to ensure that the estimated effect is not biased as a result of missing data for outcomes; this may occur if differential loss to follow-up creates an imbalance between trial groups or if attrition is high.

- An adjusted analysis that will only include baseline attainment as a covariate (not including variables used for school minimisation[39]).
- A full model, including variables used at minimisation, as well as a wider range of explanatory variables to control for potential imbalance at baseline: pupil's FSM status, pupil's gender and term of birth; school setting, school type.
- A single-level OLS regression model including school- and class-level fixed effects, incorporating clustered robust standard errors (at the school level). The variables included will be the same as the primary analysis model: baseline attainment, treatment allocation and minimisation variables.

Compliance Analysis

EM will collect data during the implementation of the programme which will be used as part of our analysis of compliance.

The complier average causal effect[40] (CACE) will be estimated to show the impact of English Mastery on the primary outcome for pupils attending complying schools in the treatment group when compared to individuals in the control group.

Compliance in the intervention group will be determined at school-level. A composite measure of compliance will be constructed out of the following items, assessed termly in all participating schools in both year 1 of the programme and year 2 of the programme.

These items were identified by English Mastery as fundamental elements of the intervention. Their assessment by English Mastery forms part of a wider assessment of compliance and fidelity conducted by English Mastery links via termly school visits with each school participating in the English Mastery programme. The overall compliance judgement for each school will be based on English Mastery's assessment of compliance across all three terms each year, with a school judged to be meeting the overall compliance threshold for each item if it achieves the termly compliance threshold in least two out of three terms each year for that item[41].

For the compliance analysis, three measures of overall compliance will be constructed according to the extent of fidelity of implementation of key elements of the English Mastery Intervention as outlined in the "compliance threshold" column in the below table.

1) **optimal compliance**, where a school would be judged to be fully compliant if and only if it meets the overall compliance threshold on all 6 items (and judged non-compliant otherwise); and

2) **partial compliance**, where a school would be judged to be partially compliant if it meets or exceeds the overall compliance threshold on a minimum of 3 out of 6 items (and judged non-compliant otherwise).

[39] Namely: 1) mean proportion of pupils achieving A+ to C in both English and Mathematics GCSEs (academic year 2017/18); 2) proportion of pupils ever eligible for Free School Meals (academic year 2017/18); 3) total number of pupils in school (academic year 2017/18).

[40] Corresponding to the average effect of the intervention for those pupils who have complied with the programme.

[41] Please note, as school visits will be taking place at different times for different schools over the course of each term, English Mastery's compliance assessment will be to some extent based on reasonable expectation of compliance based on best available evidence at the time of the school visit. For example, if a school visit is taking place in week 5 of the term, compliance with the co-planning element of the intervention will be judged based on existing evidence on co-planning practice to date. E.g., English Mastery will use the school visit to assess whether co-planning is well established, regularly occurring and whether there is reasonable expectation that this will continue. If the school has not started co-planning at time of the school visit, English Mastery will conduct a follow-up with the school later in the term to review progress towards and achievement against the termly compliance threshold.

3) Covid-19 optimal compliance, a new measure that has been designed to account for the challenges schools may encounter to full EM implementation as a result of the Covid-19 pandemic and resulting school pressures. This measure will capture whether the critical components of the English Mastery's adapted programme have been met. A school would be judged to be fully compliant if and only if it meets the overall compliance threshold on all 6 items. Further details of these critical components can be found in the table below.

In each case, the compliance measure will be binary, with each intervention school having a value of either zero or one on each of these measures. We assume that none of the control group schools can feasibly implement more than one or two of the six English Mastery elements outlined above as part of their business as usual approach to teaching and assessments. Given that this is insufficient to reach the proposed thresholds for a control school to be considered non-compliant with treatment assignment, we will assume one-sided non-compliance in our analysis[42].

We will estimate three CACEs –using optimal, partial and Covid-19 compliance measures in order to identify a lower and upper bound of the likely true CACE of the intervention for those that comply with the intervention.

[42] This is in line with a substantive assessment by English Mastery that implementation of teaching approaches resembling one or two of the six English Mastery elements would be insufficient to deem the school to be implementing English Mastery in any meaningful way.

Item	Definition	Compliance threshold (measured termly)	Overall compliance threshold	Covid-19 compliance threshold
1. Induction training attendance	Teachers delivering English Mastery attend the English Mastery induction training.	80% of teachers delivering English Mastery attend induction training.	School meets or exceeds the termly compliance threshold two out of three terms in year one and two out of three terms in year two.	80% of teachers delivering English Mastery attend induction training. The overall compliance threshold remains the same as before
2. Assessing for Mastery training attendance	English Mastery Ambassadors attend the termly Assessing for Mastery training days.	English Mastery Ambassador attends termly Assessing for Mastery training.	School meets or exceeds the termly compliance threshold two out of three terms in year one and two out of three terms in year two.	English Mastery Ambassador attends termly Assessing for Mastery training. (These are equivalent to the MyMastery+ standardisation zoom sessions). The overall compliance threshold remains the same as before
3. Co-planning time allocation	Co-planning time is allocated on a regular basis.	Co-planning is allocated fortnightly or more regularly for at least 10 weeks per term.	School meets or exceeds the termly compliance threshold two out of three terms in year one and two out of three terms in year two.	Co-planning time is allocated on a regular basis (12 sessions in 2020/21).
4. Standardised conditions for assessment	Pupils sit the termly English Mastery assessments in standardised conditions.	Pupils sit the termly English Mastery assessments in standardised conditions.	School meets or exceeds the termly compliance threshold two out of three terms in year one and two out of three terms in year two.	Pupils complete the English Mastery quiz (one quiz per term). The overall compliance threshold remains the same as before
5. English Mastery curriculum delivery	Schools teach all components of the English Mastery curriculum.	School teach the Literary Heritage curriculum component for at least 100 mins+ per week for at least 10 weeks per term, and the Mastery Writing component for at least 50 mins per week for at least 10 weeks per term.	School meets or exceeds the termly compliance threshold two out of three terms in year one and two out of three terms in year two.	The school follows English Mastery's revised recommended structure (four lessons per week and with recommended content). The overall compliance threshold remains the same as before
6. English Mastery Abassador involvement	The English Mastery Ambassador teaches the English Mastery curriculum.	English Mastery Ambassador teaches at least two lessons per week for a minimum of 10 weeks per term.	School meets or exceeds the termly compliance threshold two out of three terms in year one and two out of three terms in year two.	The English Mastery Ambassador teaches two lessons per week for 27 weeks of the year [70%].

Other subgroup analyses

Subgroup analyses will be performed, to assess whether the treatment effect varies between boys and girls, according to the number of hours of English lessons or Key Stage 2 prior attainment in English. Reports for these results will be considered indicative and no adjustment for multiple hypothesis testing will be made.

It should be noted that these subgroup analyses are likely to lack enough statistical power due to small sample sizes, so the interpretation of their results may be limited.
Again, we will be including interaction terms to assess whether there is a difference on the effect of English Mastery for each one of these subgroups of pupils.

Secondary outcomes

For all defined secondary outcomes on academic attainment (please see list below), we will be measuring the impact of the EM intervention following an intention-to-treat approach estimating a basic model similar to that of the primary outcome analysis:

$$Outcome_{ijk} = \beta_0 + \beta_1 Baseline_{ijk} + \beta_2 Intervention_k + \beta_3 Strata_k + u_{jk} + v_k + e_{ijk}$$

Where pupils (i) are clustered in classes (j) within schools (k). The intervention effect is estimated by β_2. $Strata_k$ represents a vector containing the variables (strata) utilized at minimisation. The term v_k is a school-level random effect, u_{jk} a class-level random effect, and e_{ijk} the error term. For this measure we will also be using standardized scores (z-scores), reporting confidence intervals at 95% level, and the effect size utilizing hedge's formula as previously described. As in the case of analysis of primary outcomes, the three-level multilevel model, which is described above, will be followed if class level data is missing for less than 5% of pupils.

If the decision rule to implement the preferred three-level multilevel model is not met, for all defined secondary outcomes on academic attainment (please see list below), we will be measuring the impact of the EM intervention following an intention-to-treat approach estimating a basic model similar to that of the primary outcome analysis:

$$Outcome_{ij} = \beta_0 + \beta_1 Baseline_{ij} + \beta_2 Intervention_j + \beta_3 Strata_j + u_j + e_{ij}$$

Where pupils (i) are clustered within schools (j). The intervention effect is estimated by β_2. $Strata_j$ represents a vector containing the variables (strata) utilized at minimisation. The term u_j is a school-level random effect, and e_{ij} is the error term. For this measure we will also be using standardized scores (z-scores), reporting confidence intervals at 95% level, and the effect size utilizing hedge's formula as previously described.

The outcomes to be utilized are the following:

- Spelling, grammar and punctuation attainment score (GL Progress test in English)
- Reading comprehension attainment score (GL Progress test in English)

Teacher Workload

We will also report changes due to the intervention on teachers' workload, making use of some of the items present on the teachers' survey. These items will be selected from the Teacher Workload Survey based on whether time spent on these activities is likely to be influenced by the adoption of English Mastery. We are planning to derive a composite score measuring the total time spent across a number of teaching-related activities (preparing lessons, marking essays, supervising pupils, etc.). However, should we encounter high item non-response on these measures, we may need to analyse the time spent on each activity separately.

Descriptive statistics will be conducted as part of the analysis to investigate changes over time in the hours spent on different teaching-related activities and the overall teacher workload (measured through the composite score). We will use change scores to capture the overall change in workload over the two year period (survey round 1 and survey round 3). Wave 2 of the Teacher Workload Survey was originally planned to take place in May 2020. This would have allowed us to capture the incremental change in teacher workload in year 1 (survey round 1 and survey round 2 measures), and the incremental change in teacher workload in year 2 (survey round 2 to survey round 3 measures). However, schools and teachers were under a great deal of pressure following the changes to schooling as a result of the Covid-19 pandemic. The Wave 2 survey did not collect workload information in recognition that teachers' day-to-day duties were changed and atypical during this time (for example provision of home learning in place of in-school teaching). This analysis will involve comparisons between treatment and control groups to account for time trends in these measures.

If possible, an additional regression model including the total number of weekly hours spent on teaching-related activities may be used to estimate the impact of the intervention on teachers' workload[43]. Since it was considered important to measure changes on teachers' workload due to the implementation of the English Mastery programme, this analysis will also form a component of the process study described in more detail in the following section.

Long term Impacts

We will be reporting longer term effects of the intervention, by assessing whether the English Mastery programme had any impact in the following measures:

- Attainment 8
- GCSE outcomes for the subject of English (both Language and Literature)

These analyses will be performed in 2023, and its results will be published as an addendum to the final report, on the condition that no significant issues with implementation or experimental effects are observed (e.g. control group contamination, important differential attrition).

Implementation and process evaluation

Research questions

The main research questions that will be answered by the process evaluation are:

- What perceptions do teachers have about the EM programme design and how the programme is delivered?
- What are the key challenges in the delivery of the programme and what adaptations have been made?
- What are the perceived benefits of the programme for teachers? Do teachers believe they have changed their teaching practice? How do they think practice has changed?
- What are the perceived benefits of EM on pupils from the perspective of teachers in relation to 1) pupils' enjoyment of studying English and 2) progress in reading and writing?

[43] For this purpose, a three-level, mixed effects model, accounting for baseline measures utilizing the total number of hours worked (TeacherWorkload) could be used. The basic form of the model in this case is: $TeacherWorkload = \beta_0 + \beta_1 Baseline_{ijk} + \beta_2 Intervention_k + \beta_3 Strata_k + u_{jk} + v_k + e_{ijk}$. Where pupils (i) are clustered in classes (j) within schools (k). The average effect of the intervention in the number of working hours is represented by β_2.

- What are the local and national contextual issues, including changes to school regulations and inspections, that affect delivery and adaptation?
- What are the perceptions around the cost of the programme and value for money? What are the implications of these for wider take up of English Mastery?

Implementation dimensions

The key dimensions of implementation that the process evaluation will assess are: fidelity, dosage, reach, responsiveness, usual practice and adaptation. The IPE dimensions are addressed through the analytical process. Conducting an IPE involves synthesizing findings from data collected through different research methods to report on more than one of the IPE dimensions. As such, there is not one unique research method that responds to one dimension. We will be using a combination of observations, interviews and surveys as our methods for collecting data for the IPE.

Methods

For the qualitative research component, we will adopt a nested case study design. Data collection methods will be replicated across each case study unit and synthesised across all cases. The case studies will be complemented by a quantitative research component involving a school survey of both treatment and control schools. Due to the Covid-19 pandemic, associated social distancing restrictions and in an effort to reduce additional research burden on schools, some of the IPE methods and timing of data collection have been revised, as outlined in the table below.

Data collection timetable and overview (intended delivery and reviewed post-Covid)

		Pre-intervention		Early implementation		Ongoing delivery	Fidelity and sustainability	
Treatment schools	Intended	**June 2019** Observe EM induction training in London	September 2019: Teacher workload school survey	**February 2020** **12 Treatment schools:** Telephone interviews with EM Ambassadors	**May 2020** Teacher workload school survey	**October 2020** **12 Treatment schools:** Interviews with Englih teacher and the English Mastery Ambassador. One lesson observation per school	**April 2021** **12 Treatment schools:** Telephone interviews with Year 8 English teachers delivering the programme and EM Ambassadors	May 2021: Teacher workload school survey
	Reviewed (post-Covid)				**May 2020** School survey replaced with survey about how home learning is organised	**January 2021** Interview data collection mode changed from face to face interviews to telephone/online. Lesson observations moved to April 2021	**April 2021** Lesson observations in treatment schools moved from previous term	
Control schools	Intended			**February 2020** **9 Control schools:** Telephone	**May 2020** Teacher workload school survey		**April 2021** **9 Control schools:** Telephone interviews	

	Reviewed			interviews with Heads of English	**May 2020** School survey replaced with survey about how home learning is organised		with Heads of English	

Online school surveys

The survey data will be collectedthrough three rounds of data collection, collected over the two academic years during which the trial will take place. Due to the Covid-19 pandemic and resulting changes to schooling, the Wave 2 school survey was replaced with a shorter teacher survey to capture information about how remote home learning of English was organised at the school.

Data collected will be relevant to both the impact evaluation and the IPE. All schools in the trial will be invited to take part in the survey. All teachers delivering the EM programme (Year 7 in 2019/20 and Year 8 in 2020/21), including EM Ambassadors, and the SLT Champion in treatment schools along with the key contacts in control schools (Head of English and a member of staff with line management responsibility for English and/or teaching and learning within the senior leadership team) will be invited to take part.

The survey will collect key information on school characteristics and other school interventions aiming to support the development of pupil competencies in the subject of English for the relevant cohorts (to explore business as usual and co-interventions).

The survey for treatment schools will include questions related to intervention delivery. The survey will include role-specific questions for teachers, English Mastery Aambassadors and SLT staff members. Different members of staff will respond to some different questions within the same survey.

In addition, the survey will cover the following:

- September 2019: The survey will gather contextual information on school characteristics, the school's wider approach to English teaching and where English sits within the timetable. The survey will gather information on why the school has taken up EM. There will also be some key questions on teacher workload in this and subsequent rounds of the survey.
- May 2020: In treatment schools, the survey was meant to gather data on teachers experiences and views of implementation, key challenges encountered, adaptations made and the perceived benefits of English Mastery. Control schools were meant to be asked to provide information on delivery of English lessons and any changes made to lesson plans and activities during the academic year. Due to the Covid-19 pandemic, and resulting changes to schooling and teacher workload, the teacher workload survey planned for May 2020 will be replaced by a short teacher survey about delivery of remote learning English during school closures. It captures data on whether pupils are given home learning assignments, whether the work is marked and what resources teachers use to prepare for the lessons.
- May 2021: Schools will be asked to take part in this survey prior to pupil testing. In treatment schools, data will be collected on implementation, key challenges encountered, adaptations made and the perceived benefits of English Mastery. In addition, information on the cost of delivery and future sustainability will be sought. Control schools will be asked to provide information on delivery of English lessons and any changes made to lesson plans during the academic year.

In-depth qualitative case study research

The case study sample will include 12 treatment and 9 control schools.

Primary sampling criteria will include Year 7 cohort size, number of hours of English (4/5/6) and school's geographical location.

Additionally, we will also refer to the following date in developing our sample:

i) Data from a NatCen Survey administered to teachers
ii) Compliance data from school visits, including information on attendance to training
iii) Number of school visits and other additional support requested from English Mastery.

Case study data collection

Data collection will be conducted in four phases:

Pre-intervention (June 2019)

- We will observe training days one and two in London to assess how teachers, ambassadors and SLT champions engaged with EM, training methods and materials, and how information was cascaded to them.

Phase I - early implementation (February 2020)

- Telephone interviews with EM Ambassadors who took part in the training provided by the English Mastery team. EM ambassadors are central to early intervention in their schools and facilitating buy-in from their colleagues.
- In control schools, in-depth telephone interviews will be carried out with Heads of English to inform our comparison of English Mastery teaching with English teaching as usual. It will also help us to separate delivery issues that are specific to English Mastery from those that schools are facing more generally due to wider policy and regulatory factors.

Phase II – on-going delivery (January 2021 - previously planned for October 2020)

- The intention was to visit treatment schools to observe one EM lesson, interview a Year 8 English teacher and interview the English Mastery Ambassador to explore views on how early implementation issues (identified during Phase I) were addressed.
- Due to Covid-19 social distancing measures and to reduce research burden on schools, this data collection wave has been rescheduled. Interviews with English Mastery Ambasadors will take place by phone or online in January 2021.
- Planned visits to treatment schools to observe an EM lesson and interviews with the Year 8 English teacher will take place in April 2021.

  The interviews will now additionally explore any adaptations that the school has to English Mastery implementation in light of Covid-19 pandemic.

Phase III – fidelity and sustainability (April 2021)

- Observations of one English Mastery lesson in case study treatment schools have been rescheduled to take place in this phase of data collection (moved from the previous term). An observation template will be used to note EM delivery, pupil responsiveness and any observed adaptations.
- Teacher interviews in treatment and control schools will explore barriers to teaching the new/existing curriculum and whether/how these barriers have evolved over the past two years. For control schools, we will try to ascertain any changes/issues to delivering usual practice. These interviews will be conducted over the phone/online.
- English Mastery Ambassador interviews (twelve ambassadors) will explore (i) the extent to which teachers adhered to the programme and (ii) what factors enabled/hindered ambassadors.
- As part of these interviews, we will collect data on compliance including perceptions around pathway setting, what worked well and challenges to providing additional provision for weaker pupils, attending initial training, how co-planning worked in practice,

approaches as well as barriers and facilitators to delivering teaching on all three curriculum strands. We will also ask questions about teachers' confidence in relation to the topics that they teach and perceptions around 1) changes to their subject knowledge including how this may impact on classroom delivery of English teaching and 2) impact on pupil learning.

The NatCen research team will also create a log that the English Mastery team can populate throughout the intervention to record the key issues/questions that teachers share and discuss at the webinars of the Assessing for Mastery sessions.

Analysis

All of the qualitative case study data collection encounters will be used to address all the research questions. NatCen uses the Framework approach to qualitative data analysis, which facilitates robust qualitative data management and analysis by case and theme within an overall matrix. In the Framework approach, thematic matrices are developed through familiarisation with the data and identification of emerging issues. Thematic matrices represent key themes that are orientated towards the research questions. Each thematic matrix represents one key theme, with the column headings in each matrix relating to key sub-topics, and the rows to each ‘case’ (interviewee). Data from each case is then summarised in the relevant cell. The context of the information is retained as NVivo software facilitates the linking of verbatim data extracts to relevant themes and sub-topics in the Framework.

The survey data analysis will consist of reporting on overall survey non-response, generating descriptive statistics (such as proportions and counts) on relevant variables, and analysis of teacher workload. The latter is described in more detail in the Analysis Plan / Secondary Outcome / Teacher Workload Section above.

Cost evaluation

Cost information will be collected through the process evaluation from all intervention schools. When evaluating the per pupil cost of the intervention the approach set-out in EEF’s published guidance will be followed. Calculating the average cost of delivery enables comparisons to be made with other interventions based on both the average effectiveness and costs incurred. The total cost per pupil will be calculated based on information provided by schools in the school post-intervention survey about direct and indirect costs incurred.

Costs that will be collected and reported in monetary terms include:

Schools costs

Start-up costs (financial)

- Full-day induction training – teachers
    - Return travel per teacher
- Cost of teaching cover to release teachers for training
- Additional training day – English Mastery Ambassadors
    - Return travel per teacher
- Fee – Cost of the intervention paid by schools
- Materials –study texts, and photocopying

Developers costs (financial and time)

- Cost and time taken to set up all training events

- Cost and time of providing additional support to schools

We will ask schools to provide cost information in the survey at the end of the school year. To help maximise response, we will ensure questions are kept as simple as possible. We will follow EEF guidance to calculate costs over three years[44] by calculating future costs for the programme to cover a two-year total programme duration and dividing costs by two. We will also ask the developer about costs incurred at the end of the programme.

Ethics and registration

Process for ethical approval

NatCen has a robust ethics governance procedure. Research projects are scrutinised by the NatCen Research Ethics Committee (REC). The committee consists primarily of senior NatCen staff. If necessary, external research experts or professional experts ('lay people') may also be invited to review individual studies. Depending on the nature of the research and the perceived level of risk, projects undergo either an expedited review (scrutiny by the REC Chair) or a full review by the sitting REC.

For this evaluation we believe that a full review is appropriate given the scale of the project, the range of research of tasks and the age of the children and young people involved. The REC procedure is designed to provide ethical advice and guidance, and to ensure that all research undertaken by NatCen is ethically sound and meets the ethical standards of government and other funders. The process provides reassurance to potential research participants and, where relevant, to gatekeepers through whom they are approached.

The REC has reviewed the design of this project, provided guidance that has been incorporated into this final protocol, and will continue to be involved on an ongoing basis. For example, the REC will review any changes to the study and consent and recruitment materials as they are developed.

Parental and pupil agreement to participate in the evaluation

At the beginning of the 2019/2020 academic year, schools that signed an MoU will send out letters to pupils starting their Year 7 studies in that academic year and their parents/carers. NatCen provides the letter electronically; schools print the letter and send it out. The letter explains that the school is taking part in the study and provides the opportunity to opt their children out of participating in the evaluation. Specifically, the letter informs parents/carers and pupils that the trial will involve:

- The school sending data securely to NatCen. This will include their child's name, date of birth, their Unique Pupil Number and School ID.
- NatCen linking their child's data to the National Pupil Database (NPD).
- Some pupils taking part in an English assessment in June/July 2021.
- Data from the NPD being shared with NatCen and then stored in the Education Endowment Foundation's archive (which is managed by the Fischer Family Trust) and in anonymised form with the UK Data Archive.

NatCen will provide clear instructions to schools to maintain a log of withdrawals, requests for which could be made using a return slip attached to the letter, by email or letter or verbally by speaking to a member of school staff. If parents/ carers or pupils withdraw from the evaluation, schools will not provide any information in the pupils' information form to NatCen.

[44] EEF Guidance on Cost Evaluation March 2016

Data protection

NatCen has a range of policies and practices in place to ensure secure data handling. These are summarised below.

GDPR

The National Centre for Social Research (NatCen) is the data controller and data processor for this project. This means that we are responsible for deciding the purpose and legal basis for processing data. Under chapter 2, article 6 of the general data protection regulation , the legal basis for processing data is 'legitimate interest'. A general privacy notice has been published on the study page on NatCen's website. It will be shared with all schools participating in the trial and will be included among the information shared with parents/carers and pupils at the beginning of the 2019/2020 academic year. It can be found here:
http://www.natcen.ac.uk/taking-part/studies-in-field/evaluation-of-english-mastery/

All data collected for this study will be kept securely. We will safeguard the anonymity of all participants and no school, staff member or pupil will be named in any outputs or reports. School or pupils that decide that they no longer want to take part in the study may on request have their data deleted at any point and prior to the submission of a draft report to the EEF.

At the end of the research, all pupil data will be anonymised before being archived. Once the data is archived, EEF will take on the responsibility of data controller. All personal information, and any other data held on the project, will be securely deleted once the project is complete in April 2024.

Secure data handling

NatCen has a range of policies and practices in place to ensure secure data handling. These are summarised below. We categorise all data and files to 5 different levels, dictating how they are stored, handled and transmitted. The sample data for this study is Level 3 - 'Respondent Confidential'. Only those who carry out research tasks and those who need to check or process the data will have access to personal identifiable information. Our confidentiality measures for Level 3 data include:

- **Encryption**: All staff and freelancer laptops that hold Level 3 respondent confidential data have a hard drive encrypted using PGP Whole Disk Encryption by Symantec. This means that should the laptop be lost or stolen, the data contained on the hard drive is inaccessible. The encryption used by PGP is certified to FIPS 140-2 standards. We also use encrypted digital recorders for qualitative interviews.
- **Password Policy for office-based staff:** This includes 1) complex passwords that need to be changed every 30 days, 2) a password history of 10 previous passwords automatically enforced, and 3) account locked after 5 wrong attempts.
- **Access control:** Access to project data is managed via compliant segregation. There is a strict access control policy that is limited to named authorised individuals. In addition, unique serial numbers are assigned to avoid use of personal information.
- **Data Security Plans:** There will be a project data security plan in place detailing data security procedures. Rights of access will be recorded before they are granted.
- **File Systems Auditing** A file System Auditor is used to monitor activities logging what was created, updated, moved, renamed and deleted, and when.

NatCen processes for retention and destruction of personal data exceed ISO 20252 requirements on archiving and secure deletion.

Personnel

Delivery team at English Mastery:

The project is managed by Nick Wallace (Director of English Mastery), assisted by Emily Huballah (Project Manager), and Evie Newbold (Head of School Development).

Evaluation team at NatCen:

Daniel Phillips, Director of Evaluation at NatCen will provide quality assurance at design, analysis and reporting stages. Valdeep Gill and Dr. Priya Khambhaita (Research Director) will be the project leads with overall responsibility of project delivery. Priya and Val will manage the project team and provide advice, guidance and quality assurance for the design, implementation, analysis and writing up of the IPE.

The trial manager will be Robert Wishart (Research Director, Evaluation) who will lead on the design, implementation, analysis and writing up of the Impact Evaluation. Rodrigo will be assisted on all aspects of the Impact Evaluation by Enes Duysak (Senior Researcher, Evaluation) and Ben Stocker (Researcher, Evaluation).

Members of the research team include Anna Marcinkiewicz (Senior Researcher) who will be day to day project manager and support Priya and Val with project delivery. Ella Guscott (Researcher) and Harriet Read (Research Assistant) will collect and analyse qualitative data and support all other aspects of project delivery. All researchers will work closely with other departments and specialists at NatCen including statisticians and the Operations Department.

Risks

The main risks to the project are:

Recruitment of schools and pupils (high risk). Despite great efforts placed on recruitment, it might be difficult to recruit the 100 schools originally planned to power the study. As agreed with EEF, NatCen have considered this when advising the English Mastery team on the minimum number of schools that will be needed to be recruited and retained to secure a certain level of Minimum detectable effect sizes. The agreed minimum number of schools is also intended to account for the risk of school withdrawal between year one and two of the trial.

Covid-19 related risks (high risk). Additional pressures on school due to the Covid-19 pandemic (including school closures for most children and remote learning) may create additional and new risks for the project. These include: increased school withdrawal from the evaualtion; reduced school capacity to deliver the intervention as intended; increased pupil absences reducing dosage; challenges to intended data collection e.g. observations and in-person pupil testing. The evaluation team will take steps to mimise research burden on schools, conduct evaluation activites remotely and monitor these risks closely.

Non-participation in post-intervention pupil testing and teacher surveys, particularly among control schools (low/medium risk). There is a risk that schools are unwilling to participate in the post-intervention research tasks. This will be addressed by setting out clearly the requirements for the trial in the MoU; timing payments of the fiduciary incentive of £1,500 for control schools to be paid on completion of the post-intervention research tasks; and providing schools with clear instructions at the start of the project on what needs to be done and when.

High levels of pupil withdrawal (low/medium risk). Parents/carers and pupils will receive withdrawal forms at the beginning of the 2019/2020 academic year, and the evaluation team has assumed that only small numbers of families will return these forms based on other trial responses.

However, it is possible that parents/carers may not want their child to participate in the evaluation, or a large number of withdrawal forms may be received within certain schools. This is not typically a problem in EEF trials, but the evaluation team will monitor this closely.

Access to NPD data (low risk). We will need to access NPD data for our baseline measures of pupil attainment and FSM eligibility. The new GDPR compliant procedures and processes that are being implemented by NPD and the Department of Education are may result a delay in obtaining an NPD data extract. NatCen will monitor the processing times on other similar projects and inform EEF about any potential delays to the evaluation timetable due to data access.

Timeline

Dates	Activity	Staff responsible/ leading
March – July 2019	Allocation to intervention and control groups	NatCen
June – July 2019	Observations of ambassador and teacher training sessions	NatCen
August - September 2019	Statistical analysis plan	NatCen
September 2019	Delivery of intervention begins	English Mastery
September 2019	Contacting schools about collecting parental data processing opt out form	NatCen
September 2019	First survey	NatCen
September/ October 2019	Receive Year 7 pupil information from schools (student names, dates of birth, School ID and Unique Pupil Numbers (UPN))	NatCen
February 2020	*Case study schools:* Telephone interviews with English Mastery ambassadors in intervention schools and Heads of English in control schools	NatCen
May 2020	Second survey – refocused to understand how home learning is organised during Covid-19 school closures	NatCen
January 2021 (previously scheduled for October)2020	*Intervention case study schools:* Telephone/online interviews	NatCen
April 2021	*Case study schools:* Telephone/online interviews with teachers and EM Ambassadors in intervention schools and classroom observations. Interviews with Heads of English in control schools.	NatCen
May 2021	Third Survey	NatCen
June/July 2021	Administration of pupil testing (GL marking)	NatCen
July 2021	Intervention of English Mastery ends	English Mastery
July 2021	IDEA Workshop 2	NatCen
August – November 2021	Report writing	NatCen

Appendices

Appendix 1: Theory of Change