Using the Power of Reading to improve writing attainment, a two-armed cluster randomised trial
Evaluation Protocol

Evaluator (institution): Manchester Metropolitan University
Principal investigator(s): Steph Ainsworth and Stephen Morris

Template last updated: December 2022

Evaluation summary

Project title	**Using the Power of Reading to improve writing attainment, a two-armed cluster randomised trial**
Developer *(Institution)*	Centre of Literacy in Primary Education
Evaluator *(Institution)*	Manchester Metropolitan University
Principal investigator(s)	Steph Ainsworth and Stephen Morris
Protocol author(s)	Sandor Gellen, Stephen Morris, Steph Ainsworth, and Helen Lawson
Trial design	Two-arm cluster randomised controlled trial with random allocation at school level
Trial type	Efficacy
Pupil age range and Key stage	9-10 years old, KS2
Number of schools *(at design stage)*	140
Number of pupils *(at design stage)*	3,920
Primary outcome measure and source	Writing score (KS2 NCT pre-2013 tests, Two bespoke writing tasks)
Secondary outcome measure and source	Writing self-efficacy (Experiences of Reading and Writing questionnaire) Writing enjoyment (Experiences of Reading and Writing questionnaire) Reading engagement (Experiences of Reading and Writing questionnaire) Reading attainment (NPD)

Protocol version history

Version	Date	Reason for revision
1.2 [*latest*]		
1.1		
1.0 [*original*]	03/02/2026	N/A

Contents

Evaluation summary 1
Protocol version history 2
Study rationale and background 4
Intervention 6
Impact evaluation design 9
Implementation and process evaluation (IPE) design 23
Cost evaluation design 38
Ethics and registration 38
Data protection 39
Personnel 41
Timeline 45
References 47

Tables

Table 1 Trial design 9
Table 2 Primary and secondary outcome measures with relevant baseline covariate 16
Table 3 Sample size calculations 17
Table 4 Overview of case study visits 27
Table 5 IPE methods overview 33
Table 6 Risks 42
Table 7 Timeline 45

Figures

Figure 1: Logic model 8

Study rationale and background

The Power of Reading (PoR) is a literacy programme developed by the Centre for Literacy in Primary Education (CLPE) to help teachers understand the deep connection between language, reading, and writing. Through training and resources, the programme supports primary schools in developing a high-quality, cohesive English curriculum. By using high-quality children's literature, PoR fosters oracy, reading fluency, comprehension, writing stamina, and composition from Early Years Foundation Stage to Year 6.

The programme aims to develop teachers' understanding of the ways in which high quality texts can be used to support pupils understanding of an authentic writing process, ultimately leading to gains in writing attainment. The key pedagogical principles underpinning PoR include: making explicit the link between reading and writing, and the ways that texts are structured using high quality examples; explicit modelling of authentic writing processes; integrating creative approaches that develop oracy, reading comprehension and writing (including book talk, storytelling and drama); and exploring a range of different text types written by a range of authors. Participating teachers are provided with training and access to 12 high-quality children's books along with online resources, including teaching plans, which are designed to develop their subject and pedagogical knowledge, while reducing teacher workload.

The design of PoR aligns with the broader literature around effective teaching of reading and writing. For example, the 2024 EEF Practice Review highlighted the importance of high-quality texts in the teaching of writing (Grima et al., 2024). The EEF Evidence Review found support for programmes which: explicitly model the processes involved in writing across a range of genres (e.g. Torgerson et al., 2018); introduce writing conventions within the context of creative writing (e.g. Kim et al, 2011); make strong links between reading and writing (e.g. Olsen et al., 2017); and involve in-depth training for teachers which allows them to develop a strong understanding of the writing process themselves (e.g. Kim et al., 2011). Each of these features is embedded within PoR.

A recent independent evaluation of PoR conducted by Leeds Trinity University (Doherty, 2019) showed promising results, with pupils who received PoR making, on average, six months' additional progress in reading compared to the control group. However, the evidence is limited due to a small sample (11 primary schools) the inclusion of other components in the intervention group package, such as governor training and tailored support for teachers and TAs, so the observed impact cannot be solely attributed to PoR. Additionally, the study was not an RCT: the term 'control group' is used in the report, but it appears to refer to a general comparison or benchmark group - no methodological explanation is provided regarding how this group was selected, matched, or defined - which further limits the strength of the findings. The related Power of Pictures programme[1] (PoP; Anders et al., 2021a), found that pupils exposed to PoP made, on average, one month of additional progress in writing compared to pupils in the control

[1] Another CLPE programme designed to boost children's engagement and achievement in reading and writing – focusing on developing visual literacy, enhancing reading comprehension, and fostering creative writing skills.

group[2]. Gains were also found for writing self-efficacy and writing ideation. PoR is a more comprehensive programme than PoP, covering a broader range of text genres and providing more in-depth training to teachers, yet underpinned by the same overarching principle of using high quality texts to develop pupils' confidence and proficiency as writers.

Few writing-focused programmes have gained widespread recognition, but PoR stands out as one of the most well-known, having been delivered in over 4,200 schools and worked with 5,500 teachers over the past 20 years. CLPE report 1,610 schools having CLPE membership in 2021/2 with 259 teachers having attended PoR training (CLPE, 2023). Case study and survey data within the same report suggests that the programme is well received by schools, who perceive PoR to have multiple benefits for teachers and pupils. These findings are consistent with qualitative evidence reported by Leeds Trinity University (Doherty, 2019).

The programme addresses multiple priority areas, including reading, writing composition, structured step-by-step guidance, writing conventions, and genre-based teaching. Additionally, PoR aligns with the EEF's broader mission on attainment as its structure, pedagogical principles, training content, and resources equip teachers with effective assessment strategies to better understand their pupils' needs. This enables teachers to tailor instruction, enhance engagement, and ultimately improve literacy outcomes.

In summary, there is promising evidence that PoR may positively impact pupil outcomes. However, there is a need for a robust, large-scale evaluation of PoR, which investigates its impact on pupils' writing. This study aims to build on emerging evidence regarding the potential efficacy of PoR.

The present study will comprise an impact evaluation (IE) in the form of a two-arm, parallel, cluster randomised controlled efficacy trial, in which whole primary schools are randomly assigned to either the intervention or control group. Alongside this, an implementation and process evaluation (IPE) will examine how the delivery of the programme influences its impact — or lack thereof— on pupil outcomes within the IE. The IPE will be guided by the programme's theory of change.

In each participating school, one Year 5 class (and their teacher) will be randomly selected to participate in the trial. Half of the selected classes (and their teachers) will then be randomly assigned to the intervention group, where they will be invited to implement PoR. The primary outcome of the study will be writing attainment, assessed through writing tasks completed by all pupils in selected classes (in both intervention and control schools) at the end of the intervention period.

The study will be conducted across England with no geographical restrictions beyond schools having to organise travel to the training locations, and will follow a split cohort design. The split cohort design was chosen because the delivery team faced challenges in recruiting the required number of schools within a single recruitment period. Half of the schools will be recruited,

[2] The same was found for pupils eligible for free school meals (FSM)

randomised, and take part during the 2025/26 academic year, with outcome assessments taking place in summer 2026. The remaining schools will be recruited between February and July 2026, and participate in the trial in the 2026/27 academic year, with assessments conducted in summer 2027[3]. The impact of PoR will be evaluated by comparing average writing attainment scores between the intervention and control groups in a single analysis that includes pupils from both cohorts.

Intervention

In each participating school assigned to the intervention group, the teacher of the selected Y5 class will attend four in-person training days across one school year. This allows them to directly apply what they've learned, refine their teaching practice, and evaluate its impact on pupils. School leaders can also join for the first and last sessions.

The training days equip teachers with a comprehensive understanding of a high-quality literacy curriculum. They explore how to use poetry, picture books, non-fiction, storytelling, and drama to deepen children's engagement with reading and writing. Training sessions model evidence-informed and creative teaching strategies, including:

- Oracy: book talk, storytelling, drama, poetry performance, debate, and argument.
- Reading comprehension: reading aloud, re-reading, book talk, text-marking.
- Writing: storytelling, illustration, drama, shared writing.

Teachers also learn how to model the authentic writing process and expose pupils to a range of text types, illustrative styles, themes, and subject matter from diverse authors.

The training is delivered in person by a CLPE Primary Advisory Teacher alongside an experienced classroom teacher who has successfully implemented the programme. Sessions take place at CLPE's training centres in London, Manchester, Leeds, and Nottingham.

Each participating teacher from the intervention group receives a set of 12 high-quality children's books tailored to their year group. Additionally, a second set of three theoretical books (written by / in collaboration with CLPE) helps deepen teachers' subject knowledge in effective literacy instruction. Schools also gain access to an online suite of teaching units and resources, which includes:

- Detailed teaching plans for each text to support curriculum design and whole-class instruction.
- Curriculum maps to help teachers build their English curriculum.
- Asynchronous online CPD units to support specific aspects of English teaching

[3] At the time of writing this protocol, we confirm that 69 schools have been randomised into Cohort 1. We therefore aim to recruit 71 schools for Cohort 2.

These resources aim to reduce teacher workload by providing structured guidance for classroom implementation. Teachers will also have the opportunity to hear from two published authors so that they can understand authentic creative processes.

Every training session includes gap tasks that are designed to help teachers refine their practice, trial new texts and approaches, and evaluate their impact on pupils and school culture. Tasks may involve:

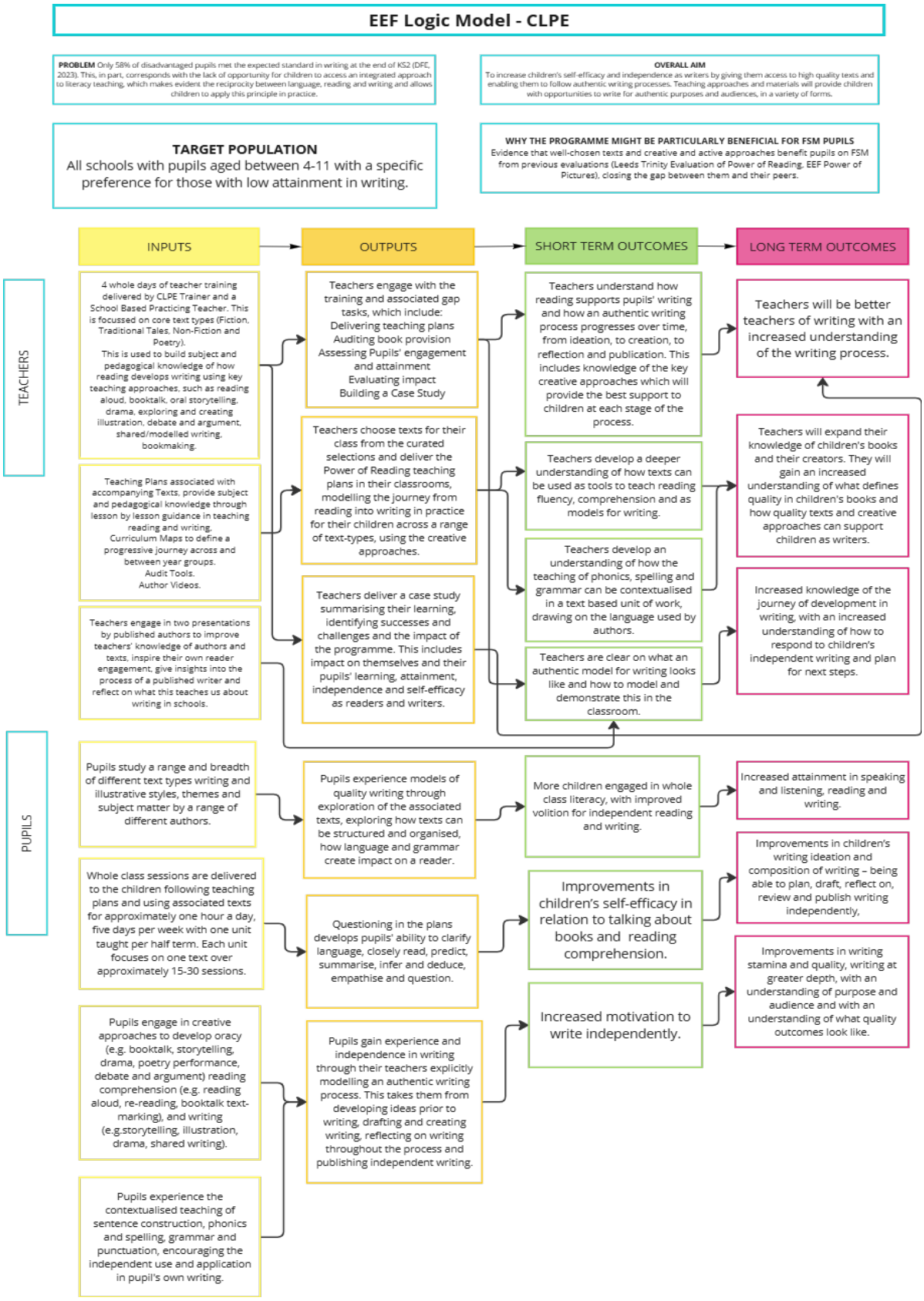
- Delivering a teaching plan.
- Auditing book provision, environment and routines.
- Assessing pupil engagement and attainment.
- Developing a classroom-based case study.

These tasks also are meant to enable teachers to apply their learning in the classroom, improving their confidence and effectiveness over time. On the final training day, teachers present a case study demonstrating how they have embedded their learning into school practice and its impact on pupils' progress, engagement and attainment in reading and writing.

The POR programme does not mandate a complete curriculum overhaul. Instead, it provides tools, resources, and guidance to help schools adopt a text-centred approach to literacy, incorporating proven creative teaching strategies. Schools can tailor implementation based on their context and needs.

Figure 1 below shows the Logic Model developed by the delivery team with facilitation from EEF ToC workshops.

Figure 1: Logic model



Impact evaluation design

Research questions

The purpose of the impact evaluation is to evaluate the effect of PoR on: (1) writing attainment, (2) writing self-efficacy, (3) writing enjoyment, (4) reading engagement, and (5) reading attainment.

Primary research question (end of Year 5)

1. What is the difference in **writing attainment** between pupils exposed to a teacher trained in the Power of Reading compared to pupils exposed to business-as-usual conditions?

Secondary research questions (end of Year 5)

2. What is the difference in **writing self-efficacy** between pupils exposed to a teacher trained in the Power of Reading compared to pupils exposed to business-as-usual conditions?

3. What is the difference in **writing enjoyment** between pupils exposed to a teacher trained in the Power of Reading compared to those pupils exposed to business-as-usual conditions?

4. What is the difference in **reading engagement** between pupils exposed to a teacher trained in the Power of Reading compared to those pupils exposed to business-as-usual conditions?

Secondary research question (NPD, end of Year 6)

5. What is the difference in **reading attainment** between pupils exposed to a teacher trained in the Power of Reading at the end of Year 6 (2026/27) compared to those pupils exposed to business-as-usual conditions?

Subgroup effects will be calculated for those pupils ever-FSM and pupils with SEND as well as for gender.

Design

Table 1 Trial design

Trial design, including number of arms		Two-arm, split cohort, cluster randomised trial
Unit of randomisation		School
Stratification variables (if applicable)		Training sites
Primary outcome	**Variable**	Writing attainment

	Measure (instrument, scale, source)	KS2 NCT pre-2013 tests, 0-40, two bespoke writing tasks
Secondary outcome(s)	**Variable(s)**	Writing self-efficacy Writing enjoyment Reading engagement Reading attainment
	Measure(s) (instrument, scale, source)	Experiences of Reading and Writing questionnaire (ERW), Writing Self-Efficacy Measure, 16-80 ERW, Liking Writing Scale, 4-20 ERW, Reading Engagement Scale 20-96 KS2 NCT reading test scaled score, 80-120, NPD
Baseline for primary outcome	**Variable**	Writing Attainment
	Measure (instrument, scale, source)	Writing Assessment Measure (WAM), 0-32
Baseline for secondary outcome	**Variable**	Writing self-efficacy Writing enjoyment Reading engagement Reading attainment
	Measure (instrument, scale, source)	ERW, Writing Self-Efficacy Measure, 16-80 ERW, Liking Writing Scale, 4-20 ERW, Reading Engagement Scale 20-96

The impact evaluation follows a parallel, split cohort, cluster-randomised controlled trial design to assess efficacy of the intervention in improving educational outcomes, specifically in writing and reading. Schools will be randomly assigned to intervention or control groups on a 1:1 basis. Teachers of a single randomly selected Year 5 class in intervention schools will be invited to attend PoR training, while those in control schools will not have access to the programme. Control schools will receive a payment of £500 per school if all required pupil data are provided and if the school facilitate administration of the required evaluation assessments (writing tasks and online questionnaires at the beginning and end of the evaluation). They will also benefit from 20% discounted PoR membership and/or training at the end of the evaluation.

The primary outcome will be writing attainment, measured using two bespoke writing tasks from KS2 pre-2013 tests after programme exposure in both groups, and for both cohorts. Secondary outcomes will assess writing self-efficacy, writing enjoyment, and reading engagement using the Experiences of Reading and Writing questionnaire (ERW), measured at the end of Year 5 for both cohorts. Furthermore, effect on reading attainment will be assessed at the end of Year 6 through the KS2 NCT reading test scaled score accessed via the National Pupil Database (NPD) again for both cohorts.

Participant selection

We aim to recruit a total of 140 primary schools for the trial, with half assigned to the intervention group. The trial will focus on a single Year 5 teacher and one randomly selected class per school, identified prior to randomisation. Schools will be recruited in two cohorts as part of a split design: 69 will participate in the 2025/26 academic year, and the remainder in 2026/27. In total, 70

teachers are expected to complete the PoR training across the two cohorts. Cohort 1 will be exposed to the programme during the 2025/26 academic year and assessed at the end of 2026, while Cohort 2 will be exposed during 2026/27 and assessed at the end of 2027.

Assuming no attrition, we anticipate a total sample of approximately 3,920 pupils (28 pupil per participating schools), with half taught by a teacher who has received PoR training in their respective intervention year. As PoR is a teacher-focused CPD programme, all pupils taught by the trained teacher are expected to benefit. Therefore, there are no pupil-level exclusion criteria. All pupils entering Year 5 in the selected class - beginning in September of their cohort year - will be included in the study sample, subject to opt-out parental consent.

Schools

School eligibility criteria have both 'hard' and 'soft' elements. The hard criteria refer to those criteria that cannot be comprised, whereas soft criteria are those that can be comprised in circumstances where there are difficulties meeting the school recruitment target.

To be eligible to participate in the trial schools must be:

- A state-maintained primary, junior, middle, or all-through schools in England.
- Not already involved in another EEF-funded trial involving Year 4 pupils in the academic year prior to their PoR participation, or Year 5 pupils during their year of involvement in PoR (i.e., 2024/25 and 2025/26 for Cohort 1; 2025/26 and 2026/27 for Cohort 2).
- Able to confirm that their Year 5 teachers in the upcoming academic year have not participated in Power of Reading or Power of Pictures training within the last five years.
- Able to provide necessary IT connectivity to enable pupils to participate in online questionnaires
- Schools that 'set' or 'stream' Year 5 pupils in literacy teaching will be excluded
- Able to release the chosen Year 5 teacher for four days of training spread across the year, with a maximum of two teachers participating in the training per school

Around 40-50% of recruited schools will come from Education Investment Areas. In addition, schools are given preference for inclusion in the trial where they do not have mixed-age group classes for literacy instruction. If schools that do have mixed-age group classes remain interested in participating, they will be added to a recruitment waitlist. Also, there is a preference for schools not to be participating in any other EEF trials in the year of intervention delivery. This is due to concerns that this may affect their ability to engage with the data collection requirements of the project (e.g. providing enumeration data etc). If schools are participating in another EEF trial and are interested in this project, they will also be added to the recruitment waitlist.

Pupils

All pupils in the selected class who are enrolled at the time of school recruitment and are due to enter Year 5 in the September following recruitment (i.e., September 2025 for Cohort 1 and September 2026 for Cohort 2) will be included in the study. Parents may withdraw their child from the study at any time. Pupils who withdraw or join after the baseline assessment will not be included in the evaluation. Those entering intervention schools later will take part in PoR but will not complete endline assessments.

Recruitment

Recruitment will take place in two phases, aligned with the split cohort design. The first phase is expected to run between February and June 2025 for schools participating in the 2025/26 academic year (Cohort 1), with a second recruitment phase taking place the following year for Cohort 2. The school recruitment process will follow these steps: upon initial signup, CLPE will identify and contact schools that meet the selection criteria. Initial information will be collected, including school name, address, telephone number, URN, and the names and contact details of anticipated Year 5 teachers (for the intervention year—i.e., 2025/26 for Cohort 1 and 2026/27 for Cohort 2). Schools will then be invited to sign an MoU, which outlines the project's purpose, anticipated benefits, key activities and timelines, data protection provisions, and responsibilities of all parties. A school will be considered 'recruited' once the MoU is signed, at which point CLPE will share the school's information with FFT and MMU.

In addition, schools must sign a separate Data Sharing Agreement specifying how personal data will be collected and shared between CLPE, the Evaluation Team, and the school. Once MMU receives the finalised school details, they will select a single class in multi-form entry schools and share the selected class name and Year 5 teacher name with FFT and CLPE, who will then relay this information back to the schools. Following this, schools will need to circulate withdrawal letters and information sheets to the parents of pupils that will enter Year 5. Pupils are given an initial window of two weeks to withdraw from the study. During this period and ongoing for these pupils, it is the responsibility of the school to maintain a record of pupils that have withdrawn and ensure that they do not release any personal information regarding these pupils to FFT, MMU or AlphaPlus. If a pupil withdraws subsequent to the two-week period, then the school will need to notify MMU using a secure link to ensure that any existing data held from pupils that wish to withdraw is destroyed. Once the two-week opt-out period has ended, FFT will collect pupil-level data for all pupils, excluding those pupils who have opted out. These records will then be verified for completeness before being transferred to MMU.

Outcome measures

The following section sets out the primary and secondary outcome measures that will be collected at endline, as well as the corresponding baseline measures that will be used as covariates in the analysis.

Primary outcome

We propose a bespoke measure of writing attainment as the primary outcome, derived from scores on two writing tasks from previous KS2 National Curriculum Tests (pre-2013) – one longer narrative (45 minutes) and one shorter non-fiction task (20 minutes). For both tasks pupils are asked to plan and create a piece of independent writing within the time given, following a standardised prompt. These tasks are appropriate for upper KS2 and have high internal consistency ($\alpha > 0.90$ for the 2010 and 2011 KS2 writing assessments; Ofqual, 2014) with clear marking criteria. For studies using bespoke tests derived from pre-2013 KS2 writing assessments, relatively low intraclass correlation coefficients (ICCs) were encountered. The Integrating English trial found an ICC of 0.18 (Culliney, et al. 2019). The Grammar for Writing study, also using a bespoke test from past KS2 writing tests obtained an ICC estimate of 0.12 (Tracey, et al. 2019).

The longer task is scored on a scale from 0 to 28 depending on how well the pupil has performed in relation to the following three strands: sentence structure and punctuation (0-8 marks); text structure and organisation (0-8 marks); and composition and effect (0-12 marks)[4]. A score of 28 means all the criteria for each of these three areas have been met to a high standard. The marker arrives at a score for each of the strands by judging which descriptors in the mark scheme provide the best fit to the nature of the pupil's writing. For example, for the sentence structure and punctuation strand, there are descriptors provided in the mark scheme which provide a guide as to what is required for a child to score 1 mark for this strand (e.g. …Some simple sentences, often brief, starting with a subject + verb…) all the way up to 8 marks (e.g. Length and focus of sentences varied to express subtleties in meaning…). Then the scores for each strand are added together to give a total score out of 40 for the longer task.

The shorter task is scored on a scale of 0 to 12 against two strands: sentence structure, punctuation and text organisation (0-4 marks); and composition and effect (0-8 marks). Again, the marker makes a judgement about how many marks should be awarded for the two strands based on the descriptors provided in the mark scheme. The scores and then added together to give a mark out of 12 for the shorter task. Finally, the longer writing task and shorter writing task are added together to calculate the overall score of writing attainment with possible scores ranging from 0 to 40.

The writing tasks evaluate pupils' ability to 'write imaginative, interesting, and thoughtful texts' and 'produce texts which are appropriate to task, reader and purpose.' These areas of focus align with the logic model which predicts improvements in pupils' writing ideation and composition, writing at greater depth, and with an understanding of purpose and audience.

The delivery team would be kept blind to the exact nature of the tasks. The writing tasks would be administered by the teacher (whole class) in the presence of invigilators from AlphaPlus Consultancy (subcontracted by MMU). The tasks will be marked by teachers who have experience of teaching and assessing writing in Key Stage 2. They will have training on how to

[4] In original mark scheme from the KS2 SATS tests, there is a fourth strand, handwriting, for which 0-3 marks could be awarded. For the purposes of this trial, handwriting will not be assessed, as it is not a focus of PoR.

mark the tasks, drawing upon the exemplification materials (annotated examples of scored writing) that were published within the KS2 SATS mark schemes. The training will also provide time for teachers to moderate each other's judgements and arrive at consensus scores for a range of sample papers.

We considered and discounted a range of alternative assessments. For example, these included the Writing Assessment Measure (WAM) (Dunsmuir et al., 2015). This measure aligns well with the logic model; however, we have concerns over the high ICCs found in the PoP (0.31) (Anders et al., 2021a) and Craft of Writing (0.33) (Anders et al., 2021b) studies (although these high conditional ICCs may also be a function of the lack of a pre-test covariate in both Anders at al., 2021a and 2021b). High ICCs would imply the need for a larger sample of schools in order to maintain power at 80 per cent all else equal. Though note we propose this as a baseline measure of writing. The Test of Written Language 4 (TOWL-4) was also considered. However, this does not have UK norms and involves pupils recounting a given story, rather than generating one using their own ideas. Similarly, the writing measures within KTEA-3 and WIAT-III UK focus on technical writing skills rather than ideation. These measures were therefore considered not to align closely enough with the logic model. The Progress in English test used in the GfW efficacy trial (Torgerson et al., 2014); was also considered but is now out of print, replaced with the Progress Test in English, which does not include an extended writing task.

On balance, the pragmatic decision was made to use two past KS2 NCT writing tasks to create the primary outcome measure, as they align well with the theory of change and avoid potentially high ICCs.

Baseline measure of writing attainment

We will collect the Writing Assessment Measure (WAM) as a baseline measure of writing. It has high content validity, concurrent validity, internal consistency, test-retest reliability, and inter-rater reliability (Dunsmuir et al., 2015). It is suitable for pupils across KS2 and is aligned with the primary outcome. We propose WAM at baseline rather than the pre-2013 KS2 tasks because the latter may not be suitable for lower-KS2. A slight adaptation will be made to the standard procedure for the WAM (Dunsmuir et al., 2015). In line with the procedure adopted in the PoP evaluation (Anders et al., 2021a), pupils will be given five minutes to plan their writing before the 15-minute writing time begins. This decision is informed by observations from the PoP trial, where the authors noted that "15 minutes was insufficient for pupils to produce a writing sample that could be meaningfully assessed" (Anders et al., 2021a, p. 18). To address this, an additional five minutes of planning time will be provided to ensure pupils have adequate time to plan and produce a meaningful piece of writing.

Secondary outcomes

We propose using the Writing Self-Efficacy Measure (WSEM) (Anders et al., 2021a), to assess **writing self-efficacy** as a secondary outcome. This reflects the assumption within the logic model that PoR will support pupils to have a better understanding of authentic writing processes, which in turn will support them to feel more confident as writers. The WSEM was adapted from

the self-efficacy for writing scale (SEWS) created by Bruning et al. (2013). It was adapted by the evaluation team of the PoP trial to make it more accessible for primary school children (Anders et al., 2021a). The wording of some of the items was slightly simplified (based on consultation with primary literacy experts by the PoP evaluation team), and the rating scale was changed from 1 to 100 to 1 to 5. We will adopt the same procedure here on the basis that this adapted version of the scale was found to be sensitive to the impact of a similar intervention (PoP) in a sample of Year 5 pupils. Cronbach's alpha for the three subscales within the original SEWS ranged from 0.85-0.90 (Bruning et al., 2013). In the WSEM (the adapted version of the SEWS), Cronbach's alpha was estimated to be 0.9, suggesting that the adapted version maintains its internal consistency. The WSEM has 16 items which pupils rate according to a five-point Likert scale (I'm sure I can't do it to I'm very sure I can do it). The range of scores is 16 to 80.

We also suggest the inclusion of a secondary outcome measure of **writing enjoyment**, given that PoR is proposed to support pupils to become more enthusiastic about/engaged in writing. The Liking Writing Scale (LWS) as was selected as it is designed to probe this affective dimension. The LWS has four items relating to the extent to which pupils enjoy writing, which they rate according to a five-point scale. Bruning et al. (2013) found Cronbach's alpha for LWS to be 0.83. The range of scores is from 4 to 20.

It was proposed at set-up that a measure of **reading engagement** should also be included as a secondary outcome. The Reading Engagement Scale (RES) (McGeown & Smith, 2023) was selected as it aligns well with the logic model, which suggests that PoR encourages children to read more and to engage more deeply with texts. The RES measures four dimensions of pupils' engagement with reading: behavioural (e.g. How much time do you spend each week reading?); affective (e.g. I care about what happens to the characters in the books I like); cognitive (e.g. I think about what I already know to help me to understand what I read); social (e.g. I enjoy discussing books in my free time). Each of the four dimensions has six items, therefore scores for each dimension can range from 6 to 24, and the total scores can range from 24 to 96. Cronbach's alpha values for the four dimensions range from 0.61 to 0.78.

The RES would be combined with the WSEM and LWS within one Experiences of Reading and Writing (ERW) questionnaire. This will take approximately eight to ten minutes to administer. Scales within the ERW questionnaire will be administered online at both baseline and endline, with oversight from Year 5 teachers.

While the primary outcome of interest is writing, the logic model suggests that reading attainment may also be improved by PoR. The pedagogical approach which teachers are encouraged to adopt involves in depth engaging with high quality texts. This includes activities, which are designed to support reading comprehension, e.g. book talk, text marking, making predictions etc. We therefore suggest a secondary measure of **reading attainment** at endline: pupils' scaled scores from the KS2 NCT reading test, which will take place in May 2027 (cohort 1) and 2028 (cohort 2). This measure will be accessed from the National Pupil Database (NPD). Originally, we planned to use pupils' Key Stage 1 (KS1) reading teacher-assessed grades as the baseline covariate measure for reading attainment. However, due to a change in design - specifically, the shift to split cohorts with an extended trial period - Cohort 2 falls within the group for whom KS1

assessments were no longer mandatory. As a result, KS1 grades will not serve as a sufficient baseline measure. To address this, we propose using the Reading Engagement Scale (RES) as an alternative baseline predictor for reading. While the RES is very new, study by McGeown and Conradi-Smith (2023) found positive, albeit modest, correlations between RES reading engagement subscales and performance on the New Group Reading Test (r = 0.13. – 0.25). Similarly, there is evidence that WAM correlates with a range of reading skills, with coefficients ranging from r = .33 to .66, as reported by Booton et al. (2025). These findings are consistent with the broader literature demonstrating strong links between reading and writing performance (Koons, 2008; Pearson Clinical, no date). Taken together, this supports the use of both the RES and WAM as suitable proxy baseline measures given that KS1 reading data are unavailable (see Table 2).

Table 2 Primary and secondary outcome measures with relevant baseline covariate

Construct	Endline		Baseline	
	Measure	Assessment length (minutes)	Measure	Assessment length (minutes)
Writing (primary outcome)	**KS2 NCT pre-2013 tests**	65	**WAM** Writing Assessment Measure	15 (plus 5 minutes planning time)
Writing self-efficacy (secondary outcome)	**ERW** Writing Self-Efficacy Measure	5	**ERW** Writing Self-Efficacy Measure	5
Writing enjoyment (secondary outcome)	**ERW** Liking Writing Scale (motivation & engagement)	2	**ERW** Liking Writing Scale (motivation & engagement)	2
Reading engagement (secondary outcome)	**ERW** Reading Engagement Scale	8	**ERW** Reading Engagement Scale	8
Reading attainment (secondary outcome)	**KS2 Reading** Scaled scores	N/A	**WAM** Writing Assessment Measure **ERW** Reading Engagement Scale	15 (plus 5 minutes planning time) 8

Sample size

Effect sizes (standardised difference in means) from previous tests of writing interventions have varied quite markedly, from -0.02 to 0.12 (Culliney, et al. 2019; Anders et al. 2021a; Anders et al, 2021b; and Tracey et al, 2019). The question is, how far are the effect sizes we see in these studies useful for predicting the potential effect of PoR. PoR is a development of PoP for which an ES=0.09 was obtained. PoR is a more advanced and intensive version of PoP. Thus, we might expect PoR to achieve a higher effect size than PoP (Anders et al., 2021a). Slavin et al (2019) provide a synthesis of writing studies and finds effect sizes that are larger 0.17 to 0.19. On balance, we might therefore expect PoR to produce an effect size of around 0.15.

Table 3 below presents the minimum detectable effect sizes (MDES) for both the full sample and the FSM subgroup, under scenarios of no attrition and 10 percent school-, and pupil-level

attrition. The MDES values indicate the smallest effect sizes that, if true, would yield statistically significant results at the 95 percent confidence level.

Table 3 Sample size calculations

		Overall	FSM	Overall	FSM
Attrition rate		**No attrition**	**No attrition**	**10% attrition**	**10% attrition**
Minimum Detectable Effect Size (MDES)		**0.14**	**0.18**	**0.15**	**0.19**
Pre-test/ post-test correlations	level 1 (pupil)	0.7	0.7	0.7	0.7
	level 2 (class)	NA	NA	NA	NA
	level 3 (school)	0.7	0.7	0.7	0.7
Intracluster correlations (ICCs)	level 2 (class)	NA	NA	NA	NA
	level 3 (school)	0.15	0.15	0.15	0.15
Alpha		0.05	0.05	0.05	0.05
Power		0.8	0.8	0.8	0.8
One-sided or two-sided?		Two	Two	Two	Two
Average cluster size		28	7[5]	25	6
Number of schools	Intervention	70	70	63	63
	Control	70	70	63	63
	Total	140	140	126	126
Number of pupils	Intervention	1,960	490	1,575	378
	Control	1,960	490	1,575	378
	Total	3,920	980	3,150	756

The MDES values are calculated using the PowerUp software (Dong & Maynard, 2013) based on the following assumptions:

- Statistical tests of the null hypothesis will be conducted at a 95% significance level, with statistical power of 80%.
- Schools will be randomly assigned to either the intervention or control group in a 1:1 ratio.
- The total expected sample size at recruitment is 140 schools, pooling together both cohorts.

[5] According to the latest education statistics, 24.3% of primary school pupils were eligible for FSM in the 2023/24 academic year.

- Each participating school will contribute one Year 5 class to the evaluation, with an average class size of 28 pupils.
- Previous studies on writing outcomes with pre-test measures have reported pre/post-test correlations ranging from 0.60 to 0.70 (Tracey et al., 2019; Stone et al., 2022). For other literacy outcomes, such as reading attainment, correlations often exceed 0.70 (Lewin et al., 2004). Therefore, we anticipate a correlation of approximately 0.70 between the WAM and the bespoke test derived from KS2 papers.
- Drawing on evidence from other EEF trials using bespoke KS2 writing tests (Culliney et al., 2019; Tracey et al., 2019), we assume an intra-cluster correlation (ICC) of 0.15 at the school level.

Under these assumptions and with no attrition, the estimated MDES is 0.14 for the full sample and 0.18 for the Free School Meals (FSM) subgroup. Table 3 also presents an alternative scenario, which is considered the primary scenario, accounting for a 10% attrition rate at both the school and pupil levels. In this case, the MDES for the full sample is 0.15, while for the FSM subgroup, it is 0.19. As noted earlier, an MDES of 0.15 is deemed appropriate for this trial.

Randomisation

Randomisation will occur at the school level with a 1:1 allocation to intervention and control groups. It will be conducted separately for each cohort. After recruitment, one Year 5 class will be randomly selected by the evaluators in each school to take part, and all pupils in that class will be eligible for inclusion, subject to opt-out parental consent. Baseline assessments will be conducted at the start of Year 5.

For Cohort 1, randomisation is scheduled for October 2025; for Cohort 2, it will take place in October 2026. In each case, randomisation will occur after schools have confirmed the teacher assigned to the selected class and baseline data has been collected. This timing helps reduce the risk of schools adjusting teacher allocations based on trial group assignment.

The randomisation will be conducted in Stata by a member of the evaluation team who will remain ‘blind’ to school identifiers, with schools then informed of their allocation. Regions tied to training centres will serve as stratifiers to guarantee sufficient allocation of intervention schools per training site.

We will use the Stata command `randtreat` (Carril, 2017) to perform stratified randomisation and manage "misfits"—instances where the number of schools in a stratum is not divisible by the number of groups. This global approach ensures balance across the entire sample. To ensure transparency and prevent modifications, the randomisation code will be documented in a Stata do-file.

Statistical analysis

Primary analysis

The primary analysis will be performed on an intention to treat basis. Consistent with EEF's statistical guidelines, a sample estimate of the intervention effect will be obtained through a regression adjusted analysis, through fitting a mixed linear model of the following form:

$$Y_{ij} = \beta_0 + \beta_1 T_j + \beta_2 X_{ij} + \beta_3 S_j + \beta_4 C_j + u_j + \varepsilon_{ij} \quad \text{.... [1]}$$

Where:

- Y_{ij} is the writing attainment score for pupil i in school j.
- T_j is a binary variable coded to '1' if school j is assigned to the intervention '0' otherwise.
- X_{ij} is the baseline writing score obtained from the WAM for pupil i in school j (this covariate is entered as a pupil level covariate and will reduce variance explained at school and pupil levels).
- S_j Is a stratum indicator
- C_j is a binary cohort indicator denoting which cohort the school belongs to (e.g., 0 = Cohort 1; 1 = Cohort 2), .

The sample estimate of β_1 is the estimated treatment effect of Power of Reading. There are random effect at the school u_j as well as a pupil level residual ε_{ij}. The two random effects are assumed to be uncorrelated, normally distributed, with mean zero and conditional variances σ_ς^2 (between school variance)σ_ε^2 (with in school variance). $\sigma_\varsigma^2 + \sigma_\varepsilon^2$, and therefore, the school level $\frac{\sigma_\varsigma^2}{(\sigma_\varepsilon^2+\sigma_\varsigma^2)}\frac{\sigma_\varsigma^2}{(\sigma_\varepsilon^2+\sigma_\varsigma^2)}$.

Estimates of the model parameters will be obtained using restricted maximum likelihood and the command 'mixed' in STATA v18. The sample estimate of the intervention effect from the model above will be converted into an effect size. This will be achieved through first fitting a simple variance components model:

$$Y_{ij} = \beta_0 + \varsigma_j + \varepsilon_{ij} \text{....[2]}$$

The confidence interval for the ES will be calculated by obtaining the upper and lower limits of the 95% confidence interval from the regression output associated with the estimate of $\hat{\beta}_1$ after fitting model [1], dividing both limits by the denominator from the expression immediately above (pooled unconditional variance). The ES, regression coefficient, 95% confidence interval for ES and continuous p-value will be reported, along with estimates of the ICCs and ES denominator.

Secondary analysis

Secondary analysis will involve estimating the effects of Power of Reading on the secondary outcomes for writing self-efficacy, writing enjoyment, reading attainment, and reading engagement. Sample estimates will be obtained from estimating four regression models (one for each outcome) similar to that used in the primary analysis, except where y_{ij} is either the writing self-efficacy, writing enjoyment, reading attainment, or reading engagement score for pupil i in school j – depending on which of the four models is considered. In these models x_{ij} is the baseline measure (or a vector of baseline measures, in the case of reading attainment) for the given outcome measure as outlined in Table 2.

Estimates of the model parameters will be obtained using restricted maximum likelihood and the command 'mixed' in STATA v18. Once again, results will be reported as effect sizes with 95% confidence intervals.

Estimation of effect sizes

The intervention effect estimated from the primary outcome regression model will be converted into an effect size. This conversion involves fitting a basic variance components model of the outcome and two level of random effects (pupil and the school level):

$$y_{ij} = \beta_0 + u_j + \epsilon_{ij}$$

Where y_{ij} refers to the outcome for pupil i in school j and u_j and ,ϵ_{ij} correspond to school and pupil level errors, respectively.

From this model we will obtain sample estimates of the unconditional variances for , u_j and ,ϵ_{ij} namely,$\hat{\sigma}_u^2$, and .$\hat{\sigma}_\epsilon^2$. The sample estimate of the effect size is then calculated as:

$$ES = \frac{\hat{\beta}_1}{\sqrt{\hat{\sigma}_u^2 + \hat{\sigma}_\epsilon^2}}$$

Where $\hat{\beta}_1$ is the coefficient estimate on the binary treatment indicator from the primary outcome analysis regression model.

The confidence interval for the ES will be calculated by obtaining the upper and lower limits of the 95% confidence interval of $\hat{\beta}_1$ from the primary outcome analysis regression output, dividing both limits by the denominator from the expression immediately above. The ES, regression coefficient, 95% confidence interval for ES and continuous p-value will be reported, along with estimates of the ICCs and ES denominators.

The same approach will be used to compute effect sizes for all secondary outcomes.

Sub-group analyses

For three subgroups we will conduct separate analyses using the same model as our primary analysis with the addition of the relevant main and interaction effects. Subgroups of interest will be pupils entering Year 5 in participating classes that are ever-FSM, those designated SEND and by gender. Subgroup analyses are of particular interest because there is preliminary evidence suggesting that PoR is effective at narrowing the gender gap, and the achievement gap between advantaged and disadvantaged pupils (Doherty, 2019).

Following EEF statistical guidance, we will conduct subgroup analyses using two approaches: the restricted subsample approach and the interaction approach. For the restricted subsample approach, the primary analysis model will be applied separately to specific subsets of pupils categorized as 'male,' 'female,' EverFSM and SEND-designated. For the interaction approach, a similar model to that used in the primary analysis will be employed, incorporating both the binary subgroup indicator and an interaction term between the subgroup indicator and treatment allocation.

Results will be reported as effect sizes accompanied by 95% confidence intervals to provide a clear understanding of the subgroup-specific effects.

Analysis in the presence of non-compliance

Following extensive consultations with CLPE, it was agreed that the completion of the first Gap Task - aimed at assisting teachers in reviewing and refining their practices, as well as evaluating their impact on pupils and the learning culture - provides a sufficient foundation for the intervention to positively influence pupils. Evidence of the completed Gap Task is collected during the second Training Day. Consequently, it was determined that attendance at the first two Training Days would serve as the minimum compliance criterion. As a result, all focal pupils in the intervention arm taught by a teacher who participated in the second Training Day may be considered compliant.

There is also a small risk that schools allocated to the control group may not adhere to their allocation but still subscribe to PoR and send their focal teacher to training. To address this, an additional compliance criterion has been introduced to identify non-compliant control schools, ensuring they are appropriately classified.

In summary, two (school-level) compliance criteria are proposed:

1. Intervention schools are considered non-compliant if there is no evidence that the focal Year 5 teacher attended the first two Training Days.

2. Control schools are considered non-compliant if there is evidence that the focal Year 5 teacher attended the first two Training Days.

Evidence will be gathered through training logs provided by the Delivery Team.

A complier average causal effect (CACE) will be estimated by constructing a binary compliance variable at the pupil level (1 = complier, 0 = non-complier). The detailed approach to compliance analysis will be outlined in the Statistical Analysis Plan (SAP).

Additional analyses and robustness checks

In additional for fitting models [1] and [2] above, the primary analysis will also include the reporting of results from fitting the following additional models. For each of these models, writing attainment will form the response variable:

Simple model containing the intervention dummy variable only, with inference performed based on standard parametric assumptions. This model enables us to examine the effects of covariates on sample estimates. Covariates are known to bias estimates particularly in small sample with bias diminishing quickly as the sample size rises (Lin, 2014).

Model containing intervention dummy, cohort indicator and stratifiers with inference performed based on standard parametric assumptions. This model contains all 'design' related covariates so that when compared to the model at the first bullet [1] above, the influence of design features can be assessed.

Simple model containing the intervention dummy variable only, with inference performed using randomisation inference, obtaining exact p-values, under the sharp null hypothesis. This model enables us to relax parametric assumptions associated with inference in order to assess their reasonableness. This approach assesses whether the result from [1] could have arisen by chance, without relying on standard statistical assumptions. It is important to note that randomisation inference proceeds under a different definition of the null hypothesis referred to as the sharp null (Imbens and Rubin, 2015). For inference performed using randomisation tests we will use the command ritest in STATA v18 (Heß, 2017).

Finally, a model including an interaction between the cohort and intervention variables to explore whether the effect of PoR differs between the two cohorts. This will help determine if the impact of the intervention varies depending on the academic year in which the schools participate.

Missing data analysis

For the primary analysis, sensitivity tests will assess whether missing follow-up data biases or imprecise estimates of the average treatment effect. Pre-randomisation missingness is unlikely to bias treatment effects but may reduce sample sizes. Post-randomisation missingness may arise from:

- Parental requests to remove children from the study
- Pupils leaving school before KS2 NCT writing tests
- Schools disengaging from endline data collection or withdrawing from the evaluation
- Pupil absence during KS2 NCT writing tests

In the first stage, we will classify missingness as Missing completely at random (MCAR), Missing at Random (MAR), or Missing not at random (MNAR) by comparing missing data rates across trial arms. If missingness exceeds 5% in both groups, we will use a multilevel logistic regression model (the "dropout model") to identify additional covariates that might be predictive of missingness. Significant covariates (95% confidence interval) will be considered explanatory for missingness.

If missingness exceeds 5% in any trial arm and covariates explain missingness, sensitivity tests will use Multiple Imputation by Chained Equations (MICE) to assess the impact of missingness. If missingness is suspected to be MNAR, sensitivity analyses using Lee bounds (Lee, 2009) will estimate treatment effect bounds under minimal assumptions (randomisation and monotonicity). Lee bounds provide upper and lower estimates for the always-responding subsample, addressing non-random attrition. These will be estimated using the `leebounds` package in STATA v18.

Detailed methods for missing data analysis, including MI and Lee bounds, will be specified in the SAP.

Implementation and process evaluation (IPE) design

The implementation and process evaluation (IPE) aims to explore the delivery of the programme and any impact of PoR on pupil outcomes (or not) within the IE. The IPE is grounded in the programme's ToC (see Figure 1). Quantitative and qualitative data will be gathered through pre- and post-intervention surveys of headteachers and teachers, together with case studies in schools.

The split cohort design provides an opportunity to explore:

- the short-term sustainability of the intervention
- whether and how schools are able to embed PoR within the school's culture
- adaptations made to the intervention.

Therefore, we will conduct case study visits at six schools during Cohort 1, with two schools per region, each visited on two separate occasions. In Cohort 2, we will visit an additional six schools, with two visits at each school. Additionally, we will revisit each of the Cohort 1 school at the end of the project. This makes a total of 30 case study visits across the study situated in 12 case study schools. The IPE will also review project documentation, as well as data from the reflection evaluation forms completed during the training days.

Research questions

The proposed research questions the IPE aims to address are as follows with associated IPE dimensions provided in brackets (EEF, 2022):

RQ1. What reading practices/programmes/interventions are:

- Happening in control schools?
- Happening alongside PoR in intervention schools?
- Replaced by PoR within intervention schools? (programme differentiation, monitoring of control groups)

RQ2. To what extent was the intervention delivered as intended in relation to:

- How often teachers delivered lessons using PoR?
- Embedding key pedagogical principles of PoR within lessons?
- Use of online PoR resources?
- Engaging with gap tasks? (fidelity, dosage, responsiveness)

RQ3. What adjustments (if any) were made to PoR?

- Why were these adjustments made? (adaptation)

RQ4. What are the barriers and facilitators to implementation?

- Which barriers and facilitators, if any, were specifically or disproportionately experienced by FSM, SEND, EAL pupils and pupils of different genders? (context, moderators)

RQ5. Did the support available within PoR enable practitioners to deliver it effectively? Including:

- Training days
- Unit plans
- Online resources
- Gap tasks associated with the training (quality, context/moderators)

RQ6. What impacts (if any) do headteachers, teachers and pupils perceive on pupils':

- Reading?
- Speaking and listening?
- Understanding of text structure and organisation?
- Grammar and language choices?
- Self-efficacy and enjoyment of writing?
- Volition and ability to develop ideas, draft, create and reflect on writing?
- Stamina and quality?
- Overall writing ability? (perceived impact, mediators)

RQ7. What impacts (if any) do headteachers and teachers perceive PoR to have on:

- Workload?
- Teachers' understanding of text-based approaches to reading fluency, comprehension, and writing?
- Teachers' understanding of authentic writing and how to model this?

- Teachers' understanding of creative approaches that support writing.
- Teachers' understanding of how the teaching of spelling and grammar can be contextualised within a text-based approach.(perceived impact, mediators)

RQ8. What evidence is there to support the assumption that gains in writing are mediated by gains in reading comprehension? (mediators)

Research methods

The main methods of data collection will include surveys, interviews, observations, and the review of relevant documents and resources. Question design will be undertaken in consultation with the Delivery Team and informed by in-house literacy specialists. Surveys will combine both closed and open-ended questions.

As the trial is a two-year split cohort design data collection will happen in both years. We will follow the same data collection plan for both years (except for an additional case study visit to the Cohort 1 schools to explore the extent to which PoR pedagogy became embedded in the school and any perceived impacts after the Cohort 1 intervention period). Preliminary analysis of the Cohort 1 IPE data will allow us to adapt the methods/research instruments as appropriate. Decisions about whether changes are implemented will balance the benefits of a two-cohort design allowing unexpected lines of inquiry to be followed within the second cohort, with the affordances of keeping the core methods the same to allow comparison and aggregations across the two cohorts.

Pre-delivery

Prior to the start of the intervention, all headteachers will be asked to complete an online survey (no longer than 15-20 minutes). The survey will elicit data about:

- reasons for signing up to the trial;
- commitment to the trial;
- current literacy practices including other literacy interventions delivered in the school
- planned changes to the way in which reading and writing is taught for the relevant academic years (in addition to PoR if allocated to the intervention arm);
- future staff CPD plans relating to teaching reading and writing (in addition to PoR if allocated to the intervention arm);
- the school's culture around reading and writing

The headteachers survey will be administered in September 2025 and September 2026 to headteachers in both control and intervention schools. All teachers (in both control and intervention schools) who will be teaching the selected Year 5 classes (or classes including Year 5 pupils) in September 2025 and September 2026 will be asked to complete a teacher survey in September (no longer than 15-20 minutes). The teacher survey will elicit data about current literacy practices and knowledge including:

- current literacy practices including other reading and writing interventions delivered in the school; estimated pupil time spent on reading and writing instruction and reading and writing practice each week;
- formal and informal CPD relating to teaching reading and writing;
- knowledge of the process of how exposure to high quality literature improves pupil writing.

The teacher and headteacher surveys will allow us to define business-as-usual. The post intervention survey (see below) will then assess the extent to which business-was usual was maintained in control schools.

During delivery

Fidelity is a key implementation dimension for this evaluation.To monitor this, two instances of each of the four training days will be observed, with structured field notes serving as the primary data source. All intervention documentation and support materials provided to teachers will be reviewed, and attendance registers will be analysed. In addition, the Delivery Team will administer short surveys and activities related to training and support to help shape their provision. These data will be shared with the evaluation team, offering insight into teachers' perceptions of the training and support received.

In addition, fidelity will be evaluated through one 60-minute structured observation that will be undertaken of teachers in each of the case study schools: six in Cohort 1, and an additional six in Cohort 2. In addition, in year 2 of the evaluation, we will observe the teachers in the six Cohort 1 schools in order to investigate if and how teachers are continuing to implement PoR. Continuation of PoR beyond the trial period will be optional, so it will be of interest to explore how many of the Cohort 1 intervention schools choose to continue/discontinue the programme and why.

Six schools will be selected as case studies to achieve a balance in terms of sampling a sufficient range of contexts while keeping data collection feasible. Case studies will be selected at random, stratified by: school size (one-form per year group versus two or more forms per year group) and the percentage of pupils eligible for free school meals (divided into terciles). This will ensure representation across diverse contexts. Since engagement as a case study school is more involved than as a non-case study school, it is possible that some schools who are invited to be a case study may choose not to take part in this aspect of the trial. If this is the case, schools will be replaced until a sample size of six is met.

Two half-day visits to each case study school will be undertaken in Cohort 1 and then again in Cohort 2 case study schools. The first visit will be mid-way through the intervention and the second will be towards the end of the intervention. During each visit, the following activities will be undertaken:

- 60-minute classroom observations
- 30-minute teacher interviews (year 5 teachers)
- 20-minute headteacher interviews

- 20-minute pupilpupil group interviews (six to eight pupilspupil per school)

The planned timeline for case study visits is as follows.

Table 4 Overview of case study visits

Visit	Date	No. of Schools
Cohort 1: Case Study Visit 1	Jan/Feb 2026	6
Cohort 1: Case Study Visit 2	April/May 2026	6
Cohort 2: Case Study Visit 1	Jan/Feb 2027	6
Cohort 2: Case Study Visit 2	April/May 2027	6
Cohort 1: Follow-up visit	June/July 2027	6

During the second visit to Cohort 1 schools, we will carry out a combination of classroom observations and interviews with staff and pupils. To minimise additional burden on teachers, the scope of data collection will be adjusted according to each school's context—for example, whether PoR is viewed as embedded in school culture and practice, or whether literacy approaches have remained largely unchanged.

The headteacher will be sent interview questions in advance and advised to invite a senior manager with responsibility for literacy in KS2 to join the interview if deemed necessary. That is, if the headteacher does not feel able to respond fully to the questions, they can bring in someone with more direct involvement who may be able to provide more detail. Teachers will be asked to identify a balanced group of pupils based on gender, FSM, SEND and EAL (to the extent that this is possible depending on the profile of the class and parents and pupils providing consent/assent).

Headteacher/senior leader interviews will focus on:

- how PoR was integrated with current practices;
- perceived impact on:
  - teaching reading and writing;
  - reading and writing culture in the school;
  - staff development and knowledge of teaching reading, and the link between high quality texts and improved pupil writing;
  - workload
- how they themselves (or other senior staff) have supported the intervention;
- perceived enablers and barriers, as well as contextual factors influencing the delivery of the intervention (both positive and negative);
- future plans.

Teacher interviews at the first visit will focus on:

- experiences of the training and support including the resources provided;

- fidelity and dosage;
- perceived enablers and barriers;
- how PoR was integrated with current practices (i.e. what did it replace);
- perceived usefulness of elements of PoR;
- required adaptations and identified gaps (what else did they cover outside PoR).

Teacher interviews at the second visit will focus on:

- any updates to questions asked at the first interview;
- perceived impact on:
  - teaching reading and writing;
  - reading and writing culture in the school;
  - staff development and knowledge of the link between reading high quality texts and improved pupil writing;
  - pupils' writing self-efficacy; writing enjoyment; reading engagement; and reading attainment (reading; speaking and listening, understanding of text structure and organisation; grammar and language choices; self-efficacy and enjoyment of writing; volition and ability to develop ideas, draft, create and reflect on writing; stamina and quality).
- planned future use;
- recommendations for other teachers.

Pupil focus groups will focus on:

- their experiences of PoR, what they like and do not like about it;
- how it compares to other ways of learning to write that they have experienced;
- perceived impact on:
  - their reading ability;
  - their writing ability;
  - attitudes to writing;
  - school reading culture
  - any other impacts the pupils associate with PoR.

Post-delivery

Post-intervention surveys will be administered to headteachers and teachers in both Cohort 1 and Cohort 2 in June/July 2026 and June/July 2027.

Headteachers at intervention schools will be asked to complete an online survey (no longer than 20-30 minutes). This will elicit data about the impact of PALSUK on the school and the costs involved:

- commitment to the trial;
- current literacy practices including other reading interventions delivered in the school;
- aside from PoR what changes in literacy practices took place in 2025/26 (Cohort 1) and 2026/26 (Cohort 2), both planned and unplanned;
- CPD offered in relation to teaching reading (in addition to PALS-UK);

- perceived impact on:
  - the school reading culture;
  - staff knowledge of the link between exposure to high quality texts and writing attainment;
  - pupils' writing attainment
  - pupils' reading attainment;
  - future literacy practices;
- costs associated with delivering PoR.

At the end of the intervention, headteachers at control schools will be asked to complete an online survey (no longer than 15-20 minutes). The purpose of this survey is to establish whether or not business as usual has been maintained in relation to teaching reading. This survey will elicit data about:

- current literacy practices including other reading interventions delivered in the school;
- what changes in literacy practices took place in 2025/26 and 2026/27 (planned and unplanned);
- CPD offered in relation to teaching reading;
- future plans teaching reading.

All Year 5 teachers (or those teaching classes including Year 5 pupils) in intervention schools will be asked to complete a survey after the intervention has been delivered (no longer than 20-30 minutes). The survey is designed to gather teachers' perceptions about the delivery of PoR. The teacher survey will elicit data about:

- current literacy practices including other reading interventions delivered in the school; estimated pupil time spent on PoR approaches;
- experiences of PoR training;
- formal and informal CPD relating to teaching reading undertaken;
- dosage;
- adaptations made;
- identified gaps in the programme and how they were addressed;
- perceived enablers, challenges, and benefits of PoR and specifically SLT support;
- impact of PoR on:
  - workload;
  - understanding of text-based and creative approaches to text-based approaches, to reading fluency, comprehension, and writing
  - Impact on pupils' writing self-efficacy; writing enjoyment; reading engagement; and reading attainment.

All Year 5 teachers (or those teaching classes including Year 5 pupils) in control schools will be asked to complete a survey at the end of the intervention period (no longer than 15-20 minutes). The purpose of this survey is to establish whether or not business as usual has been maintained in relation to teaching reading. This survey will elicit data about:

- current literacy practices including other reading interventions delivered in the school (to be compared with responses to baseline survey);
- the extent to which current literacy practices have developed over the year (i.e. has business as usual changed in any way);
- estimated pupil time spent on reading and writing instruction and practice each week;
- formal and informal CPD relating to teaching reading undertaken.

Survey completion rates will be maximised using the following strategies:

- A £10 voucher provided to each participant who completes both the baseline and endline survey;
- Two reminder emails to be sent following the initial invite;
- Careful consideration of the length of each survey to balance the minimisation of burden with comprehensiveness of the questions in order to address the research questions.

The delivery team will also be interviewed to elicit their perceptions on the delivery of PoR, the training and support offered, any adaptions made, barriers and challenges and how they were addressed, and the impact of PoR on schools, teachers, and pupils. These interviews will take place towards the end of both the Cohort 1 and Cohort 2 intervention periods to capture any differences in terms of the delivery between the two cohorts.

Implementation dimensions

Compliance is addressed and defined under the impact evaluation described above.

Monitoring of control groups (RQ1) Verification of "business as usual" will be achieved through headteacher and teacher surveys administered at baseline (September 2022) across all schools, and again at post-test for control schools. Post-test surveys will gather information about current practices, enabling comparison with baseline data, as well as any specific changes that occurred during the academic year. This provides a means of triangulation.

Programme differentiation (RQ1) The IPE will explore what sets PoR CPD apart from what teachers have had in the past (through the surveys and case study interviews). More broadly the IPE will also explore the differences between the translation of PoR into classroom practice and business-as-usual (again, through the surveys and case study interviews).

Perceived impact of PoR will be assessed through the intervention headteacher and teacher surveys and the case study interviews (staff and pupils).

Fidelity (RQ2) will be assessed through multiple data sources: records of training and support delivery (including independent observations of training events), the PALS-UK observation checklist used during four structured classroom observations (covering classroom set-up and pupil activities), and teacher reports of adaptations to the programme gathered in post-test surveys. Case study interviews will further contribute to fidelity data, given its importance to successful implementation.

Dosage (RQ2) understood as a key aspect of fidelity, will be captured through teacher self-reports in post-test surveys and via light-touch weekly logs completed by intervention teachers.

Responsiveness (RQ2) (the extent to which participants engage with the intervention) will also be assessed. This will be evaluated through teacher surveys, case studies and lesson observation, while pupil responsiveness will be examined through lesson observations and pupil focus groups.

Adaptation (RQ3) We will explore the extent to which teachers adapted the PoR approach to their specific contexts and pupils through the surveys, interviews and classroom observations, as well as by observing the discussions during the training days and reviewing teachers' reflective tasks associated with the training.

We will explore any enabling/constraining factors associated with implementation **context (RQ4, 5)** through the surveys, case studies and discussions at the training days.

Possible **mediators (RQ6, 7, 8)** of any impacts will be explored through the survey responses and case study/delivery teams interviews.

Perceived impact (RQ6, 7) from a range of stakeholders will be investigated through the case study interviews and teacher/headteacher survey responses.

The **quality (RQ5)** with which the intervention is delivered will be examined through the classroom observations, interviews, and survey responses, as well as through review of teachers' responses in the training tasks.

Potential bias will be mitigated in several ways. Respondent bias will be minimised by gathering data from multiple stakeholders (e.g. headteachers, teachers, pupils) and through multiple methods (surveys, interviews, observations). Rigour will be ensured through the use of structured observation tools, standardised interview schedules, and carefully designed survey questions. Surveys will be distributed online, with at least two reminders issued to reduce non-response bias and the offer of incentives (see above). In a previous trial using a similar approach (but without the incentives) we achieved a response rate of 68% for the staff surveys (Morris et al, in prep), which we hope to increase further with the addition of incentives. In the event that more schools volunteer for case study participation than are required, selection will follow pre-specified criteria to ensure a representative spread of school contexts. Similarly, pupils for focus groups will be selected by teachers, who will be asked to identify a sample reflecting diversity in gender, ability, and other relevant criteria. Qualitative data will be coded and analysed thematically using a consistent framework applied by all researchers. A reflexive approach to analysis (Braun & Clarke, 2022) will be used to minimise researcher bias.

Analysis

Quantitative data will be analysed using descriptive statistics and mediation analysis. Quantitative analysis will investigate the potential mediation effect of reading comprehension,

on writing attainment. This analysis will help us understand mechanisms of change (RQ8) and will support explanations for the findings of the IE.

Qualitative data from open survey responses, observations, interviews, and focus groups will be analysed in NVivo using thematic analysis (Braun & Clarke, 2022) with a mixed coding approach. Interview recordings will be transcribed by a professional transcription service. Full verbatim transcripts (including pauses and stutters) are not required for this study and will not be used, given the additional cost.

A coding framework, informed by the logic model (e.g. SLT support, teacher engagement), will be applied deductively, while inductive coding will capture additional or unexpected themes. This dual approach allows for the identification of unanticipated mediators and outcomes. Following Braun and Clarke's (2022) guidance on reflexive thematic analysis, the process will involve familiarisation with the data, systematic coding, and iterative development and refinement of themes, with steps repeated as needed.

The selection of codes and themes will address the IPE research questions by identifying patterns in participants' experiences and perceptions of the intervention (Kiger & Varpio, 2020). The analysis will be primarily deductive, rooted in the logic model. For example, the dataset will be interrogated in relation to participants' views on the role of SLT support (or lack thereof) in the delivery of PoR, exploring how this influences practice and highlighting potential challenges to avoid. By adhering to Braun and Clarke's (2022) framework, the analysis will move beyond description to provide a nuanced interpretation that situates the significance of the findings in relation to the IPE research questions.

These data will allow the research questions to be addressed by capturing the experiences of both intervention and control schools (business as usual). Findings from the IPE will support the interpretation of the impact analysis and create opportunities to generate further hypotheses about potential mediators and sources of variation in treatment effects. Particular attention will be given to elements of the logic model not directly tested in the impact analyses, including SLT support, teacher engagement, pupil attitudes and confidence, and teachers' knowledge of how reading high quality . Both quantitative and qualitative evidence from the IPE will be drawn on to explore these aspects.

The use of structured approaches, including statistical and thematic analysis, will ensure rigour in the analyses. Quality assurance procedures will be implemented to safeguard the reliability of IPE findings. These will include training for junior team members where needed, clear documentation of analytical procedures, systematic record-keeping, and regular meetings of the analysis team to identify and resolve issues promptly.

Table 5 IPE methods overview

IPE dimension	RQs	Research methods	Data collection methods	Sample size and sampling criteria	Data analysis methods
Monitoring of control/ comparison group	RQ1	**Surveys (pre/post)**	Online questionnaires	140 headteachers and 140 teachers	Descriptive statistics; thematic analysis: deductive coding
Programme differentiation	RQ1	**Surveys (pre/post)**	Online questionnaires	140 headteachers and 140 teachers	Descriptive statistics; thematic analysis: deductive coding
		Interviews	Structured interviews	140 headteachers (or other senior leaders) and 140 teachers (and up to 140 teaching assistants)	Thematic analysis: deductive and inductive coding
Fidelity	RQ2, 3	**Surveys (post)**	Online questionnaires	70 headteachers and 70 teachers – intervention	Descriptive statistics; thematic analysis: deductive coding
		Interviews	Structured interviews	70 headteachers (or other senior leaders) and 70 teachers (and up to 70 teaching assistants) - intervention	Thematic analysis: deductive and inductive coding
		Observation	Structured observations	Classroom observation in 12 case study schools Observation of training - 4 observations of each of the	Frequency counts; deductive coding

				4 training days (2 in Cohort 1 and 2 in Cohort 2)	
		Document analysis	Structured document analysis	Analysis of delivery documentation and training tasks/reflection activities (while observing the training days	Frequency counts; deductive coding
Dosage		Surveys (post)	Online questionnaires	70 headteachers and 70 teachers – intervention	Descriptive statistics; thematic analysis: deductive coding
		Document analysis	Logs of how many PoR lessons taught – collected by delivery team	70 headteachers and 70 teachers – intervention	Frequency counts
Responsiveness	RQ2, 6	Surveys (post)	Online questionnaires	70 headteachers and 70 teachers – intervention	Descriptive statistics; thematic analysis: deductive coding
		Observation	Structured observations	Classroom observation in 12 case study schools Observation of training - 4 observations of each of the 4 training days (2 in Cohort 1 and 2 in Cohort 2)	Frequency counts; deductive coding
		Document analysis	Logs of how many PoR lessons taught – collected by delivery team	70 headteachers and 70 teachers – intervention	Frequency counts
Adaptation	RQ3	Observation	Structured observations	Classroom observation in	Frequency counts; deductive coding

				12 case study schools Observation of training - 4 observations of each of the 4 training days (2 in Cohort 1 and 2 in Cohort 2)	
		Surveys (post)	Online questionnaires	70 headteachers and 70 teachers – intervention	Descriptive statistics; thematic analysis: deductive coding
		Interviews	Structured interviews	70 headteachers (or other senior leaders) and 70 teachers (and up to 70 teaching assistants) - intervention	Thematic analysis: deductive and inductive coding
Context/ moderators	RQ4, 5	Observation	Structured observations	Classroom observation in 12 case study schools Observation of training - 4 observations of each of the 4 training days (2 in Cohort 1 and 2 in Cohort 2)	Frequency counts; deductive coding
		Surveys (post)	Online questionnaires	70 headteachers and 70 teachers – intervention	Descriptive statistics; thematic analysis: deductive coding
		Interviews	Structured interviews	70 headteachers (or other senior leaders) and 70 teachers (and up to 70	Thematic analysis: deductive and inductive coding

				teaching assistants); around 72 pupils - intervention	
Mediators	RQ4, 7, 8	Observation	Structured observations	Classroom observation in 12 case study schools Observation of training - 4 observations of each of the 4 training days (2 in Cohort 1 and 2 in Cohort 2)	Frequency counts; deductive coding
		Surveys (post)	Online questionnaires	70 headteachers and 70 teachers – intervention	Descriptive statistics; thematic analysis: deductive coding
		Interviews	Structured interviews	70 headteachers (or other senior leaders) and 70 teachers (and up to 70 teaching assistants); around 72 pupils - intervention	Thematic analysis: deductive and inductive coding
Quality	RQ5	Observation	Structured observations	Classroom observation in 12 case study schools Observation of training - 4 observations of each of the 4 training days (2 in Cohort 1 and 2 in Cohort 2)	Frequency counts; deductive coding
		Surveys (post)	Online questionnaires	70 headteachers and 70	Descriptive statistics; thematic

				teachers – intervention	analysis: deductive coding
		Interviews	Structured interviews	70 headteachers (or other senior leaders) and 70 teachers (and up to 70 teaching assistants); around 72 pupils - intervention	Thematic analysis: deductive and inductive coding
Perceived impact	RQ 6,7,8	Surveys (post)	Online questionnaires	70 headteachers and 70 teachers – intervention	Descriptive statistics; thematic analysis: deductive coding
		Interviews	Structured interviews	70 headteachers (or other senior leaders) and 70 teachers (and up to 70 teaching assistants); around 72 pupils - intervention	Thematic analysis: deductive and inductive coding
		Document analysis	Structured document analysis	Analysis of delivery documentation and training tasks/reflection activities (while observing the training days	Frequency counts; deductive coding

Cost evaluation design

Cost evaluation will determine the expenses associated with delivering the intervention during the trial. The research questions derived from this objective are:

1. What are the estimated delivery costs of the Power of Reading intervention per school?
2. What are the estimated delivery costs of the Power of Reading intervention per pupil?
3. What would be the estimated cost per school and per pupil of implementing Power of Reading over three years?

Consequently, the cost evaluation takes the form of a Cost Feasibility analysis, serving as a guide to the affordability of PoR, rather than a comparison between PoR and an alternative intervention. In accordance with the EEF's cost evaluation guidance (EEF 2023), anticipated ingredients will be categorised as either pre-requisite[6], start up, and recurring costs.

Ingredients might relate to, for instance:

- Programme fees: Reflecting subscription fees, and school access to training and materials based on market value.
- Staff time for teacher training, preparation, and delivery of PoR, with a separate identification of the cost of new hires and supply staff.
- Any additional (unpaid) staff time supporting the delivery of PoR

Program fees calculation will rely on information provided by the Delivery Team. Additionally, cost data will be collected through post-intervention headteacher and teacher surveys participating in the implementation and process evaluation.

Ethics and registration

Ethical clearance was granted by Manchester Metropolitan University following an initial submission on 27th January 2025. The approval process required detailed documentation on project design, ethical procedures, participant information sheets, consent and withdrawal forms, the Memorandum of Understanding, and privacy notices.

The school recruitment process follows a structured approach: the Delivery Team identifies and contacts schools that meet the selection criteria, collecting key details such as school

[6] Pre-requisites refer to the resources necessary for programme implementation that are already available within schools. Although these goods are not included in the main cost estimates of the programme - since they do not require additional financial outlay from schools - they will be identified and included in a sensitivity analysis as per the EEF cost guidance (2023).

information, contact details, and the names of Year 5 teachers. Schools are then invited to sign a Memorandum of Understanding, which outlines project details, objectives, expected benefits, timelines, data protection considerations, and the responsibilities of all parties. Additionally, schools must sign a separate Data Sharing Acknowledgement, which defines how personal data will be collected and shared between the Delivery Team, the Evaluation Team, and the school.

Parents of Year 4 pupils of the selected class receive a withdrawal notice, giving them a two-week response period, with the option to withdraw their child at any time. Once this period ends, FFT collects baseline data from each school. The trial will be registered on the Open Science Framework registry once this protocol is finalised.

The trial is registered on the Open Science Framework (OSF). Please refer to https://doi.org/10.17605/OSF.IO/YVQZF

Data protection

MMU and CLPE are independent data controllers for this project. They make decisions about how and what personal data is used in accordance with the purposes set by the EEF. There will be some instances where CLPE will also act as data processor for MMU, where CLPE will collect data on behalf of MMU (CLPE will ask schools to sign the data sharing agreement between each school and MMU). FFT and AlphaPlus are data processors. The EEF will become data controller for the data once it is archived at the end of the project.

As public authorities conducting research and analysis in the public interest which has undergone ethical approval, MMU uses the following lawful bases for the processing of:

- Personal data: 'Public Task' – GDPR Article 6(1)(e);
- Personal data defined as special category: 'Research purposes in the public interest' – GDPR Article 9(2)(j).

We are not collecting any personal data defined as special category in this project although we are collecting SEN (EHC or Support plan), FSM and EAL status for pupils which is considered by the DfE to be sensitive data.

The project involves collecting writing assessments and surveys from pupils, survey, and interview data from headteachers and teachers, observations of Power of Reading training and lessons, and interviews with and pupils. Data will be processed by MMU to ascertain the impact of Reading Plus on the pupil outcomes above, and to make judgements about compliance and fidelity, as well as stakeholders' experiences of Power of Reading.

All assessment data will be accessed and analysed by the Evaluation team (Manchester Met). For the purpose of this project, some personal data (staff names/contact details; pupil names) will be shared between MMU and CLPE, MMU and FFT, and MMU and AlphaPlus (see privacy notices for parents and school staff for full details Privacy Notices.

MMU will access the National Pupil Database as part of the evaluation to access pupils' KS1 and KS2 assessment data. For the purposes of archiving and further research, pupil-level impact data (e.g., writing assessment scores) and pupils' personal data (e.g., name, date of birth, UPN) will be shared with EEF's Data Archive manager (FFT). FFT will match pupil records with the Pupil Matching Reference (PMR), an identifier that is stored in the National Pupil Database, and is meaningless to anyone outside the Department for Education (DfE). All other personal data that can identify an individual pupil will be then deleted prior to archiving. The remaining data will then be archived in the Office for National Statistics' Secure Research Service (ONS SRS). After anonymisation, Data collected as part of evaluations funded by the Education Endowment Foundation (EEF) are archived in order to estimate the long-term impact of interventions. Further information about how EEF store and process data for follow-up analysis can be found here: https://educationendowmentfoundation.org.uk/privacy-notices/privacy-notice-for-the-eef-data-archive.

Pupil and staff data will be treated with the strictest confidence, given a unique code immediately after collection and prior to analysis in order to reduce risk, and stored securely in line with General Data Protection Regulation (GDPR) and the Data Protection Act 2018 (DPA).

MMU shall ensure that a data sharing agreement is in place as required by the GDPR and DPA. This document will clearly outline the data sharing and protection responsibilities of the five parties involved with this arrangement (Manchester Met, CLPE, FFT, AlphaPlus and EEF). We will not use pupils' names or the name of the school in any report arising from the research. We expect that your pupils will enjoy their involvement in the project, and they will be free to withdraw at any time.

So that the processing of personal data relating to the pupils is fair, lawful, and transparent, we will use a parent information sheet, parental withdrawal form, and a privacy notice for parents agreed with the University's Data Protection Officer. Parents are free to withdraw their child from the evaluation and/or ask to have any or their child's information deleted until September 30th, 2027, without giving a reason. Schools will need to keep a record of which children have withdrawn from the project to ensure that children whose parents have 'opted out' are not included in any evaluation activities (including any related data sharing). Please note that withdrawn pupils will be involved in the intervention itself, but not the evaluation (e.g. they will not participate in any testing). Requests to withdraw or delete information after this point may not be possible but will always be considered.

Schools can withdraw from the evaluation at any time, without giving a reason. They can ask for the school's data to be deleted at any time, but it might not always be possible. If a request is made to withdraw any personal data up until 31st August 2027, we will make sure this is done. However, analysis of the baseline and endline survey data will begin in November 2025 and July 2027, respectively. After these points, anonymised staff survey responses may already have been included within the analysis, and it may not be possible to remove them.

The information collected will be used for research purposes only and no information that can identify individuals will be used for any other purpose. Any personal data collected and held by

Manchester Met, CLPE, FFT and AlphaPlus will be destroyed in accordance with the GDPR when it is no longer required, and no later than 31st July 2029.

Personnel

Delivery team – Centre of Literacy in Primary Education

Farrah Serroukh is Research and Development Director at the Centre for Literacy in Primary Education (CLPE): She works closely with the teams within and across CLPE, EEF and MMU to ensure the successful delivery of this trial from inception to evaluation.

Anjali Patel is Lead Advisory Teacher at the Centre for Literacy in Primary Education (CLPE). She brings decades of educational experience to this EEF trial in partnership with the evaluation team at MMU. She is responsible for overseeing trial delivery, ensuring that schools are supported to implement the Power of Reading programme effectively through a quality-assured training programme and adherence to protocols and trial conditions.

Dr Jonny Rodgers is the trial Project Manager at the Centre for Literacy in Primary Education (CLPE). He is responsible for managing all aspects of the project delivery.

Evaluators – Manchester Metropolitan University

Dr Steph Ainsworth, Education and Social Research Institute, Manchester Metropolitan University: Steph Ainsworth is a Reader in Education. Her expertise includes primary English teaching and the assessment of reading and early language skills. Steph is the joint Principal Investigator and is responsible for design and management of the impact and process evaluation.

Prof Stephen Morris, Policy and Evaluation Research Unit, Manchester Metropolitan University: Stephen Morris is Professor of Evaluation. He specialises in experimental/quasi-experimental evaluation designs. Steve is the joint Principal Investigator and is responsible for the design and management of the impact evaluation

Sandor Gellen, Policy and Evaluation Research Unit, Manchester Metropolitan University: Sandor is a Senior Research Associate with expertise in evaluating programmes using quantitative and small-n mixed methodologies. Sandor is a Co-Investigator and responsible for analysis of the impact evaluation data and managing the relevant data flows.

Dr Helen Lawson, Education and Social Research Institute, Metropolitan University: Helen Lawson is a Research Fellow, She has substantial experience managing/supporting education evaluations (e.g. Hodgen et al., 2023; TASO, 2023). She has expertise in IPE design and theory-based evaluation methods. Helen is a Co-Investigator and responsible for project management.

Prof Cathy Lewin, Education and Social Research Institute, Metropolitan University: Cathy Lewin is Professor of Education. She has extensive experience of mixed-method evaluation of school-based interventions, including educational technology and inclusive education. Cathy will draw on this experience to provide independent quality assurance during the evaluation (e.g. reviewing data collection instruments, the protocol, final report).

Risks

Table 6 Risks

Risk	Likelihood	Impact (1-3)	Detail/preventive measure
Recruiting sufficient numbers of schools	Medium	2	• The delivery team will send prospective emails to schools in collaboration with their partners National Literacy Trust to broaden reach. This approach will be complemented by adverts on social media and emails to contacts within the delivery and evaluation team's existing networks, e.g. partnership schools, local authorities, research schools, etc. EEF will also promote the trial through its website, social channels, and networks. • Reducing burden on schools, e.g. FFT enumerating the sample; using a short writing measure at baseline. • Intervention schools are asked to contribute £300 per school plus VAT towards the costs of the PoR training. This represents a small contribution towards the usual cost of the training, which is on average £1450+ VAT . • Control Schools are offered payment of £500 per school if all required data are provided and the Evaluation team is given access to the school to conduct all assessments with participating Year 5 pupils. • Interested schools that do have mixed age group classes will be added to a recruitment waitlist. Similarly, if schools are participating in another EEF trial and are interested in this project, the possibility will be held open for them to also be added to a recruitment waitlist.

Missing data and sample attrition	High	2	Schools recruited to the trial may decide to withdraw, and this sample loss might both reduce precision of statistical estimates and introduce bias. Attrition will be minimised through the following steps: • The randomised controlled trial model will be explained to schools during recruitment. • The value of control schools will be explained in initial discussions. • Effective communication strategy throughout the trial to keep all schools engaged. • Financial incentives will be provided to control schools that choose to remain in the trial (see above). • Aim to over recruit in the main trial to allow for some attrition (see sample size scenarios above). • Schools participating in the trial will be offered PoR at a discounted rate once the trial has ended. • Mis-recording of identifying data for pupils, mitigated through carrying out extensive checks on pupil records prior to randomisation. • Recruitment documentation stresses the importance of the study encouraging children and their parents to remain in the trial. More broadly, where attrition occurs, steps will be taken in analysis to test various assumptions regarding missingness and assess consequences for bias and precision using approaches such as multiple imputation of the estimation of bounds.
Contamination and measuring compliance	Medium	2	Strong communications strategy to limit contamination. Communicating expectations clearly to control schools. We will identify what control schools do in relation to teaching reading/use of other reading programmes through the teacher/headteacher surveys. We will work closely with developers to arrive at a definition of compliance.
Staff shortages and retention in the evaluation team	Medium	2	Research projects spanning extended durations often encounter turnover among research staff. With a substantial pool of suitably qualified and experienced personnel, coupled with adaptable workload and staff management systems, we can consistently maintain adequate staffing and effective project management. Additionally, we will establish procedures such as maintaining a comprehensive variable library and log for each data source, as well as implementing appropriate handover processes if necessary.
Poor communications between stakeholders	Low	3	As an integral aspect of our project planning, we intend to conduct frequent meetings with the Delivery Team, especially as we approach significant milestones such as school recruitment and enumeration. Additionally, we will consistently provide our project plans and risk management documents to the developers, fostering alignment between our management processes and those of the project team. This synchronization will enable us to collectively and effectively address any emerging challenges that may arise.

Schools being unable to support supervised testing in schools	Medium	3	The new DfE guidelines in relation to Research with Children and Young People state that researchers should not be left alone with a pupil 1:1 even if they have a DBS. This might cause difficulties as schools are currently experiencing staff shortages and so it is predicted that they will find it difficult to release staff to supervise testing. To mitigate this risk, AlphaPlus test administrators will be flexible in their approach to this, suggesting with schools in advance a number of options for how these guidelines might be met in ways that minimise the burden to schools, e.g., testing could take in a corridor in sight of other staff and pupils; in the staff room, PPA room, or another space where there are other people present (teachers engaged in their PPA activities in the background, office staff, teaching assistants, etc.).
Small number of FSM children for analysis	Medium	2	Prioritising recruitment of schools in EIAs to support FSM subgroup analysis.

Timeline

Table 7 Timeline

Dates	Activity	Staff responsible/ leading
Nov – Dec 2024	Start-up meetings/theory of change	MMU/CLPE/EEF
Dec 2024 – Jan 2025	Data governance/MoU, parental withdraw & data processing notices drafted and agreed	MMU
Dec 2024 – Jan 2025	Finalise trial/IPE design	MMU
Jan – Feb 2025	Ethical approval	MMU
Feb – July 2025	School recruitment (Cohort 1)	CLPE
September – October 2025	Administration of headteacher baseline survey (Cohort 1)	MMU
August 2025	Protocol	MMU
Sept 2025	Class selection (Cohort 1)	MMU
Sept – Oct 2025	Baseline assessment (WAM, WSEM, LWS, RES) (Cohort 1)	AlphaPlus
Oct 2025	Randomisation (Cohort 1)	MMU
Oct - Dec 2025	SAP	MMU
Sept – Oct 2025	Teacher baseline survey (Cohort 1)	MMU
Nov 2025	Observe first training session/intervention commences (Cohort 1)	MMU
(tbc with developers)	Observe remaining training sessions (Cohort 1)	MMU
Nov 2025 – June 2026	Intervention delivery (Cohort 1)	CLPE
Jan – Feb 2026	First case study visits (Cohort 1)	MMU
Apr – May 2026	Second case study visits (Cohort 1)	MMU
Mar – May 2026	Recruit and train endline data collectors	MMU
Jun – July 2026	Endline teacher/headteacher surveys (Cohort 1)	MMU
(end of) May – July 2026	Endline pupil data collection in schools (Bespoke writing assessment, WSEM, LWS, RES) (Cohort 1)	AlphaPlus
Feb – July 2026	School recruitment (Cohort 2)	CLPE
Sept – Oct 2026	Administration of headteacher baseline survey (Cohort 2)	MMU

Sept 2026	Class selection (Cohort 2)	MMU
Sept – Oct 2026	Baseline assessment (WAM, WSEM, LWS, RES) (Cohort 2)	AlphaPlus
Oct 2026	Randomisation (Cohort 2)	MMU
Sept – Oct 2026	Teacher baseline survey (Cohort 2)	MMU
Nov 2026	Observe first training session/intervention commences (Cohort 2)	MMU
(tbc with developers)	Observe remaining training sessions (Cohort 2)	MMU
Nov 2026 – June 2027	Intervention delivery (Cohort 2)	CLPE
Jan – Feb 2027	First case study visits (Cohort 2)	MMU
Apr – May 2027	Second case study visits (Cohort 2)	MMU
Mar – May 2027	Recruit and train endline data collectors	MMU
Jun – July 2027	Endline teacher/headteacher surveys (Cohort 2)	MMU
(end of) May – July 2027	Endline pupil data collection in schools (Cohort 2)	AlphaPlus
July – Sep 2027	Data linking, cleaning, and structuring	MMU
Sep – Oct 2027	Analysis of IPE data	MMU
Sept – Oct 2027	Analysis for the IE	MMU
Nov – Dec 2027	Reporting – main evaluation report	MMU
Jan – Mar 2028	Analysis within the ONS SRS (end of KS2 reading assessments)	MMU
Apr – May 2028	Reporting - addendum	MMU

References

Anders, J., Shure, N., Wyse, D., et al. (2021a). *Power of Pictures Evaluation Report*. London: Education Endowment Report. Available at Power of Pictures - trial | EEF (educationendowmentfoundation.org.uk)

Anders, J., Shure, N., Wyse, D., et al. (2021b). *The Craft of Writing Evaluation Report*. London: Education Endowment Report. Available at Power of Pictures - trial | EEF (educationendowmentfoundation.org.uk)

Booton, S. A., Menon, P., Kyriacou, M., & Murphy, V. A. (2025). Exploring the writing attainment gap: profiling writing challenges and predictors for children with English as an additional language. *Reading and Writing, 38*(3), 843-863.

Braun, V., & Clarke, V. (2022). Thematic analysis: a practical guide. Los Angeles: Sage.

Bruning, R., Demspey, M., Kauffman, D. F., McKim, C., & Zumbrunn, S. (2013). Examining Dimensions of Self-Efficacy for Writing. Journal of Educational Psychology, 105(1), 25-38. https://doi.org/10.1037/a0029692

Carril, A. (2017). Dealing with misfits in random treatment assignment. *The Stata Journal, 17*(3), 652–667.

CLPE (2023). *Power of Reading Impact Report*. London: Center for Literacy in Primary Education

Culliney, M, Moore, N, Coldwell, M, & Demack, S. (2019). *Integrating English Evaluation*. London, Education Endowment Foundation.

Doherty, J (2019) *Diminishing the gap for disadvantaged pupils in Key Stage 2 Reading*, Exceed teaching schools, exceed academies trust, Leeds Trinity University

Dong, N., & Maynard, R. (2013). PowerUp!: A Tool for Calculating Minimum Detectable Effect Sizes and Minimum Required Sample Sizes for Experimental and Quasi-Experimental Design Studies, *Journal of Research on Educational Effectiveness*, 6(1), 24-67, DOI: 10.1080/19345747.2012.673143

Dunsmuir, S., Kyriacou, M., Batuwitage, S., Hinson, E., Ingram, V., & O'Sullivam S. (2015). An evaluation of the Writing Assessment Measure (WAM) for children's narrative writing. *Assessing Writing, 23*, 1-18. https://doi.org/10.1016/j.asw.2014.08.001

Education Endowment Foundation (2022). Implementation and process evaluation guidance for EEF evaluations. London: Education Endowment Foundation. Retrieved from: https://d2tic4wvo1iusb.cloudfront.net/production/documents/evaluation/evaluation-design/EEF-IPE-Guidance-August-2022.pdf?v=1761902435

Education Endowment Foundation (2023). Cost evaluation guidance for EEF evaluations. London: Education Endowment Foundation. Retrieved from: https://d2tic4wvo1iusb.cloudfront.net/production/documents/evaluation/evaluation-design/Cost-Evaluation-Guidance-Feb_2023.pdf?v=1710150030https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/Cost_Evaluation_Guidance_2019.12.11.pdf

Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. London: W. W. Norton & Company

Grima, G., Hooper, A., Mullins, L., Redmond, B., Sharott, P., & Ashworth, B. (2024). *Understanding current practice and research priorities in teaching Writing.* London: Education Endowment Foundation.

Heß, S. (2017). Randomization inference with Stata: A guide and software. *The Stata Journal, 17*(3), 630-651.

Imbens, G. W., & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge university press.

Kim, J. S., Olson, C. B., Scarcella, R., Kramer, J., Pearson, M., van Dyk, D., ... & Land, R. E. (2011). A randomized experiment of a cognitive strategies approach to text-based analytical writing for mainstreamed Latino English language learners in grades 6 to 12. *Journal of Research on Educational Effectiveness, 4*(3), 231-263.

Koons, H. H. (2008). *The reading-writing connection: An investigation of the relationship between reading ability and writing quality across multiple grades and three writing discourse modes* (Doctoral dissertation, The University of North Carolina at Chapel Hill).

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies, Oxford University Press, 76*(3), 1071-1102.

Lin, W., (2014). Comments on Covariance Adjustments for the Analysis of Randomized Field Experiments. *Evaluation Review 38*, 449–451.

Lewin, C., Morris, S., Ainsworth, S., Gellen, S., & Wicker, K. (2024). *Peer Assisted Learning Strategies UK (PALS-UK): A whole class reading approach—Evaluation Report*. London, Education Endowment Foundation.

Maisuria, A., Roberts, N., Long, R., & Donechi, S. (2023). Teacher recruitment and retention in England. *HC Library, Briefing Paper*, *7222*, 12.

McGeown, S., & Smith, K. C. (2023). Reading Engagement Matters! A New Scale to Measure and Support Children's Engagement with Books. The Reading Teacher, 77(4), 462-472. https://doi.org/10.1002/trtr.2267

Ofqual (2014). Marking Reliability of the Key Stage 2 National Curriculum English Writing Tests in England. Coventry: Ofqual. Available at https://assets.publishing.service.gov.uk/media/5a81d19de5274a2e87dbfa4d/0214_Ofqual-marking-reliability-of-the-ks-2-nc-english-writing-tests-in-england.pdf

Pearson Clinical, (no date). *Wechsler Individual Achievement Test – Third UK Edition. Pearson Clinical.* Retrieved August 26, 2025, from https://www.pearsonclinical.co.uk/en-gb/Store/Professional-Assessments/Academic-Learning/Reading/Wechsler-Individual-Achievement-Test---Third-UK-Edition/p/P100009274

Slavin, E.R., Lake, C., Inns, A., Baye, A., Dachet, D., Haslam, J. (2019). *A Quantitative Synthesis of Research on Writing Approaches in Years 3 to 13*. London: Education Endowment Foundation.

Stone, G, Andrade, J, Martin, K, & Styles, B. (2022). *Helping handwriting shine - Evaluation Report,* London: Education Endowment Foundation.

Tauchmann, H. (2014). Lee (2009) treatment-effect bounds for non-random sample selection. *The Stata Journal*, *14*(4), 884–894.

Torgerson, D. J., Torgerson, C. J., Ainsworth, H., Buckley, H., Heaps, C., Hewitt, C. E., & Mitchell, N. (2014). *Improving Writing Quality: Evaluation report and executive summary*. London: Education Endowment Foundation.

Tracey, L, Boehnke, J, Elliot, L, Thorley, K., Ellison, S, & Bowyer-Crane, C. (2019). *Grammar for Writing - Evaluation report*, London, Education Endowment Foundation.