English Mastery
Evaluation Protocol

Sheffield Hallam University:
Dr Martin Culliney, Dr Karen Daniels, Sean Demack, Lewis Clark

Education Endowment Foundation

PROJECT TITLE[1]	Evaluation of English Mastery: A Randomised Controlled Trial
DEVELOPER (INSTITUTION)	Ark Curriculum
EVALUATOR (INSTITUTION)	Sheffield Hallam University
PRINCIPAL INVESTIGATOR(S)	Dr Martin Culliney
PROTOCOL AUTHOR(S)	Dr Martin Culliney, Dr Karen Daniels, Sean Demack, Lewis Clark
TRIAL DESIGN	Two-arm, three-level cluster randomised controlled trial with random allocation at school level
TRIAL TYPE	Efficacy
PUPIL AGE RANGE AND KEY STAGE	11-13, KS3
NUMBER OF SCHOOLS	72
NUMBER OF PUPILS	5184
PRIMARY OUTCOME MEASURE AND SOURCE	English attainment (GL Progress Test in English)
SECONDARY OUTCOME MEASURE AND SOURCE	Reading; Spelling, Punctuation and Grammar (GL PTE subtests)

Protocol version history

VERSION	DATE	REASON FOR REVISION
2.1	February 2026	Amended a formatting error on page 7 and updated the format to match the latest EEF template
2.0	July 2024	Revisions made: • Adding further detail on outcome measures and correcting a mistake in stratification measures listed (p8) • Updating sample size figures to match SAP (p9-12) • Amending compliance indicators to match SAP (p13-14) • Amending timeline for NPD application and SAP publication (p21)
1.0 [*original*]	October 2022	N/A

[1] Make sure that the project title here matches the title of the document. Please ensure that there is an identification as a randomised trial in the title as per CONSORT requirements.

Table of contents

Study rationale and background

English Mastery is designed to improve attainment in English by providing a substantial and ongoing programme of support of curriculum-integrated professional development for teachers. This has the potential to develop knowledge and expertise in English teaching, enabling teachers to deliver a coherent and cumulative Key Stage 3 English experience for pupils.

The English Mastery approach from Ark Curriculum Plus is distinct from other Mastery approaches in that it is underpinned by pedagogical pillars: accumulation of knowledge; discrete grammar teaching; Tier 2 vocabulary; and use of standardised, norm-referenced assessments. These are delivered through an English curriculum with the following strands:

- Literary heritage: introduces pupils to canonical texts and modern classics. The lessons follow a knowledge-rich approach to teaching English. English literature, the world, history, culture and heritage are taught, as are connections between these subjects.
- Writing Mastery: introduces pupils to grammatical concepts and rules in a logical and cumulative sequence. The rules are taught discretely from reading. The lessons follow an interleaved structure and pupils deliberately practise concepts to the point of mastery.
- Reading for pleasure: introduces pupils to shared reading of contemporary texts using a structured approach. This strand is used to teach vocabulary explicitly.
- Termly pupil assessment

Research evidence suggests that these pedagogical pillars and curriculum strands have the potential to impact on pupil attainment in English. The use of whole texts in English teaching is essential to engagement with texts and reading for pleasure (Cremin, 2014), and the Ark strategy ensures that pupils with weaker reading skills can access abridged texts and study at a 'foundation level', progressing where appropriate to a 'traditional level'.

Furthermore, teacher knowledge of children's literature and its use in the classroom is paramount to children's engagement with reading. Effective teachers of English have strong subject knowledge and plan purposeful writing opportunities with meaningful cross-curricular links. Where pupils gain automaticity in the technical skills of writing, such as spelling and punctuation, and where pupils have a good understanding of the use of grammar, they are more able to focus on the process of effective composition.

In addition, the English Mastery approach is underpinned by the principles and practices of metacognition (EEF 2021a). The following features of effective practice have been identified in secondary schools (EEF 2021b): the inclusion of targeted vocabulary; reading of complex texts through active engagement and application of existing knowledge; and the breaking down of reading of complex texts, and of complex writing tasks.

EEF previously funded an efficacy trial of English Mastery, yet this was severely affected by the Covid-19 pandemic and the closure of schools to most pupils from March to September 2020 and again from January to March 2021. Delivery of the programme was disrupted as schools were unable to conduct in-person lessons. Furthermore, it was not possible to collect the outcome

assessment data required for the impact evaluation. This element of the trial was therefore not completed.

The implementation and process evaluation was able to report findings despite the difficulties faced during the trial period. It showed that 80% of teachers in intervention schools responding to an end of intervention survey believed that English Mastery had reduced their planning and marking workload, 70% perceived that it had improved overall progress in English and 61% reported that their teaching of classic literature had improved. A majority of teachers were also satisfied with the programme overall, and with the training and support in particular.

Two key outcomes for pupils were increased enjoyment of English and a higher level of writing accuracy. However, some experienced teachers reported that the programme was overly prescriptive and limited creativity in lesson planning. Adherence to co-planning and formative and summative assessments was also found to be mixed.

Specific changes made to the English Mastery programme since the previous trial include:

- Further integration of assessment procedures and the use of these to inform planning.
- Improving emphasis and clarity on the importance of co-planning and how it contributes effectively to the high-quality delivery of the programme.
- Additional support for teachers in adapting planning to meet the needs of SEND, EAL and FSM pupils, or those at the lower end of the attainment scale.
- Review of texts included in the programme to include contemporary texts.

Pupil engagement in English Mastery

The implementation and process evaluation of the previously funded efficacy trial of English Mastery suggested an increase in pupil enjoyment of English. Literacy engagement involves motivation, decision making, thinking, effort expenditure, emotion and the commitment of time. The structured approach to teaching texts in English Mastery has the potential to strengthen pupil cognition and progress in reading and writing, and to improve pupil self-efficacy and confidence. Hempel-Jorgenson et al (2018) suggest a 'virtuous cycle' of increased confidence through self-efficacy that increases pupil reading motivation. Motivation is linked to appropriate goals, perceptions of self-efficacy (Guthrie and Wigfield, 2000) and belonging to peer groups and social relations (Wentzel and Wigfield, 1998, see also DfE, 2012).

Improving overall engagement in English has been seen as particularly crucial when supporting pupils who have fallen behind or those from disadvantaged groups, who may benefit from teaching that impacts positively on self-efficacy, motivation and engagement (Ng and Graham, 2018; 2006; Guthrie et al., 2007).

Intervention

Name

English Mastery

Why (theory/rationale)

The aim of the English Mastery Secondary programme is to improve achievement in the subject of English in Key Stage 3. This is done by providing teachers with subject-specific training, curriculum materials and ongoing in-school support and coaching.

Who (recipients)

Key Stage 3 pupils (for this trial, Years 7 and 8, ages 11-13).

What (procedures)

English Mastery is delivered through in-class teaching during regular timetabled English lessons as a whole-curriculum programme. It is a comprehensive programme with two pathways: the traditional curriculum for pupils working at the expected level for their age, and the foundation curriculum for those below the age-expected level. If pupils attain the age-expected level in two consecutive assessments, then it is recommended that they advance from foundation to traditional.

What (materials)

Weekly departmental co-planning materials are provided to aid teachers in adapting the English Mastery lessons for their pupils. Teachers also receive training and support through induction, additional training for the KS3 lead, remote development sessions, biannual school coaching visits, 'Assessing for Mastery' standardisation sessions each term and optional webinar workshops.

Pupil assessments, undertaken termly, are central to the intervention. Teachers have the chance to compare assessments with those from other schools. All teachers receive access to the MyMastery platform with the full range of curriculum resources and on-demand training materials, classroom footage and training events.

All pupils receive the same dosage and study the same topics, but the foundation curriculum is designed to be accessible to lower attaining pupils (such as through use of abridged texts).

Who (provider)

The intervention is designed by Ark Curriculum Plus, who also run the training and provide all ongoing support to the teachers delivering English Mastery in schools.

How (format)

Schools recruited to the trial and randomised into the intervention group send all KS3 English teachers for external training held by the developer. A member of the English department must be nominated to become the English Mastery Lead (EML) to lead the implementation of the programme in their school. The EML leads weekly co-planning sessions to support teachers to plan their implementation of the curriculum together. Teachers then deliver English Mastery through normal timetabled English lessons in school. This intervention therefore targets the entire pupil cohort in participating schools.

Where (location)

The intervention is delivered in classrooms. Schools in England with KS3 pupils studying English are eligible. This project is part of the DfE's Accelerator Fund (AF) and there is a requirement that a minimum of 50% will be recruited from three Regional School Commissioner (RSC) regions of England (West Midlands, East Midlands and the Humber, and the North[2]). Teacher induction training is provided offsite. Ongoing support is provided online or in school.

When and how much (dosage)

The sample will be split into two cohorts, recruited in Spring 2022 and 2023 respectively. For the first cohort, KS3 English teachers at schools randomly allocated to the intervention group for the first delivery cohort will attend English Mastery training in September 2022. English Mastery will then be delivered to pupils through their usual timetabled English lessons during the 2022/3 and 2023/24 school years. The 2022/23 Y7 cohort continue with the programme through Y8. The second cohort will follow the same timetable starting one year later, with two years of delivery beginning in Autumn 2023. Recruitment to the second cohort will begin from Autumn 2022, with the main bulk of recruitment activity expected to take place from January 2023. Randomisation will take place at Easter 2023, and again in June if necessary, with teacher training in June/July 2023.

Tailoring (adaptation)

The intervention is designed for all KS3 (Y7-9). However, due to funding constraints the current trial will only include Y7-8. This applies to both recruitment cohorts.

Control condition

Control schools in this trial should teach KS3 during the study period under business-as-usual conditions. Incentive payments totalling £1,250 will be paid to control schools for taking part in the trial. These will be paid in three instalments: £500 for submitting pupil and class data in Autumn 2022 (or Autumn 2023 for cohort 2), £250 for fulfilling all evaluation requirements by the end of the first year, and £500 for completing the pupil outcome assessment due to be taken in school during summer 2024 for cohort 1 and summer 2025 for cohort 2. Control schools are not able to purchase English Mastery on a commercial basis while the trial is ongoing.

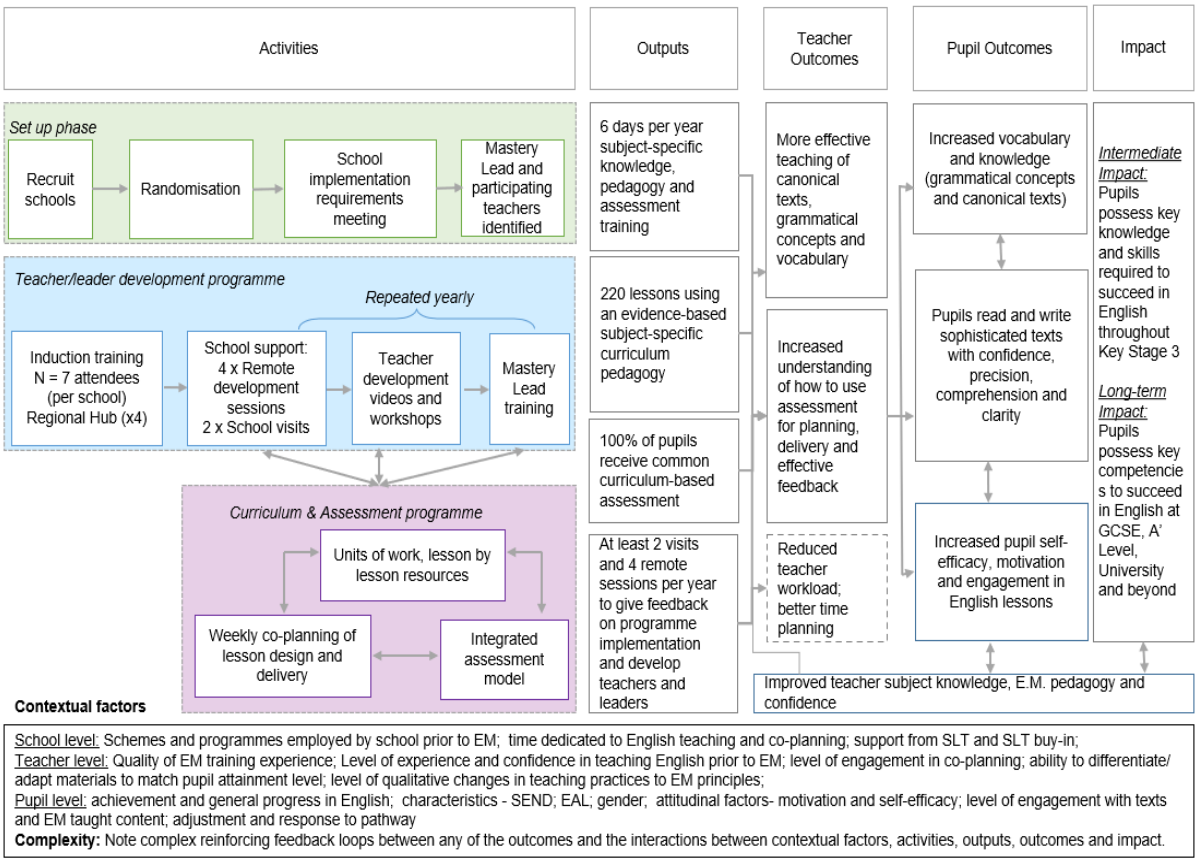
Theory of change and logic model

The Ark Curriculum Plus English Mastery approach is designed to improve attainment in English by providing a substantial and ongoing programme of support and curriculum-integrated professional development for teachers. This has the potential to secure the development of knowledge and expertise in English teaching, enabling teachers to deliver a coherent and cumulative Key Stage 3 English experience for pupils.

In defining a theory of change we draw on the work of Maxwell and Coldwell (2018, p269), who follow Merton in conceptualising theories as those "that lie between the minor but necessary working hypotheses ... and the all-inclusive systematic efforts to develop a unified theory" (1968, p.39). The aim is to understand "the mechanisms that mediate between the delivery (and receipt) of the program and the emergence of the outcomes of interest" (Weiss 1998, p.57). These mechanisms are detailed in the logic model presented in Figure 1.

[2] As defined by Regional Schools Commissioner boundaries (see here for local authorities included).

Figure 1: English Mastery logic model



Impact evaluation

Research questions

1. What is the impact of English Mastery on English attainment of Y8 pupils, as measured by the GL Progress Test in English (PTE)?
2. What is the impact of English Mastery on two dimensions of English (Spelling, Grammar and Punctuation, and Reading Comprehension) measured by the GL PTE?
3. What is the impact of English Mastery on English attainment of Y8 FSM pupils, as measured by GL PTE?
4. What is the impact of English Mastery on Y8 pupils with lower prior attainment, EAL, and SEN, as measured by GL PTE?

Design

This two-arm, three-level clustered efficacy trial will recruit and randomise schools in two cohorts (Y7 in 2022/23 and 2023/24). Pupils will be clustered into classes and classes clustered into schools. The baseline measure is prior attainment in English, as measured by KS2 test results from the NPD. The primary outcome is the total raw score on the GL PTE13, to be taken by Y8 pupils in summer 2024 (cohort 1) or summer 2025 (cohort 2).

Table 1: Trial design

Trial design, including number of arms		Two-arm, three-level cluster randomised
Unit of randomisation		School
Stratification variables (if applicable)		Geographic area Setting/streaming practice OFSTED rating
Primary outcome	variable	English attainment (overall)
	measure (instrument, scale, source)	GL Progress Test in English total raw score (0-65 scale)
Secondary outcome(s)	variable(s)	Reading Spelling, punctuation, and grammar
	measure(s) (instrument, scale, source)	GL Progress Test in English (Reading 0-32 scale, Spelling, punctuation and grammar 0-33 scale),
Baseline for primary outcome	**variable**	English attainment
	measure (instrument, scale, source)	KS2 Reading, GPS (combined, 0-120 scale)
Baseline for secondary outcome	**variable**	KS2 English KS2 Reading (0-50 scale), GPS (0-70 scale) combined, from NPD
	measure (instrument, scale, source)	KS2 Reading and GPS from NPD (as for primary outcome but separate scales)

Randomisation

A three-level design has been adopted, with pupils clustered into classes and classes clustered into schools. Two cohorts of schools were recruited, with the first taking part for two years from Autumn 2022-23, and the second taking part for two years from Autumn 2023-24. The original plan was to have only one cohort, but recruitment difficulties in Spring 2022 meant that more schools were needed to reach the target sample size. Within the first cohort, schools that had signed the MoU before 22 June were allocated to either the intervention or control group in two batches on that date, a second batch was allocated on 13 July 2022. The second cohort of schools was recruited during Spring term 2023 and informed of their allocation around Easter 2023.

Schools were allocated to the intervention or control group through minimisation, using the MinimPy software. School level minimisation minimises spillover risk. This approach increases the chance of good balance between the two conditions on the three key variables: four geographical areas (West Midland, North, East Midlands and Humber, other non-Accelerator Fund (AF) regions), setting/streaming practices (yes or no) and OFSTED rating (Outstanding/good or satisfactory/requires improvement). It also allows additional schools to be allocated in the second cohort while accounting for the characteristics of the first cohort.

Geographical region is included to produce a more even spread of schools across the participating areas. Setting and streaming practices are important due to the expected effect on class-level clustering in the outcome data. OFSTED rating has been used as an indicator of school quality as attainment data is not available for the two years prior to the trial.

The following data was collected from schools before randomisation: contact details for school contacts and English teachers, whether a school uses setting/streaming in KS3 English, the number of timetabled hours for English each week, and information about prior use of approaches similar to English Mastery. After meeting these requirements, schools were included in the allocation process.

Schools undertook to provide the full name, date of birth and unique pupil number for each pupil in the 2022/23 Y7 cohort at the start of the Autumn term so that KS2 test results and pupil characteristics can be obtained from the NPD. Schools must also specify the class and teacher for each pupil and provide an update on this in 2023/24. For the second cohort, Y7 data was submitted in Autumn 2023/24 with the Y8 update due in 2024/25.

Participants

The target was to recruit 100 schools, with at least 50% based in either the West Midlands, the North of England, or the East Midlands and the Humber regions[3]. Recruitment was led by the developer. It was agreed that schools partaking in the previous trial as members of the control group would be eligible for this trial if they had no exposure to English Mastery and were willing to accept the outcome of randomisation, although ultimately no such schools were recruited.

The recruitment target of 100 schools was not met by the original deadline in summer 2022, with only 39 schools recruited. It was agreed that these schools should start the trial as intended, with intervention schools taking part in English Mastery during the 2022/23 and 2023/24 school years.

[3] As defined by Regional Schools Commissioner boundaries (see here for a summary).

A second cohort was then recruited in 2022/2023, with a further 33 schools agreeing to take part in the study during 2023/24 and 2024/25. In the end, 72 schools were recruited across the two study cohorts. The power calculations presented below are based on this number. An alternative scenario with 10% attrition at the school level, leaving 66 schools, is also presented.

Statistical sensitivity is commonly illustrated using a Minimum Detectable Effect Size (MDES), the smallest effect a proposed RCT design could detect as statistically significant (often set as p<0.05) with a statistical power of 80% or higher. Equation 1, drawn from Kelcey et al. (2017), was used to estimate the MDES for this trial design. The equation has been used for estimating the MDES and the figures were checked with PowerUp! (Dong et al., 2015, sheet BCRA4_3f). Estimates are based on previous research into class level clustering (Demack 2019) and evaluations of KS3 maths (Boylan et al 2015, Demack et al 2021).

Equation 1: MDES calculations

$$MDES_{3LCRT} \sim M_{K-L-2}\sqrt{\frac{1}{P(1-P)}}\sqrt{\frac{ICC_{sch}(1-R^2_{sch})}{K}+\frac{ICC_{class}(1-R^2_{class})}{JK}+\frac{(1-ICC_{sch}-ICC_{class})(1-R^2_{pup})}{nJK}}$$

For this 3-level clustered design, the MDES is influenced by:

- n=number of pupils per class, allowed to vary between **3 (Estimated number of pupils eligible for Free School Meals, FSM)** and **12 (ITT sample**); J = number of classes per school (fixed at **6**); K = number of schools (**72**)
- P = proportion of schools allocated to intervention group (=0.5 at protocol, 0.45 at randomisation)
- L= number of (level 3) covariates used (which will include: group membership, school OFSTED rating, setting/streaming and 3 geography dummies, ~ **6** variables)
- M_{K-L-2} is the group effect multiplier value of the t-distribution for a 2-tailed test with alpha=0.05 & beta=0.80, equal variances assumed with (K-L-2) ~ (72-6-2= 64) degrees of freedom $\boldsymbol{M_{64}}$ for 72 schools (or $\boldsymbol{M_{57}}$ for 65 schools allowing 10% attrition)
- ICC_3 is the school level ICC (proportion of variance at level 3) ~ unknown but estimated at 0.10 (from first efficacy trial)
- ICC_2 is the class level ICC (proportion of variance at level 2) ~ unknown but estimated as 0.40 with the assumption that setting/streaming in KS3 English will be common across recruited schools
- R^2_{pup} = proportion of within-class pupil level variance that is reduced by covariate(s) - pupil level explanatory power ($\boldsymbol{R^2_1}$ **= 0.49** based on a correlation of 0.70)
- R^2_{class} = proportion of between classes within-schools variance that is reduced by covariate(s) - class level explanatory power ($\boldsymbol{R^2_2}$ **= 0.81** based on a correlation of 0.90)
- R^2_{sch} = proportion of between-schools variance that is reduced by covariate(s) - school level explanatory power – ($\boldsymbol{R^2_3}$ **= 0.16** based on a correlation of 0.40)

With a total of 72 schools (as randomised), the outcome testing would include 5,184 pupils (1,296 FSM pupils) in 432 classes and the resulting MDES estimates are 0.21 (overall) and 0.22 (FSM). These are the figures reported below (Table 2). The MDES does not improve with more than 12 pupils per class, which justifies the sampling approach as discussed below.

MDES estimates assume that the only difference between the intervention and control groups is their allocation (one will experience an intervention, the other will not). All other differences are assumed to be random. This assumption is reasonable at the point of randomisation and when there is little or no attrition but becomes weaker with increasing attrition. Therefore, indicative MDES estimates provide an indication of sensitivity but should be interpreted with caution. With 10% attrition (66 schools), the indicative MDES estimates are 0.22 (overall) and 0.24 (FSM). With 10% attrition from the 66 schools that submitted the pupil data required to take part in the study (59 schools), the indicative MDES estimates are 0.24 (overall) and 0.25 (FSM).

Sample size calculations

Table 2: Sample size calculations

		Protocol		Randomisation		SAP	
		OVERALL	**FSM**	**OVERALL**	**FSM**	**OVERALL**	**FSM**
Minimum Detectable Effect Size (MDES)		0.18	0.19	0.21	0.22	0.22	0.23
Pre-test/ post-test correlations	level 1 (pupil)	0.70	0.70	0.70	0.70	0.70	0.70
	level 2 (class)	0.90	0.90	0.90	0.90	0.90	0.90
	level 3 (school)	0.40	0.40	0.40	0.40	0.40	0.40
Intracluster correlations (ICCs)	level 2 (class)	0.40	0.40	0.40	0.40	0.40	0.40
	level 3 (school)	0.10	0.10	0.10	0.10	0.10	0.10
Alpha		0.05	0.05	0.05	0.05	0.05	0.05
Power		0.8	0.8	0.8	0.8	0.8	0.8
One-sided or two-sided?		2	2	2	2	2	2
Average cluster size		12	3	12	3	12	3
Number of schools	Intervention	50	50	35	35	32	32
	Control	50	50	37	37	34	34
	Total	100	100	72	72	66	66
Number of classes	Intervention	300	300	210	210	192	192
	Control	300	300	222	222	204	204
	Total	600	600	432	432	396	396
Number of pupils	Intervention	3600	900	2520	630	2304	576
	Control	3600	900	2664	666	2448	612
	Total	7200	1800	5184	1296	4752	1188

Randomisation was conducted prior to the collection of pupil data from schools as the developer wanted to inform schools of their allocation to allow adequate time to prepare their curriculum plans for the start of the following academic year. The exact number of pupils and classes in each school was not then known. As such, the figures provided here are based on the assumptions

held at that stage. Six of the 72 schools randomised (Cohort 1: intervention n=1, control n=2; Cohort 2: intervention n=2, control n =1) did not submit pupil data and will be counted as school-level attrition. However, as pupil data was not collected from these schools, it will not be possible to calculate their pupil attrition with complete accuracy. The 66 remaining schools constitute the baseline sample. Summary statistics on these schools are presented in the SAP.

Approach to selecting the ITT sample

Two approaches were considered. The first was to include all pupils in the participating Y7 cohorts. This would be costly because of the higher number of test papers to purchase but would allow the widest range of subgroup analyses due to the larger sample size. It may also be more practical for schools to test the whole year group rather than a selected sample. Moreover, it is possible that providing test papers and marks for all pupils could serve as an incentive for schools to participate, as they may find use for the results in tailoring support for their pupils.

A second approach would be to select a sample of the Y7 pupil cohort. This might comprise three classes (with an estimated 24 pupils per class) or a sample of pupils from across the Y7 cohort (potentially all classes, estimated at 6 classes of 12 pupils). Whilst selection at class level might seem preferable by minimising school burden, by the end of Y8 pupils in the selected classes would probably be dispersed across a larger number of classes anyway. Demack et al (2022 under review) found considerable pupil movement between Y7 and Y8 classes for an evaluation of a KS3 mathematics programme. This pattern could conceivably be replicated in KS3 English.

For the power analysis, we have used a stratified random sample of pupils drawn from across the cohort (6 classes of 12 pupils per school) to reduce testing costs. We estimate that there will be 6 FSM pupils per class; therefore 3 FSM pupils in each sample of 12 pupils taken from each class.

Clustering of PTE at school and class levels

We have assumed relatively high clustering of the PTE outcome at the class level as setting/streaming in Y7/8 English is likely to be common across recruited schools. Whilst there has been some recent evidence on class level clustering in KS3 maths (Demack, 2019), the situation for KS3 English is less clear. Data collected during school recruitment will provide some empirical details on school setting/streaming policy that can be used to update these power analyses for the Statistical Analysis Plan (SAP). We assume a class level ICC of 0.40 which means that 40% of the variance in the PTE outcome would be within schools, between English classes. At the school level, we assume an ICC of 0.10 (as used in the first trial). Our design assumes that a total of 50% of the variance in the PTE outcome would be at the structural level (predominantly between classes but also between schools) with the remaining 50% found within classes (either between pupils or between bands of pupils if within-class attainment banding is used).

Covariate explanatory power

Class level clustering presents a methodological problem; for a given sample size, stronger clustering results in higher MDES estimates (lower statistical precision). However, if this clustering is due to setting/streaming based on pupil attainment in English, the ICC value increases, and so does class level explanatory power. We have assumed a slightly smaller class level explanatory power (R^2_{class}=0.81) than was observed by Demack et al (2022 under review) for secondary mathematics because we assume that setting/streaming is less common in KS3

English compared with KS3 Maths. We assume a lower explanatory power at pupil (R^2_{pupil}=0.49) and school (R^2_{school}=0.16) levels.

Outcome measures[4]

Baseline measures

Whilst it would be ideal to use KS2 Reading and Grammar, Punctuation and Spelling as baseline measures, KS2 SATs did not take place in 2020 or 2021 and the possibility of tests being disrupted again in 2021/22 cannot be discounted. KS1 results will be used as baseline data if it is not possible to use KS2 from 2021/22. While KS1 scores are only available as a three-point scale, these tests will have been taken by the relevant cohort and data should be available from NPD in full, making this a reliable backup plan should problems affect KS2 in 2022.

Primary outcome

GL Assessment Progress Test in English will be used as the outcome measure. The paper version of this test will be used as schools are likely to favour this format. The outcome assessment will be conducted under usual test conditions during the summer term of 2024 for cohort 1 and summer 2025 for cohort 2. Schools will receive detailed instructions on test administration from the evaluation team, who will be responsible for delivering the test materials to schools and sending the completed papers to GL Assessment for marking. Printed guidance provided by the test publisher will also be sent to schools and the evaluation team will be available to provide support and respond to queries about the assessment process.

Secondary outcomes

Subscales from the GL PTE13 test used as the primary outcome will serve as secondary outcomes, as in the previous trial. Specifically, four subscales are provided in GL PTE13: Spelling, Punctuation and Grammar, Reading Comprehension Narrative and Reading Comprehension Non-Narrative. These will be combined into aggregated measures of 1) Reading and 2) Grammar, Punctuation and Spelling to match the two KS2 baseline measures. Using secondary measures derived from the primary outcome assessment reduces burden on school staff administering the test on behalf of the evaluation team.

Compliance

Compliance will be assessed at the school level. The compliance measures specified in Table 3 were used in the previous trial of English Mastery and will be adopted in the current evaluation, although minor changes have been made to item 2 and item 7 has been added to reflect developments in the expectations on participating teachers during the interim.

This information will be collected by the developer in the course of their work with intervention schools and will be shared with the evaluators at the end of the trial. Compliance will be defined as a binary variable, so schools are either compliant or not compliant. Schools must meet the threshold for each indicator to be classed as compliant.

[4] Please see the Statistical Analysis Guidance.

Table 3: Compliance indicators and thresholds

Item	Compliance indicator	Overall threshold
1. Induction training attendance	KS3 teachers attend EM induction training	80% of teachers delivering EM attend induction training.
2. Mastery Lead training attendance	EM Lead attends a termly training session (Mastery Lead workshop or Assessing for Mastery).	Attendance in at least 2/3 terms in Year 1 and in Year 2
3. Co-planning time allocation	Co-planning allocated fortnightly at least 10 weeks per term.	At least 2/3 terms in Year 1 and in Year 2
4. Standardised conditions for assessment	Pupils sit the termly EM assessments in standardised conditions.	At least 2/3 terms in Year 1 and in Year 2
5. English Mastery curriculum delivery	Schools teach Literary Heritage for 100 mins per week and Mastery Writing for 50 mins per week for at least 10 weeks per term.	At least 2/3 terms in Year 1 and in Year 2
6. English Mastery Lead involvement in teaching	EM Lead teaches at least two lessons per week for a minimum of 10 weeks per term.	The English Mastery Lead teaches two lessons per week for 27 weeks of the year [70%].
7. Support uptake	EM Lead facilitates touchpoints with their school	Four touchpoints each year: Two school visits and two development sessions

These variables will be used to estimate the Complier Average Causal Effect (CACE). The purpose of the CACE analysis is to estimate the impact of English Mastery for pupils deemed to have 'complied' with the intervention.

CACE will be estimated using two stage least squares (2SLS) regression (Gerber and Green, 2012). The first stage will model the pupil-level compliance variables using the same explanatory variables used in the headline ITT analyses along with additional school level items that are available via the school census as included in Table 3. This will be a multilevel logistic regression model used to generate predicted compliance (1 or 0) for use in the second stage model. The second stage models will use predicted compliance in place of the group identifier variable in the ITT analyses specified above to generate the CACE estimates.

Analysis

A multilevel approach will be adopted, with pupils clustered into classes and classes clustered into schools. Multilevel linear regression models will be constructed for the GL PTE primary outcome. KS2 English attainment (Reading and GPS combined) will be used as the baseline covariate for analyses of the primary outcome. The first model will only include the school level group identifier (an outcome only model). The second model will also include KS2 English as a covariate at the pupil, class and school levels[5]. The final model will also include variables used within the minimisation (geographical hub area, setting/streaming dummy and attainment). This final model will form the headline ITT impact analysis for the PTE primary outcome.

Follow-on ITT analyses will focus on the impact of the English Mastery programme on English attainment for disadvantaged pupils, as defined by the NPD variable EVERFSM_6. The same three model stages used for the headline ITT analyses will be used here. Further exploratory analysis

[5] These will be centred so that the school level will be centred on the mean for all 120 schools; the class level will be centred around the school mean; the pupil level will be centred around the class mean.

will be undertaken for the following subsamples: the two delivery cohorts, pupils with low prior attainment on KS2 English, pupils classed as EAL, and pupils classed as SEN.

For each model, the coefficient of the school-level dummy variable used to distinguish 'intervention group' pupils within the 60 schools who will receive the English Mastery programme from 'control group' pupils will be converted into Hedges' g effect size statistics with 95% confidence intervals.

Implementation and process evaluation

The evidence-informed logic model (Figure 1, p7 above) consists of the following activities:

Teacher Development programme

- Induction training (for teachers and school leaders)
- Remote development sessions x4
- School visits (at least 2)

Curriculum and Assessment programme

- Units of work, lesson by lesson resources
- Best practice videos
- Weekly co-planning of lesson design and deliver
- Integrated assessment model

The vision and impact of English Mastery includes:

For pupils

- Increased vocabulary and knowledge
- Reading and writing sophisticated texts with confidence, precision and clarity
- Increased self-efficacy, motivation and engagement in English lessons

For teachers

- Delivering more effective teaching using a knowledge-rich approach
- Delivering more effective teaching through more impactful use of planning time
- Reduced workload

The effectiveness of the English Mastery Teacher Development programme will primarily be evidenced by pupil outcomes in English, through the impact evaluation. The IPE design is underpinned by change processes presented in the logic model (Figure 1) and is designed to generate perception data focused on school leader and teacher experiences of the quality of training in English Mastery and its materials, and changes to their planning, teaching and assessment practices and workload. We will examine teacher perceptions of the impact of English Mastery on their confidence in teaching English and the perceived impact of English Mastery on pupil progress and self-efficacy, including those of different groups such as SEND, EAL, FSM. Pupils will also be asked to talk about their progress, confidence and enjoyment in English.

Contextual factors

Successful delivery of English Mastery may depend on a range of contextual factors at programme, school, teacher and pupil level. From our extensive experience of

evaluating school-based interventions (such as Culliney, Daniels, Coldwell, Booth and Demack, 2021, Demack et al forthcoming) we would expect:

- School-level factors (senior leader support for English Mastery; school buy in; alignment with other school priorities
- Teacher-level factors (motivation and skills of teachers, years of teaching English and previous training of English programmes)
- Pupil-level factors (achievement and progress in English, characteristics such as FSM, EAL, SEND, gender, attitudinal factors, adjustment to pathway, *foundation or traditional*)
- Programme factors (such as potential variation of quality of training, fidelity to programme and adherence to principles and practices of English Mastery
- Wider system factors (such as alignment with other policies around English teaching)

Research questions

a) Are any changes in self-reported teacher workload associated with English Mastery?
b) To what extent does the English Mastery training, its scheduling and materials equip teachers and leaders to implement a coherent and cumulative English curriculum?
c) In what ways has English Mastery led to changes in practice in the teaching of English, including for supporting the needs of different pupil groups (such as EAL, SEND and FSM, and those on traditional/foundation pathways)?
d) How have the English Mastery curriculum materials in particular led to changes in the teaching of English?
e) Have English Mastery assessment protocols led to changes in teacher use of formative and summative assessment to inform planning and teaching?
f) What are teacher perceptions of the impact of EM on the attainment and enjoyment of English for different groups of pupils?
g) What are pupil perceptions of the impact of English Mastery on their self-efficacy, confidence, and enjoyment in English?
h) What are school perceptions of the programme representing value for money?

Research methods

1. **Evidence review and discussion with stakeholders** to build an agreed evidence-informed logic model and agree data collection methods.

2. **Observations of training events**, conducted by experienced literacy specialists, focussed on engagement, alignment with expected content, and process. Our previous pilot evaluation of REclassIN for EEF developed a tool to check the focus, content and delivery of CPD against current evidence on effective CPD (Maxwell et al, 2018), which could be adapted for this evaluation. We would examine the CPD training and associated materials, and examine attendance at training events and other engagement with English Mastery support, to gauge fidelity.

3. **School visits** to understand fidelity, influences on implementation and the extent to which practice is aligned with expectations. We will look to visit a sample of schools which is balanced

on the characteristics used in the minimisation. As this is a two-year evaluation, our team of literacy experts will visit the same sample of 10 schools once during Year 7 and again in Year 8 to examine perceived impacts of the intervention across the study period. English specialists from the evaluation team will conduct a field visit in each of ten schools beginning in January 2023 (one term into the programme). These visits will be divided equally between the first and second cohorts of schools, so the first visits to intervention schools in the second cohort will begin in January 2024. This timing is intended to ensure sufficient opportunity for the intervention to have become embedded.

SHU will contact schools and request their participation in the fieldwork. The first visit will investigate schools' usual practice in English teaching, the quality of English Mastery training, support and materials, and also any contextual factors at school, teacher or pupil level that are either facilitating or inhibiting the success of the programme. Visits will be repeated in January of the second delivery year (when pupils are in Y8) at the same schools where possible. This will allow data collected one term into the programme at Y7 to be compared with data collected from the same schools at Y8. In addition to the focuses of the first visit in Y7, this second visit will provide detailed insights into schools' experiences of implementing and gaining knowledge and experience of delivering the English Mastery programme, capturing their ongoing confidence. We are also interested in pupil movement between the foundation and traditional stages of the programme, and any impact on pupil perceptions of self-efficacy.

During each school visit, SHU researchers will observe at least one English Mastery session. This will be followed by an interview with the teachers leading the lessons and a focus group of pupils (selected by the school) whose lesson was observed. Observations will be guided by a schedule designed to reflect the principles of practices of English Mastery. The EM lead in school will also be interviewed. Interview schedules will be guided by the outcomes of English Mastery from the evidence informed logic model. Findings will inform the post-intervention teacher survey.

4. **Pre- and post-intervention teacher survey** to examine understanding and practices around teaching English. The post-intervention survey will address contextual issues raised in the school visits and from the delivery team. The survey will also examine changes in classroom practice associated with English Mastery, with a focus on the 'pedagogical pillars' of the programme, teacher perceptions of their confidence in a mastery approach, and impact on workload. Where schools are identified as running multiple interventions to raise attainment in English for this age group, we will ask schools to discuss any perceived impacts, benefits or limitations that this has on the successful implementation of English Mastery. We will also explore any perceived impact on the English outcomes for pupils, practice in the teaching of English or teacher workload.

Analysis

Pre- and post-intervention surveys

Descriptive statistics will be generated from online survey data to summarise teacher experiences of the programme, and other issues such as workload that are relevant to the research questions. Analysis will include pre/post-intervention comparisons to assess change over the trial period. Comparisons between respondents in the intervention and control groups will also be conducted where possible.

School visits

Data analysis will be guided in how far observations/ materials reflect the features of effective pedagogy as described by Ark, and the principles and pillars of practice underpinning the English Mastery Programme. All focus groups and interviews will be transcribed. Inductive coding will be conducted using NVivo software to enable a cross-case analysis which will highlight patterns and recurring themes (Miles et al. 2019). In addition, this approach will facilitate a *variable-orientated* approach to case analysis, with a focus on variables and their relationships. As this analysis will be conducted during the first and second year of delivery for both cohorts, cross-case themes can be compared to examine change in teacher and pupil perceptions over time.

Table 4: IPE methods overview

Research methods	Participants/ data sources (type, number)	Data analysis methods	Research questions addressed	Implementation/ logic model relevance
Evidence review	Literature search, discussion with stakeholders			
Discussion with developers	Phone interview	Thematic inductive analysis	RQ b, c	Teacher Development Programme and Curriculum Assessment Programme
Observations of training events	EM trainers, teachers	Focus group English specialists	RQb	Teacher level outcomes/ quality and impact of training
School visits	Observations of EM lesson Teacher	Thematic inductive analysis	RQa-g	Teacher and pupil level outcomes
School visits	Interviews with EM lead and teacher/s		RQ a-f,h	Teacher and pupil level outcomes
School visits	Pupil Focus groups		RQg	Pupil level outcomes (self-efficacy, motivation, engagement)
Pre- and post-intervention teacher survey	Online survey	Descriptive statistics	RQa-f	Pre- business as usual, contextual factors Post- teacher and pupil level outcomes

Cost evaluation

Cost calculations will be based on data provided by the English Mastery developers. This will be conducted in line with EEF cost evaluation guidance to produce per pupil costs over three years. Specific items to be considered will include time for teacher training and preparation (but not delivery as this takes place during normal lessons) and costs of learning materials.

Ethics and registration

The evaluation was approved by the SHU ethics committee on 26 January 2022 (Ethic Review ID: ER40773698).

The trial has been publicly registered: https://www.isrctn.com/ISRCTN76965629

Schools will receive a participant information sheet that must be sent to the parents/carers of all pupils taking part in the trial before their data is shared with the evaluation team. This will be sent to schools once the pupils have started Y7 at the trial schools (in September 2022 for cohort 1, September 2023 for cohort 2). It will contain details about the intervention and the evaluation, a statement on data protection along with links to further documentation on data sharing and other relevant matters, and a slip for parents/carers to return to the school if they wish for their child to not take part in the evaluation.

Data protection

This EEF trial is part of a wider Department for Education (DfE) funded programme called the 'Accelerator Fund'. SHU is the data controller for the English Mastery evaluation, and the processing of personal data is defined under GDPR as a specific task in the public interest. As data is being processed for the purpose of academic research, the main aim of which is to improve reading ability among school pupils, the legal basis for processing is as a 'Public Task' (Article 6 (1) (e)). https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/public-task/

Special category data, specifically English as an Additional Language (EAL), Special Educational Needs and Disabilities (SEND) and Free School Meals (FSM) status, will be accessed from the National Pupil Database and processed for the purpose of scientific research as permitted under GDPR Article 9 (j). Pupil names, dates of birth and Unique Pupil Numbers (UPN) obtained from schools will be used to access this information. https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/special-category-data/

Specifically, we are processing this data to determine if the English Mastery programme has different effects on different subgroups of pupils. EEF was established with a remit to break the link between family background and educational attainment, and all EEF projects conduct subgroup analysis on FSM pupils. We are also interested in the effectiveness of the English Mastery programme for pupils who have EAL and SEND as we believe that it could be particularly beneficial to them. Further details can be found in the privacy notice: https://www.shu.ac.uk/-/media/home/research/sioe-rke/privacy-notices/eef-english-mastery-privacy-notice.docx.

Personnel

Dr Martin Culliney is Senior Research Fellow in SIOE. He has worked as project manager and data analyst on evaluations for several funders including EEF. Martin is lead author on EEF evaluation reports for the literacy interventions Integrating English (2019) and REACH Primary (2021).

Dr Karen Daniels leads developments in English across Primary and Early Years Initial Teacher Education in SIOE. Her research interests include literacy pedagogy and the social, emotional and cognitive dimensions of reading. She has worked with clients including Booktrust, Learning School Alliances and Multi Academy Trusts, and on SSIF projects including a focus on English Mastery. Karen led the IPE strand of the EEF REACH Primary trial and is a co-author of the evaluation report (2021).

Sean Demack is a Principal Research Fellow in SIOE and Deputy Head of Research Centre. He has extensive experience in educational RCTs. Sean was co-director on the EEF ScratchMaths trial and is currently directing the EEF Realistic Maths trial.

Lewis Clark is a quantitative researcher in SIOE. He has extensive experience working on large scale school-based research and evaluation projects. Lewis has also supported a number of RCTs, including the EEF-funded Adventure Learning trial.

Martin Illingworth is Senior Lecturer in English Education and course leader for Secondary English and Drama in SIOE. He is a consultant teacher with The National Association for the Teaching of English and an associate with Independent Thinking. Martin has published books including Creative Approaches to the teaching of English & Teaching English Language 16-19. He worked as expert advisor on a DfE SSIF Contextualised Grammar evaluation.

Dr Marie Helks is Senior Lecturer in English Education and has considerable expertise in teacher education and classroom pedagogy for English. She has published on grammar teaching and its impacts.

Risks

Risk	Mitigation	Adjusted risk
Staff departure	Very low turnover.	Low
Recruitment problems	Approach enough schools including control group from previous trial. GL PTE test data offered as incentive.	High
Difficulties in outcome test administration	Dedicated administrative staff in constant contact with schools via named contact, with ample time allocated to keeping in touch with schools throughout project.	Medium
School closures/lockdowns	Schools will be last thing to close as in past. If external visitors are restricted, fieldwork can be conducted remotely.	Medium
Pupil/teacher attrition	Whole classes/cohorts take part, so statistical sensitivity more affected by schools than individuals. Pupil and teacher data for each class collected at baseline and outcome. Contact with schools maintained throughout trial.	Low

Timeline

Table 4: Timeline

Dates	Activity	Who?
Dec 21-Jan 22	Set-up meetings Ethical approval Draft MoU, consent and information forms	All
Feb 22	Design IPE instruments IDEA workshop Evidence review	SHU
Feb-Jun 22	Data collection from schools (cohort 1) Pre-intervention teacher survey (cohort 1)	SHU/schools
Mar 22	Trial registration	SHU
Jun 22	Protocol	SHU
Jun 22	Randomisation 1	SHU
Sep 22	Teacher training/observations (cohort 1)	Ark/SHU
Sep 22	Pupil/class data collection Y7 (cohort 1)	SHU/schools
Oct 22-Jun 24	Delivery of English Mastery (cohort 1)	Ark/schools
Jan 23-Mar 24	Conduct IPE school visits (cohort 1)	SHU/schools
Feb-Jun 23	Data collection from schools (cohort 2) Pre-intervention teacher survey (cohort 2)	SHU/schools
Jun 23	Randomisation 2	SHU
Jun-Sep 23	Teacher training/observations (cohort 2)	Ark/SHU
Sep 23	Pupil/class data collection Y8 (cohort 1) & Y7 (cohort 2)	SHU/schools
Oct 23-Jun 25	Delivery of English Mastery (cohort 2)	Ark/schools
Jan 24-Mar 25	Conduct IPE school visits (cohort 2)	SHU/schools
Jun 24	Outcome testing (cohort 1) Post-intervention teacher survey (cohort 1) Collect fidelity data (cohort 1)	SHU/schools
Aug 24	Publish SAP	SHU/EEF
Mar 25	NPD Application	SHU
Jun 25	Outcome testing (cohort 2) Post-intervention teacher survey (cohort 2) Collect fidelity data (cohort 2)	SHU/schools
Sep 25	Data analysis	SHU
Jan 26	Report first draft	SHU
Apr 26	Final report	SHU

References

Allen, R., Jerrim, J., Parameshwaran, M. and Thompson, D. (2018) *Properties of commercial tests in the EEF database*. Available at: https://d2tic4wvo1iusb.cloudfront.net/documents/evaluation/methodological-research-and-innovations/Research_Paper_1_-_Properties_of_commercial_tests.pdf

Bloom, H.S. (2006) *The Core Analytics of Randomized Experiments of Social Research. MDRC Working Papers on Research Methodology*, August 2006. Available at https://www.mdrc.org/publication/core-analytics-randomized-experiments-social-research

Bloom, H, S., Richburg-Hayes, L. and Black, A. R. (2007) 'Using Covariates to Improve Precision for Studies That Randomize Schools to Evaluate Educational Interventions' *Educational Evaluation and Policy Analysis* 29(1), 30–59

Boylan, M., Demack, S., Willis, B., Stevens, A., Adams, G. and Verrier, D. (2015) *Multiplicative Reasoning professional development programme: evaluation*. Available at: https://www.gov.uk/government/publications/multiplicative-reasoning-professionaldevelopment-programme

Boylan, M., Wolstenholme, C., Demack, S., Maxwell, B., Jay, T., Adams, G. and Reaney, S. (2019) *Longitudinal evaluation of the mathematics teacher exchange: China-England*. Available at: https://www.gov.uk/government/publications/evaluation-of-the-maths-teacher-exchange-china-and-england

Coldwell, M. and Maxwell, B. (2018) 'Using evidence-informed logic models to bridge methods in educational evaluation', *Review of Education*, 6(3), 267–300

Cremin T., Mottram, M., Collins, F., Powell, S., and Safford, K. (2014) *Building Communities of Engaged Readers: Reading for Pleasure*. London: Routledge.

Culliney, M., Daniels, K., Booth, J., Coldwell, M., and Demack, S. (2021) *REACH Primary Evaluation Report*. Education Endowment Foundation. Available at: https://d2tic4wvo1iusb.cloudfront.net/

Davies, M., Read, H. and Guscott, E. (2022) *English Mastery Evaluation Report*. https://d2tic4wvo1iusb.cloudfront.net/documents/projects/English-Mastery-Final.pdf?v=1651210865

Demack, S. (2019) *Does the classroom level matter in the design of educational trials? A theoretical and empirical review*. https://educationendowmentfoundation.org.uk/public/files/Publications/Does_the_classroom_level_matter.pdf

Demack, S. (2021) The statistics of pupil segregation in England. *Radical Statistics* 129, 45-69 https://www.radstats.org.uk/no129/Demack129.pdf

Demack, S., Boylan, M., Culliney, M. and Wolstenholme, C. (forthcoming) *Evaluation of Realistic Maths*. Available at: https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/realistic-maths-education

DfE (2012) Research Evidence on Reading for Pleasure. Crown Copyright.

Dracup, T. (2014) "The Politics of Setting." Available at: https://giftedphoenix.wordpress.com/2014/11/12/the-politics-of-setting/

Dong, N., Kelcey, B., Maynard, R. and Spybrook, J. (2015) *PowerUp! Tool for power analysis*

EEF (2021a) Metacognition and self-regulated learning. Guidance report. https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/metacognition

EEF (2021b) *Improving Literacy in Secondary Schools*. https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/literacy-ks3-ks4

Guthrie, J. T. and Wigfield, A. (2000). 'Engagement and motivation in reading' in M.L. Kamil, P.B. Mosenthal, P. D. Pearson, and R. Barr (Eds.), Handbook of Reading Research (3rd Ed.). New York: Longman.

Guthrie, J.T., Hoa, L., Wigfield, A., and Tonks, S. (2007) 'Reading motivation and reading comprehension growth in the later elementary years' Contemporary Education Psychology. 32(3), pp282-313

Hempel-Jorgensen A., Cremin, A.T., Harris, D. and Chamberlain, L. (2018) 'Pedagogy for reading for pleasure in low socio-economic primary schools: beyond 'pedagogy of poverty'?' *Literacy* 52(2), pp 86-94.

Kelcey, B., Spybrook, J., Phelps, G., Jones, N. and Zhang, J. (2017) 'Designing large scale multisite and cluster randomized studies of professional development' *The Journal of Experimental Education*, 85(3) pp389-410

Maxwell, B., Daniels, K., Demack, S. and Stevens, A (2018) *Mastery English and Maths. Learning Unlimited Teaching School Alliance (LUTSA): Strategic School Investment Fund (SSIF) Round 1 Project Evaluation*

Ng, C. and Graham, S. (2018) 'Improving literacy engagement: enablers, challenges and catering for students from disadvantaged backgrounds' *Journal of Research in Reading* 41(4), pp 615-624

Sedgewick, E. (2001) *Teaching Literacy: A creative approach*. London: Continuum.

Wentzel, K. and Wigfield, A. (1998) 'Academic and Social Motivational Influences on Student's Academic Performance' *Education Psychology Review*. 10(2) pp. 155-175