Evaluation of the Peer Assisted Learning Strategies for Reading UK (PALS-UK) intervention, a two-armed cluster randomised trial

Evaluation Protocol

Evaluator (institution): Manchester Metropolitan University
Principal investigator(s): Cathy Lewin and Stephen Morris

PROJECT TITLE	**Evaluation of the Peer Assisted Learning Strategies for Reading UK (PALS-UK) intervention, a two-armed cluster randomised trial**
DEVELOPER (INSTITUTION)	Dr Emma Vardy (Nottingham Trent University) and Dr Helen Breadmore (University of Birmingham), supported by Dr Luisa Tarczynski-Bowles (Nottingham Trent University)
EVALUATOR (INSTITUTION)	Manchester Metropolitan University
PRINCIPAL INVESTIGATOR(S)	Cathy Lewin and Stephen Morris
PROTOCOL AUTHOR(S)	Steph Ainsworth, Sandor Gellen, Cathy Lewin and Stephen Morris
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)*	108[1]
NUMBER OF PUPILS *(at design stage)*	4185
PRIMARY OUTCOME MEASURE AND SOURCE	Reading attainment (New PiRA Summer 5 Test)
SECONDARY OUTCOME MEASURE AND SOURCE	Reading comprehension (WIAT-III UK) Oral reading fluency (WIAT-III UK, Multi-dimensional fluency scale Reading self-efficacy (Feelings about reading questionnaire) Motivation for reading (Feelings about reading questionnaire)

Protocol version history

VERSION	DATE	REASON FOR REVISION
1.0 [*original*]		N/A

[1] We assume a total number of 108 schools, which should give us a buffer for attrition at school level given that our recruitment target is 120 schools

Table of contents

Study rationale and background

Peer Assisted Learning Strategies for Reading (Fuchs et al., 1997) is a whole class intervention designed to improve children's oral reading fluency and reading comprehension. The version of the programme evaluated here is PALS-UK, based on the PALS grades 2-6 programme, with materials and training adapted for the UK and modernised by Dr Emma Vardy and Dr Helen Breadmore. Pupils work in pairs, taking it in turn as coach and reader as they engage with four activities: partner reading, re-tell, paragraph shrinking and prediction relay. Sessions last 35 minutes, conducted three-times-a-week over 20 weeks. The logic model predicts that repeated reading with peer feedback will support all aspects of fluency: accuracy, automaticity and prosody, while the tasks of re-tell, paragraph shrinking and prediction relay will support reading comprehension. Taken together, the intervention has the potential to develop pupils' fluency, self-efficacy in reading, motivation for reading, reading comprehension and reading attainment.

The EEF's Teaching and Learning Toolkit reports that peer tutoring approaches can have positive impacts on pupil outcomes. PALS-UK exemplifies a number of common elements of effective peer tutoring interventions, including the provision of training to both teachers and pupils and the use of structured activities to support high quality peer interactions. PALS-UK also exemplifies elements of another promising strand of the EEF's Teaching and Learning Toolkit, reading comprehension strategies.

Peer interventions are low cost and have been found to generate moderate/high effect sizes (EEF, 2018; Topping et al., 2011). The PALS approach has been extensively implemented in the United States and a number of studies have reported positive impacts on pupil outcomes (WWC, 2012). However most experimental research of PALS has been small-scale and developer-led, and there is a lack of independent evaluations investigating PALS in later primary years. UK research is limited and results of same-age peer tutoring programmes disappointing (Lloyd, Edovald, et al., 2015). PALS-UK has been identified as having potential for scale-up due to its structured whole class, same-age approach. An evaluation of PALS-UK will strengthen the evidence-base around peer-assisted reading programmes. We will evaluate the causal mechanisms proposed within the logic model through careful selection of primary and secondary measures and a comprehensive mixed methods approach (Morris et al., 2016). This large randomised controlled trial aims to provide a robust estimate of the impact of PALS-UK on pupils' reading outcomes. The EEF previously conducted an efficacy trial of PALS-UK, however, the trial was disrupted by the Covid-19 pandemic. As a result, the EEF has commissioned this second trial of PALS-UK.

The evaluation design will integrate three strands: an impact evaluation, implementation and process evaluation, and a cost evaluation. The former will consist of a randomised control efficacy trial allowing us to assess the impact of the PALS-UK intervention on Year 5 children's reading attainment, reading comprehension, oral reading fluency, reading self-efficacy and motivation for reading. The implementation and process evaluation will allow us to assess the extent to which the intervention was implemented as intended, identify any barriers to implementation, and investigate the causal assumptions underpinning the intervention. The cost evaluation will evaluate the costs incurred by schools in relation to implementation of PALS-UK.

This trial is informed by a previous EEF efficacy trial of PALS-UK (Culora et al., 2022), which was conducted recently but was impacted by the COVID-19 pandemic. The current trial seeks to address three important shortcomings of the previous trial relating to COVID-19 disruption:

large attrition (25.8% school level and 36.1% pupil-level attrition) which was due to the COVID restrictions in place, the delay of post-test measures due to lockdown restriction, leading to a likely diluted impact of PALS-UK on the intervention, and failure to complete the measurement of reading comprehension and oral reading fluency measures. As stated in the previous evaluation report, 'a key missing link here is whether exposure to the programme first results in improvements in reading fluency and comprehension' (Culora et al., 2022, p.100). The current design will allow post-test measures to be taken close to the end of the intervention and for robust measures of the intermediate outcomes, reading comprehension and oral reading fluency, to be collected to allow evaluation of key aspects of the logic model. The logic model itself has been augmented to include elements which were observed to be important in the previous trial (see next section). A full list of changes from the previous trial to the intervention and evaluation are detailed below.

The intervention for the current trial will be largely the same as for the previous evaluation, but with a few minor changes to the support provided to teachers. For example, the manual and initial and top-up training materials have been updated following feedback from the first trial. The book list has been updated. As in the previous trial, ongoing support will be provided for teachers in the intervention schools in the form of 'just-in-time' support. This support will be more extensive and include more components which can be accessed asynchronously online, such as additional support videos, FAQs and discussion boards, as well as twilight support webinars. Unlike the previous trial, in this trial schools will be encouraged to use the baseline reading attainment data to inform how children are paired in the intervention. Teachers will be encouraged to change pairs every four weeks. When pairing the children, first readers will remain first readers and second readers will remain second readers throughout the twenty weeks.

Recruitment procedures will be similar to the first trial except the exclusion criteria have been revised to focus on the three required Regional School Commissioner Regions, schools must not be taking part in another Accelerator Fund efficacy trial or another EEF literacy project in Year 5 in the 2022-23 academic year, and schools must provide access to technology for online assessment of reading. The decision was made to recruit schools who were in the control group within the previous EEF efficacy trial. This decision might lead to some schools being assigned to the control group for a second time, potentially increasing the risk of attrition. However, this risk was considered to be relatively small, when weighed against the need to recruit enough schools within a relatively short timeframe to achieve the required statistical power for an efficacy trial. The full exclusion criteria are listed below.

There were a number of changes made to the evaluation from the previous trial. In the previous trial the PiRA Autumn 5 test was administered at baseline in the autumn term. In the current trial, baseline testing will occur in the summer term of year 4 and the PiRA Summer 4 test will be used. This is because PiRA is designed to be administered in the second half of a term (after a certain amount of content has been delivered). So the PiRA Autumn 5 test is not appropriate for administration in September 2022 as much of the content covered would not yet have been taught. In addition, we wish to minimise the burden on schools at the beginning of the intervention in the Autumn term and allow time for randomisation, notifying schools and arranging the initial training. The WIAT-III UK-T will be used to assess the secondary outcomes of reading comprehension and oral reading fluency. In the previous trials, measurement of these secondary outcomes could not be completed because of disruption related to the pandemic. In addition, we will also include a complementary measure of fluency

alongside the WIAT-III UK-T. While the WIAT-III UK-T oral reading fluency subtest provides a basic measure of fluency (number of words correct per minute), the additional measure, the Multidimensional Fluency Scale (MDFS) (Rasinski, 2004) provides a qualitative measure of fluency based on judgements of: expression and volume, phrasing, smoothness and pace. As in the previous project, the measure of self-efficacy will be completed pre and post-test online. Within the previous trial a measure of motivation for reading was taken at the same time as reading self-efficacy, but only the reading self-efficacy data was analysed. Within the current trial motivation for reading will also be included as a secondary outcome. The timing of the primary outcome post-test will also be different. In this trial the testing will take place in the term following the end of intervention (Summer 2023), whereas in the previous trial testing was delayed by at least 6 months due to pandemic-related disruption. The administration of the post-tests will also differ. In the previous trial, the post-tests were administered by the teachers because of 'no-visitor' policies in response to the pandemic. In the current trial post-testing will be administered by the evaluation team. In the previous trial, the WIAT-III UK-T the reading comprehension subtest was administered remotely. In the current trial both the oral reading fluency and reading comprehension subtests will be conducted face-to-face to avoid the potential logistical issues associated with administrating a fluency test remotely and given that normative data from the WIAT-III UK-T were collected face-to-face.

Further points of departure from the previous trial are as follows. We will hold schools previously allocated to control as a reserve pool to approach if we encounter significant challenges in recruitment. We will work with our partner, FFT Education (part of the Fischer Family Foundation), to automate collection of enumeration data. Unlike in the previous trial, randomisation won't be stratified by region. Instead, we introduce FSM eligibility as the second stratifying variable alongside school size. The previous trial used age-standardised PiRA scores in the primary analysis as an outcome measure. The current trial will use raw (unstandardised) PiRA scores, and will control for the effect of age by adding it to the multilevel regression model as a pupil level covariate. The previous trial analysed only a single class in each school, and therefore used two-level clustered designs (pupils nested in schools). We will use three-level designs to account for school-, and class level intra-cluster correlations. The previous study used three sub-categories: children who speak English as an additional language (EAL), children eligible for Free School Meals (FSM) and high vs. low reading achievers for their sub-group analysis. In the current trial, the subgroups of interest will be pupils that are ever-FSM, designated SEND and pupils scoring in the lowest quartile on the baseline PIRA test. Further exploratory analysis will examine the effects of PALS-UK for EAL pupils, but we will include only those children whose score on the baseline reading assessment falls in the lower half of the sample distribution (see the Analysis section for details). The previous trial set the following compliance criteria: (1) Attendance at all PALS-UK initial training sessions; (2) Completion of the four weeks of training to the manual; (3) Completion of PALS-UK delivery (minimum of 12 weeks delivered). The current trial defines compliance based on two criteria: (1) Attendance at PALS-UK training and (2) Completion of the four weeks of pupil training in line with the manual (see section on Compliance below for justification).

In relation to the IPE, in comparison to the original trial of PALS-UK, this trial will collect more data. The original trial only conducted a post-intervention survey with teachers in the intervention arm of the trial. In this evaluation, all teachers will also be surveyed at baseline, although after randomisation. This is because schools would like the outcome of randomisation as soon as possible to assist with planning (ie first week in September) and Year 5 teachers will not be confirmed until that point so it is not possible to survey them in the summer term 2022. All teachers will also be surveyed after the intervention has completed

including those in control schools. The evaluation team will conduct independent observations of PALS-UK on two separate occasions through onsite case study visits; observations were not undertaken by the evaluators in the first trial. The first evaluation collected interview data from headteachers and teachers at case study schools remotely. These data will be collected during face-to-face visits. We will also collect data from pupils through conducting focus groups in case study schools. The previous trial did not collect qualitative data from pupils at case study schools.

Intervention

Peer Assisted Learning Strategies for Reading UK (PALS-UK) is a whole-class intervention designed to improve reading comprehension and oral reading fluency. The intervention is centred around children working together in pairs engaging in a set of structured reading activities three times a week for 20 weeks. For the first four weeks children are trained on the PALS-UK activities and how to work well together, then in the following 16 weeks self-direct their learning. Each PALS session takes approximately 35 minutes (which includes moving/set up times) and the intervention is aimed at children in Key Stage 2. Here we focus on children in Year 5 (aged 9-10 years). Within each pair, one of the children is given the role of first reader. This is the child who the teacher considers to have a higher current attainment of reading relative to the other child in the pair (the second reader). Within the activities children take turns to act as both coach and reader.

Each of the PALS sessions follows the structure below:

1. **Partner reading** (10 minutes): Each child reads for 5 minutes while the other child coaches them.
2. **Retell** (2 minutes): The second reader retells what has just happened in the text that they have just read.
3. **Paragraph shrinking** (10 minutes): The first reader reads the next paragraph and then is coached to summarise the paragraph.
4. **Prediction relay** (10 minutes): The children take turns to predict what will happen in the next half a page, they then read the page and check whether the prediction comes true.

Partner reading aims to develop oral reading fluency whilst retell, paragraph shrinking and prediction relay aim to support reading comprehension. All activities are designed to develop reading self-efficacy and motivation for reading as they involve scaffolded tasks in a safe learning environment supported by peers.

Participating Year 5 teachers will receive face-to-face training from the developer team: Dr Emma Vardy (Nottingham Trent University) and Dr Helen Breadmore (University of Birmingham). The first session, which will take place before the intervention begins, will last for one day and provide the teachers with training around the skills needed for reading comprehension as well as a detailed description of the intervention. This training day will also provide the teachers with the opportunity to practice the PALS activities and to discuss key implementation issues such as how to pair pupils appropriately. The second training session will be online and will be a shorter top-up session (half a day), which will take place 4 weeks into the intervention. This session is designed to support teachers in selecting books appropriately, changing pupil-pairings effectively (this is supposed to happen every 4 weeks), and providing opportunities to share good practice.

Once the teachers have received their pre-intervention training, PALS then begins with 4 weeks of whole-class training for the children. Children are trained to work effectively in their pairs and to work through the sequences of activities set out above. Within each training

session the teacher will introduce a new skill so as not to overload the children. All participating teachers will be provided with a PALS-UK manual which provides detailed instructions for the whole-class training. PALS-UK lasts for 20 weeks in total: 4 weeks of whole class training, followed by 16 weeks of the PALS sessions with children changing pairs every 4 weeks. Pairings are based on the teacher's assessment of the children's current reading attainment. Teachers perform a median split based on teacher assessment of reading level informed by the baseline PiRA scores (New PIRA Test 4, Summer), but teacher's knowledge of pupils abilities will be also included in this judgement. The top half of the class will be assigned as first readers and the bottom half of the class will be assigned as second readers - these assignments will be maintained throughout the 20-week intervention. Initially it is recommended that teachers pair the highest attainer from the above-median group with the highest attainer of the below-median group and so on. This is to ensure a difference between the two levels of reading attainment but one that is not too large. Some flexibility is allowed in the pairing process so that teachers can take into account other factors (e.g. personality) that might affect the pupils' ability to work well as a pair, and teachers are advised to shuffle pairings around if any of the pairings are not working well. Book selections will be made by the teacher on the basis of the reading attainment of each pair (training will be provided on this at the top-up training), with children given some choice over the text to help with motivation levels. Children with a severe SEND or with particular English language difficulty can have a teaching assistant join their pair to act as an additional coach if appropriate. The intervention will take place in the Autumn and Spring terms within year 5 (October 22-April 23).

To support schools with the delivery of PALS-UK, four observations during the 20 weeks will take place. These observations are meant to be supportive and help teachers. Two observations will be completed by research assistants from Nottingham Trent University; the first observation will take place during the initial four weeks of pupil training and the second observation will take place during the 16 weeks of delivery. These observations will help to identify schools where additional support is needed. Two further observations will be completed by the nominated peer observer within the school; this is to explore from a professional opinion how PALS-UK fits within the school environment. Both of these observations will take place during the 16-week part of the programme. A structured observation sheet developed by Prof Kristen McMaster will be used by the peer observer and the research assistants.

Control schools will not receive any training and will not deliver the PALS-UK intervention. They will follow a 'business as usual' approach to the teaching of reading. The incentive provided to control schools will be £500 on the completion of the post-testing. This is in contrast to the payment of £100 intervention schools will need to make to contribute to delivery costs. At the end of the trial, control schools will be invited to join a mailing list to find out more about PALS-UK and may use some of the £500 to access training and the manual. Both intervention and control schools will have access to the PiRA data for participating year 5 children in their school as an additional incentive.

The logic model (below) lays out the theoretical relationships between elements of the intervention and the pupil outcomes that will be measured in the impact evaluation. PALS-UK is primarily designed to develop oral reading fluency and reading comprehension. The emphasis on regular practice at reading aloud, re-reading and receiving feedback from a peer is predicted to improve reading fluency (Rasinski, 2003). The other activities of re-tell, paragraph shrinking and prediction relay are each designed to encourage children to make meaning as they read, thus supporting them with opportunities to develop their reading comprehension skills.

As the children read for meaning they are likely to develop their vocabularies which in turn will support both reading comprehension (once you become familiar with reading a word in context you acquire an understanding of what it means) and reading fluency (it is quicker to read a word that you know, and if you understand what the words mean, you are more likely to read with appropriate expression/intonation). The relationships between vocabulary, reading comprehension and oral reading fluency are therefore likely to be reciprocal (Breadmore et al., 2019; Pikulaki & Chard, 2011).

Through regular practice at reading (and through the associated gains in fluency and comprehension), reading becomes less effortful and therefore children are likely to experience increases in reading self-efficacy, motivation, confidence and positive attitudes towards reading (Peura et al., 2019; Reis et al., 2008). Changes in each of these outcomes might, in turn, lead to a child reading more often for pleasure, which is likely to lead to further gains still in terms of reading self-efficacy, confidence etc (Breadmore et al., 2019; Clark & Rumbold, 2006). As well as influencing the child at the individual level, an increase in reading for pleasure amongst pupils, can help to create/sustain a positive reading culture at the class/school level.

This broader school reading culture is also likely to be influenced by the teachers' knowledge of reading, which is also targeted within the intervention. The training sessions are designed to improve teachers' understanding of the skills required to be a skilled reader and provide them with pedagogical strategies to promote these skills within a supportive environment centred around peer learning. The potential influence of PALS-UK on pupil vocabulary and teacher knowledge were highlighted within the previous trial and were therefore added to the current logic model.

An additional outcome predicted to improve as the intervention progresses is peer support skills, given that pupils receive 4 weeks of training in how to work well within pairs and how to take responsibility within their pairs for conducting the sequence of PALS activities.

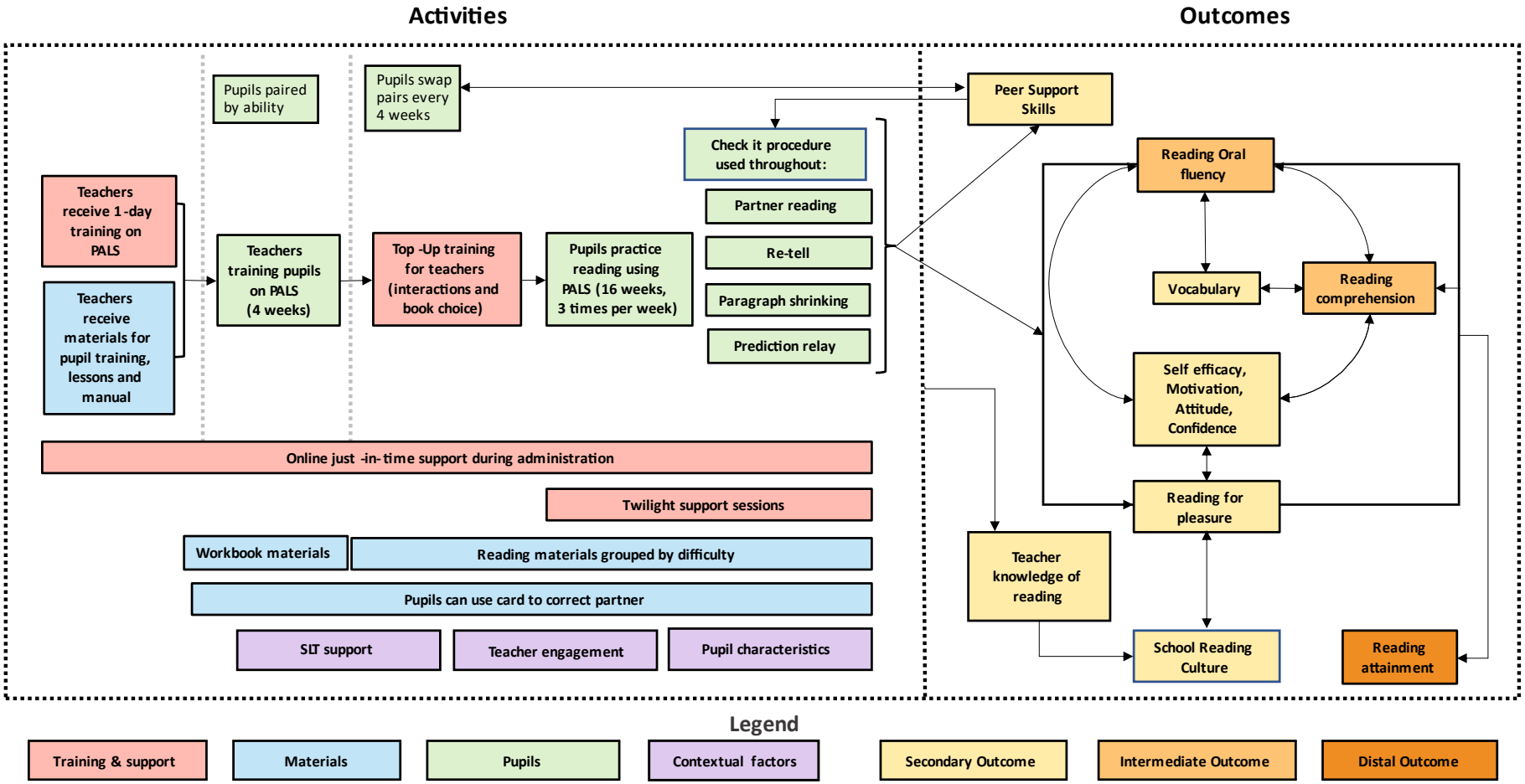
Each of the intermediate and secondary outcomes mentioned above are likely to drive gains in the more distal outcome of reading attainment given that oral reading fluency, reading comprehension, reading self-efficacy, motivation for reading and reading for pleasure have all be shown to be predictors of reading attainment (Carroll & Fox, 2017; Clark & Rumbold, 2006; Rasinski et al., 2005).

In order to ensure that the intervention is delivered as intended, teachers will receive initial training plus top-up training (at a key point in the intervention after the whole class training). They will also be provided with a manual, videos demonstrating good practice and common mistakes, materials and scripted lesson plans for pupil training, and a selection of books. The developers will also provide additional 'just-in-time' support throughout the intervention in the form of discussion boards, twilight webinars and the opportunity for teachers to contact the developers with any questions or problems at any time. The report from the previous evaluation suggests that the intervention can be delivered with excellent fidelity (e.g. 100% teachers completing the survey reported that they implemented the full 20 weeks of the intervention and 95% of teachers attended both training sessions).

There are a number of contextual factors that might affect implementation and impact: support from the senior leadership team, teacher engagement, fidelity and pupil characteristics. These are all factors that will be considered in the implementation and process evaluation.

Figure 1. PALS Logic Model

Reference Template



Impact evaluation

PRIMARY RESEARCH QUESTION

1. What is the difference in the average score for reading attainment among Year 5 pupils in schools exposed to PALS-UK, compared to Year 5 pupils in control schools exposed to business as usual conditions?

SECONDARY RESEARCH QUESTIONS

1. What is the difference in the average score for oral reading fluency (rate) among Year 5 pupils in schools exposed to PALS-UK, compared to Year 5 pupils in control schools exposed to business as usual conditions?
2. What is the difference in the average score for reading fluency (multi-dimensional) among Year 5 pupils in schools exposed to PALS-UK, compared to Year 5 pupils in control schools exposed to business as usual conditions?
3. What is the difference in the average score for reading comprehension among Year 5 pupils in schools exposed to PALS-UK, compared to Year 5 pupils in control schools exposed to business as usual conditions?
4. What is the difference in the average score for reading self-efficacy among Year 5 pupils in schools exposed to PALS-UK, compared to Year 5 pupils in control schools exposed to business as usual conditions?
5. What is the difference in the average score for motivation for reading among Year 5 pupils in schools exposed to PALS-UK, compared to Year 5 pupils in control schools exposed to business as usual conditions?

EXPLORATORY RESEARCH QUESTIONS

1. What is the difference in the average score for reading attainment among pupils who are entitled to Free School Meals (FSM) in schools exposed to PALS-UK, compared to the FSM pupils in control schools exposed to business as usual conditions?
2. What is the difference in the average score for reading attainment among pupils with special educational needs (SEND) who are in schools exposed to PALS-UK, compared to the pupils with SEND in control schools exposed to business as usual conditions?[2]
3. What is the difference in the average score for reading attainment among pupils scoring in the lowest quartile on the baseline New PIRA test in schools exposed to PALS-UK, compared to the pupils scoring in the lowest quartile on the baseline New PIRA test in control schools exposed to business as usual conditions?
4. What is the difference in the average score for reading attainment among pupils for whom English is another language (EAL) and whose score falls in the lower half of the sample distribution on the baseline New PIRA test in schools exposed to PALS-UK, compared to the same subgroup of Year 5 pupils in control schools exposed to business as usual conditions?

[2] Some pupils with SEND who - based on the school's judgement - are unable to complete the PiRA at baseline will be excluded from the trial prior to randomisation. Therefore, this research question only applies to those SEND pupils who stay in the trial.

Design

Table 1: Trial design

Trial design, including number of arms		Two-arm, stratified and cluster-randomised trial at the school level
Unit of randomisation		Schools
Stratification variables (if applicable)		School size (one-form per year group versus two or more forms per year group) Proportion of year group that are currently free school meals (split across the median sample proportion)
Primary outcome	variable	Reading attainment
	measure (instrument, scale, source)	Reading attainment (New PIRA Summer 5 Test) [3]
Secondary outcome(s)	variable(s)	Oral reading fluency (rate); Oral reading fluency (multi-dimensional); Reading comprehension; Reading self-efficacy; Motivation for reading
	measure(s) (instrument, scale, source)	WIAT-III UK-T[4]: reading comprehension and oral reading fluency subtest Multi-dimensional Fluency Scale[5] Feelings about reading questionnaire (measures reading self-efficacy and motivation for reading)[6]
Baseline for primary outcome	**variable**	Reading attainment
	measure (instrument, scale, source)	Reading attainment (New PIRA Summer 4 Test)
Baseline for secondary outcome	**variable**	Reading attainment; Reading self-efficacy; motivation for reading
	measure (instrument, scale, source)	Progress in reading attainment (New PiRA Summer 4 Test)[7] Reading Self-Efficacy Questionnaire

[3] https://www.risingstars-uk.com/series/assessment/rising-stars-pira-tests

[4]https://www.pearsonclinical.co.uk/store/ukassessments/en/Store/Professional-Assessments/Academic-Learning/Comprehensive/WIAT-III-UK-for-Teachers/p/P100009239.html

[5] See Rasinksi (2004) Available at https://files.eric.ed.gov/fulltext/ED483166.pdf

[6] Feelings about reading questionnaire; the first part, measuring reading-self-efficacy, is adapted from Carroll & Fox (2017). Available at https://doi.org/10.3389/fpsyg.2016.02056. The second part, measuring motivation for reading, adapted from the scale used in the previous trial, is pending publication (Vardy, Breadmore and Carroll, in prep).

[7] Scales for WIAT-III UK-T III and the Multi-dimensional Fluency Scale are not administered at baseline. Therefore, baseline scores for oral reading fluency (rate), reading fluency (multi-dimensional) and reading comprehension are derived from the baseline New PiRA assessment.

The impact evaluation is a two-group parallel cluster-randomised controlled trial, efficacy study, with schools allocated to intervention and control at random on a 1:1 basis. Randomisation will be stratified in order to achieve balance on key school level covariates. The proposed covariates are school size (one-form per year group, two or more forms per year group), and the proportion of year group that are free school meals.

The primary outcome will be a measure of reading attainment derived from the New PiRA reading test, delivered online at baseline. This is a digital assessment that will be administered to pupils as a whole class. The children will work through the on-screen reading test using available school devices (i.e. computers or tablets). Reading attainment will be measured using the New PiRA reading test both prior to randomisation and then post-exposure to PALS. Schools' experiences of online administration of New PiRA at baseline will be reviewed and if considered necessary, administration at post-test will be paper-based. If the school is selected to deliver PALS-UK, all Year 5 teachers at the school will receive the PALS-UK manual, training, resources and support needed to deliver the intervention. Intervention schools will deliver the intervention in all classes, and the full year-group cohort samples (Year 5) will be tested.

Secondary outcomes will be measures of reading fluency and comprehension obtained from the WIAT-III UK-T standardised assessment tool and the Multi-dimensional Fluency Scale, as well as measures of reading self-efficacy and motivation for reading, derived from the Feelings about reading questionnaire. The secondary analysis will involve estimation of effects on the reading self-efficacy and motivation reading outcomes for the full year-group cohort samples (Year 5). Effects on the outcomes measured by the WIAT-III UK-T and MDFS will be estimated for a randomly selected subset of 10 pupils in order to contain costs, given that the instrument needs to be delivered face-to-face by trained researchers.

Randomisation

By the end of the summer term 2022 we will have confirmed the identities of each participating school and pupils within these schools and we will have collected baseline reading assessments alongside reading self-efficacy and motivation for reading measures for each pupil. Prior to randomisation at the beginning of the Autumn term 2022, we will sample one class per school at random and within the chosen class, 10 pupils will be selected for additional data collection. These children will complete the WIAT-III UK-T reading comprehension and oral reading fluency subscales later on in the summer term of 2023. Only when these tasks are complete will randomisation occur and its outcome be communicated to the Delivery Team, who will in turn notify schools.

A cluster randomised trial is proposed with whole schools allocated to intervention and control groups. Randomisation will be stratified in order to achieve balance on key school level covariates. The proposed covariates are:

- School size: single-form entry versus multiple-form entry
- Proportion of year group that are currently eligible for free school meals (split across the median sample proportion)

Randomisation will proceed as follows:

- A random number seed will be determined and stored;
- Schools will be listed in descending order by URN within each stratum;
- Within each stratum schools will be allocated a random number from a uniform distribution to four decimal places;
- Schools will be re-ordered within each stratum on the basis of the random number they are assigned and on a descending basis;
- Within each stratum two groups of schools will be formed by splitting the ordered list of schools in half – the first group will be Group A the second group Group B;
- A coin toss will then determine which group is assigned to the intervention – if the coin toss is heads, Group A is intervention (and Group B control), if it is tails Group A is control (Group B intervention).

This process will be repeated within each stratum. Using STATAv17 statistical software, randomisation will be conducted by a Research Associate based in the Policy Evaluation and Research Unit who will be blinded to the identity of school at randomisation.

Participants

Schools

The focal cohorts are pupils entering Year 5 in September 2022 in primary schools recruited to the trial. Pupils within these cohorts will be identified during Year 4, prior to their entry into Year 5.

Schools that are state-funded primary schools located in following English school commissioner regions: The North, the East Midlands and Humberside, and the West Midlands are eligible to participate in the study, but with the following schools *excluded*:

- Schools taking part in another Accelerator Fund efficacy trial or another EEF literacy project targeting the cohort of pupils who will be in Year 5 in the 2022-23 academic year. This includes the following projects:
    - Learning Language and Loving It;
    - The 5Rs Approach to GCSE Maths Resits;
    - English Mastery;
    - Reciprocal Reading;
    - Children's University;
    - Teacher Choices;
    - Thinking Doing Talking Science.
- Schools that took part in the previous EEF-funded PALS-UK trial that were allocated to the intervention group.
- Schools that cannot provide access to technology for online assessment of reading.
- Schools that for whatever remaining reason are using or have used PALS-UK.

Schools will be approached to take part in the study by the Delivery Team. Once the school signs the study MoU and the Data Sharing Acknowledgement, the Delivery Team will collect some basic background information from the school. This information will be sent to MMU and FFT in electronic form and be used to generate an initial record in the trial data base.

Pupils

All year 4 pupils enrolled at the time of school recruitment are included in the study, except for those pupils with SEND who - based on the school's judgement - are unable to complete the PiRA at baseline. These pupils, together with those who join the schools after baseline assessment will not be included in the evaluation. Children joining schools in the treatment group at a later time will still take part in the intervention but they will not be assessed. Parents have the right to withdraw their child at any time throughout the study duration.

Sample size calculations

Table 2: Sample size calculations

		OVERALL	FSM
Minimum Detectable Effect Size (MDES)		0.214	0.246
Pre-test/ post-test correlations	level 1 (pupil)	0.7	0.7
	level 2 (class)	0	0
	level 3 (school)	0	0
Intracluster correlations (ICCs)	level 2 (class)	0.05	0.05
	level 3 (school)	0.10	0.10
Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		Two	Two
Average cluster size for level 1 (per level 2 unit)		20[8]	5[9]
Average cluster size for level 2 (per level 3 unit)		1.30[10]	1.30
Assumed attrition rate		10%	10%
Number of schools	Intervention	54	54
	Control	54	54
	Total	108	108
Number of pupils	Intervention	2093	523
	Control	2092	523
	Total	4185	1046

The Table above provides an assessment of statistical power. Minimum detectable effect sizes are calculated using the software PowerUp (Dong & Maynard, 2013). We assume whole schools are randomised 1:1 to intervention and control, and schools are stratified into four blocks on the basis of proportion of FSM students (split across the median sample proportion)

[8] We report the harmonic mean here to account for varying cluster size. Calculations are based on the average class sizes in English primary schools in 2020/21 as reported by the National Statistics (see Appendix 2).
[9] Based on the previous trial, we assume that 25% of the pupils will be eligible for free school meals.
[10] We report the harmonic mean here to account for varying cluster size. Calculations are based on the average number of classes in English primary schools in 2020/21 as reported by the National Statistics.

and form of entry. We aim to recruit 120 schools in the trial but will assume 10 per cent school-loss-to-follow-up (n=108). On average across the sample as a whole we would expect around 25 Y5 pupils per class and 1.55 classes per school[11].

The table above presents the Minimum Detectable Effect Sizes (MDES). These are the smallest true effects that will yield sample estimates with Types I and II long term error rates of 5 and 20 per cent respectively. Calculations were based on a proportion of variance explained in the outcome by covariates of 0.49[12]. It is assumed that the attrition rate is equal in the two arms of the trial. We use a three-level clustered design (pupils nested in classes nested in schools) and base our calculations on intra-cluster correlation ICC 0.10 at the school-level and 0.05 per cent at the class-level. We assume a low class-level ICC of 0.05 which is consistent with previous research that has a three-level design equivalent to that proposed here (Boylan et al., 2018; Jay et al., 2017). This is also in line with the widespread practice estimating class-level ICC as being half of what is found at the school-level within primary education (Demack, 2017). Typically, EEF studies with a two-level design assume an ICC at the school level of 0.20, which is conservative. We have reduced this due to clustering at the class level and because we have results from the previous PALS evaluation commissioned by EEF, which shows school level ICCs of 0.14 (Culora et al., 2022). Furthermore, a recently published EEF Research paper reports an ICC at the school-level of 0.10 at Key Stage 2 for both Maths and English subjects (Allen et al., 2018).

The assumed 10 percent attrition rate was based on previous studies reporting relatively low school-level attrition (e.g., Jay et al., 2017). The previous PALS evaluation (Culora et al., 2022) reported high levels of attrition at school-level, yet the reduced sample size was due to attrition owing to the COVID challenges that schools faced, which resulted in various issues, such as delayed outcome testing.

Using the parameters above and assuming a continuous, normally distributed outcome, and 10 percent attrition at school level, the MDES is 0.214 for the total sample. Using the same parameters and assuming that on average there are around 10 FSM eligible pupils in each school (and 6 pupils in each class), the MDES is 0.246.

Outcome measures

Primary outcome

The primary outcome is a measure of reading attainment derived from the New PiRA reading test. New PiRA is UK-standardised, and EEF-shortlisted with high test reliability (Cronbach's alpha above 0.9), face validity (it is written to follow the national curriculum guidelines) and concurrent validity, showing a strong relationship with national test scores. It can be administered to the whole class, minimising cost and school disruption. It is designed around termly expectations, providing appropriate forms at pre- and post- test (i.e. summer term year 4 and summer term year 5). The digital version of New PiRA, which will be used in this trial, offers a cost-effective option. The New PiRA instrument (New PIRA Test 4, Summer) will be administered at baseline to pupils at the end of the summer term 2022, when pupils are in

[11] Please note that MDES calculations are based on the harmonic mean and not the arithmetic mean reported here.

[12] Assumed pre-post correlation of 0.7 as per the power calculation table above, (based on KS1-KS2 correlation in the NPD, as reported by Allen et al., 2018). Thus, R2 will be 0.49.

Year 4. The assessments will be administered online by schools but the process overseen by MMU. Data collection will be complete by the end of July 2022.

The instrument will be administered by teachers and teaching assistants. Tests are scored automatically using the online tool. The test takes between 40 and 50 mins to complete.

New PiRA is structured around the following content domains (for KS2):

- Vocabulary - explain the meaning of words in context;
- Comprehension - retrieve and record information/identify key details;
- Summary - summarise main ideas from the text;
- Inference - make inferences from the text, explaining and justifying with evidence from the text;
- Prediction - predict what might happen;
- Structure - identify/explain how content is related and contributes to meaning;
- Impact - identify/explain how meaning is enhanced through choice of words and phrases;
- Comparison - make comparisons within the text.

The content domains of comprehension, summary, and prediction are likely to be particularly sensitive to PALS-UK given that retelling supports comprehension (Dunst et al., 2012), paragraph shrinking involves focusing on the key main ideas, which is a key part of summarising, and prediction relay activities are designed to support students in predicting upcoming events from what they have read (Fuchs et al., 1997). The other content domains measured by New PiRA, while not the explicit focus of the intervention, may also be improved by PALS-UK. For example, gains in fluency may free up cognitive resources for comprehension of inferences in texts (Klauda & Guthrie, 2008) and lead to gains in vocabulary knowledge (Yildirim et al., 2013).

A pupil can score a maximum of 40 marks and a minimum of zero on the Year 4 summer New PiRA reading test. A raw score from the baseline New PiRA reading test will be used as a covariate in the primary and secondary analysis that is discussed further below.

Follow-up data collection will take place toward the end of the Summer term 2023. All Year 5 students will be required to complete a New PiRA reading test and Feelings about reading questionnaire (which will also be administered at baseline) for a second time (the reading self-efficacy and motivation for reading outcomes, measures using the Feelings about reading questionnaire, and the timing of such observations are discussed further below). A subset of pupils will also be required to complete two subtests from the WIAT-III UK-T and the MDFS fluency measure; these measures are discussed further below.

At follow-up the New PiRA (New PIRA Test 5, Summer) will be administered online or using traditional paper-based materials by the evaluation team; a decision will be made about the mode of administration following a review of the baseline administration of New PiRA. Administration by the team rather than teachers and teaching assistants is preferred so that those responsible for administering the tests can be kept blind to the intervention/control status of the pupils and schools, something which would be impossible if assessments at follow-up were administered by teachers.

A student can score a maximum of 45 marks and a minimum of zero in the Summer Year 5 New PiRA assessment. The raw score from the follow-up New PiRA reading test will be used as the dependent variable in the primary analysis.

Secondary outcomes

Secondary outcomes include measures of (1) **oral reading fluency (rate)**; (2) **oral reading fluency (multi-dimensional)**; (3) **reading comprehension**; (4) **reading self-efficacy; (5) motivation for reading.**

Reading fluency (rate) and Comprehension

A fine-grained measure of fluency that moves beyond rate and accuracy is desirable, given that the different dimensions of fluency have been shown to have independent effects on reading comprehension (Klauda & Guthrie, 2008). Therefore, we will be using two measures of fluency. Measuring *Fluency (rate)* will provide us with a basic measure of fluency operationalised as average number of words read correctly per minute, whilst measuring *Fluency (multi-dimensional)* provides a qualitative assessment of fluency over the dimensions of: Expression and Volume; Phrasing; Smoothness; and Pace.

The outcome of Fluency (rate) will be based on the oral fluency subtest of the Wechsler Individual Achievement Test – III for Teachers (WIAT-III UK-T). As mentioned, the test provides a measure of the average number of words read correctly per minute. Two passages are read by the pupil. A score is obtained through taking the total word count for the two passages minus the errors made, divided by the time taken to read the passages. This quantity, a rate per second, is then multiplied by 60 to convert the measure into seconds. The total score for pupils aged 9 and 10 years can range between 2 and 3,000.

Furthermore, the reading comprehension subtest of the same tool will be used to measure the secondary outcome of Comprehension. The subtest provides a score based on responses to a range of literal and inferential comprehension questions.

Administration times are approximately 25–40 minutes for the full assessment. According to EEF's test database, subtests should rarely take more than 10 minutes to administer with the exception of the reading comprehension subtest which may take older participants longer. It is estimated that administration of the two WIAT-III UK-T subtests will take 30 minutes per pupil. Unlike New PiRA, WIAT-III UK-T will not be administered online but instead be administered face to face, in schools, to the 10 pupils per school sampled for the assessment.

Reading fluency (multi-dimensional)

The freely-available, Multi-dimensional Fluency Scale (MDFS; Zutell & Rasinski 1991; Rasinski, 2004), produces scores ranging from 4 to 16 and provides a complementary measure of fluency by providing a qualitative assessment of the outcome over the dimensions of: Expression and Volume; Phrasing; Smoothness; and Pace. MDFS will be used to provide additional information about fluency (expression and volume, phrasing and smoothness), as the pupils read the WIAT-III UK-T texts. The MDFS has been shown to be a reliable and valid measure of fluency (Paige et al., 2014).

Administrators of the test will be trained in how to assess fluency using the MDFS.

Assessors need to have been trained for valid administration of the WIAT-III UK-T and MDFS. This raises the costs of data collection. To reduce costs, we propose collecting oral reading

fluency (in conjunction with the MDFS) and comprehension measures at post-test only and from a subset of pupils in participating schools.

Reading self-efficacy and Motivation for reading

The final secondary outcome measures are reading self-efficacy and motivation for reading. These outcomes will be measured together using an adapted form of the Feelings about reading questionnaire. This questionnaire will be administered to the whole class, taking around ten minutes to complete. The first part of the questionnaire, measuring reading self-efficacy has 20 items. The second part of the questionnaire measuring motivation for reading has 10 items. This two-part questionnaire was used in the previous PALS-UK trial, although only the reading self-efficacy data was analysed. Subsequent analyses conducted by the delivery team suggest that motivation for reading and reading self-efficacy contribute to unique variance in reading attainment (Vardy, Breadmore and Carroll, under review), therefore both scales will be analysed within the current trial. The questionnaire has a Likert scale structure (7 points) and possible score range from 30 to 210. In the previous trial the motivation for reading scale used a 4 point scale, but analysis of the data suggested that better sensitivity might be gained through using a 7 point scale for both parts of the questionnaire (Vardy et al., in prep). The motivation to read scale developed by Vardy et al. (in prep) is underpinned by self-determination theory (Deci and Ryan, 1985) and has been shown to have high reliability (Cronbach's α = .83). The reading self-efficacy scale is based on self-efficacy theory (Bandura, 1993) and is adapted from Carroll and Fox's (2017) original version of the scale with minor revisions to the phrasing of a few items and an additional item added to more directly link to the PALS-UK intervention. This has a Cronbach's α value of .90 (Vardy et al., in prep). This instrument will be administered online at both baseline and follow-up to all pupils together with the New PiRA. Similarly to the New PiRA, schools will be responsible for baseline administration, and the evaluation team will administer at follow-up.

Compliance

Compliance is defined at the school level, based on completion of programme activities, as recorded by the Delivery Team. Therefore, once an intervention school is said to have complied, we cannot rule out pupils in that school benefiting in some way from the intervention, even if there is limited evidence that the pupil has directly engaged with PALS.

Other definitions of compliance might be considered, such as partial completion of PALS or partial or incomplete delivery of the intervention to some students. However, to recover a valid estimate of the complier average causal effects of PALS we have to be assured that those deemed non-compliant by virtue of our definition of compliance cannot in any way have benefited from the intervention. If those deemed non-compliant by our definition, gain some form of benefit from any of the activities of the developers, our estimate of complier average causal effects based on such a definition will be biased. The definition of compliance discussed here attempts to ensure that those in our sample deemed non-compliant cannot have gained any benefit from the activities of the PALS team.

The PALS-UK Delivery Team have provided criteria which should inform a measure of compliance for schools allocated to the intervention (Table 3). We therefore propose an indicator of compliance that will take the value '1' if:

1. At least one teacher from an intervention school attends the initial training event[13]; and
2. There is evidence from that teacher's school that one or more pupils completed all four weeks of training.

In all other cases the indicator will take the value zero.

In summary, both minimal pupil and teacher training must have taken place for a school to be compliant and the assumption is that without both forms of training, the intervention (however it emerges or materialises) cannot lead to changes in pupils' reading. In other words, minimal patterns in training receipt, as set out in the bullets above, generates an experimental contrast – that is something that is distinct when compared to business as usual conditions prevailing in the control group.

Table 3. Compliance criteria

Compliance criterion	Data source (collected by)	Compliance indicator
Attendance at PALS-UK training	Registers of attendance (Delivery Team)	At least 1 teacher per school attends 1 day initial training
Completion of the four weeks of pupil training in line with the manual	Monitoring logs (Delivery Team) RA observations (Delivery Team) Top-up training (Delivery Team)	Evidence that pupils completed all four weeks of training in line with manual. Confirmed by data collected for each class from any of three sources: • Online monitoring logs • RA observations • Confirmation during attendance at top-up training

Analysis

Primary analysis

The intervention schools will deliver the intervention in all classes, and the full year-group cohort samples (Year 5) will be tested at follow-up. Recruiting multi-form entry schools (with multiple Y5 classes) as well as single form entry schools means that pupils will be nested/clustered in classes and classes nested/clustered in schools. Sample estimates of the causal effect of PALS-UK on the average New PiRA score of pupils at follow-up will be obtained from a multi-level linear regression model taking the following, three-level, form:

$$Y_{ijk} = \beta_0 + \beta_1 T_k + \beta_2 X_{ijk} + \beta_3 Z_{ijk} + \beta_4 S_k + w_k + u_{jk} + e_{ijk} \ldots\ldots [1]$$

Here, Y_{ijk} represents a continuous response, in our case the pupil's reading test score, for pupil 'i' in class 'j' and school 'k'. T_k is an indicator variable taking the value one if school 'k' is an intervention school and zero otherwise. X_{ijk} is a continuous measure of reading attainment obtained from the baseline New PiRA assessment for pupil 'i' in class 'j' and school 'k'. Z_{ijk} is a measure of a pupil's month of birth obtained from the baseline demographic data for pupil 'i' in class 'j' and school 'k', and S_k is a collection of school-level stratum variables. w_k is a school-level random effect, u_{jk} is a class-level random effect and ε_{ijk} a pupil level residual term. The sample estimate of β_1 is the intention to treat estimate of the effect of PALS on the outcome.

[13] Note it is not possible for teachers to attend the top-up training event unless they have first attended the initial training. It is also not possible for pupils to receive training in PALS unless their teacher or at least one teacher in the school has attended initial training.

The school level random effect is assumed to be distributed normally in the population with zero mean and variance θ^2, the class level random effect similarly with variance τ^2. If the variance of e_{ijk} is σ^2, then the two intraclass correlation coefficients at the school and class levels are:

$$ICC_k = \frac{\theta^2}{\theta^2 + \tau^2 + \sigma^2}$$

$$ICC_j = \frac{\tau^2}{\theta^2 + \tau^2 + \sigma^2}$$

Similar trials (e.g. Gorard et al., 2017; Humphrey et al., 2020; O'Here et al., 2019; Rudd et al., 2017) previously funded by EEF have typically ignored clustering at the class level. This seems in part because empirical evidence suggests that class-level ICCs are typically very low for this year group[14]. Our decision not to ignore class level clustering is informed by EEF research which argues that failing to account for class-level clustering can lead to difficulties (Demack, 2019).

Sensitivity testing for the primary analysis

For the primary outcomes derived from the New PiRA test, three further analyses will be performed. These analyses aim to sensitivity test key assumptions underpinning the chosen primary analysis described above.

The first form of sensitivity analysis for the primary outcome involves a reduced regression model that takes the form of equation [1] above but with the pupil baseline measure of reading attainment excluded:

$$Y_{ijk} = \beta_0 + \beta_1 T_k + \beta_2 Z_{ijk} + \beta_3 S_k + w_k + u_{jk} + e_{ijk} \ldots\ldots [2]$$

This specification permits us to assess the extent to which inclusion of the baseline PiRA test score as a covariate influences the precision of the sample estimates. The second form of sensitivity analysis will mirror the regression model used for the primary analysis set out at equation [1] but with the age standardised PiRA score obtained at follow-up as the dependent variable instead of the raw score. This specification will omit the month of birth covariate previously included but will otherwise remain as equation [1]. This second specification will enable us to assess how far age-standardisation may influence results.

A final sensitivity test would run two separate analyses on that part of the 'as analysed sample' that has a three-level structure of pupils nested within classes and classes within schools. The first analyses would involve fitting a three-level model as equation [1] on this subset of the trial sample. The second analysis would fit the same model but allowing only for a two-level structure without clustering of students within classes. Results from these two analyses can be compared to inform judgements about whether including the class level in that part of the sample in which the Year 5 cohort is multi-form has any substantive consequences for the results of the primary analysis.

[14] The Dialogic Teaching evaluation conducted by Jay et al. (2017) that used GL Progress Test in English, Maths and Science as outcomes and KS1 test score as the baseline covariate for Y5 pupils reports class-level ICCs between 0.01 and 0.04. Another EEF trial, the ScratchMaths evaluation (Boylan et al., 2018) involving all Y5 and Y6 pupils had similarly low ICC scores (0.01 and 0.02).

Hypothesis tests for the treatment effects in each specification will be reported in the form of p-values and frequentist 95% confidence intervals. Regression estimates for treatment effects will be converted to effect sizes consistent with Hedges' g, as discussed below.

Secondary analysis

The secondary analysis will involve estimation of effects on the reading self-efficacy and motivation for reading outcomes for the full year-group cohort samples (Year 5). It also involves estimating the effects on the MDFS outcome and the WIAT-III UK-T outcomes for the subset of 10 pupils per school selected at random for more extensive testing.

Sample estimates of average causal effect for reading self-efficacy, motivation for reading, oral fluency and comprehension will be obtained from fitting regression models to the relevant data consistent with the Equation [1] above, and using the same statistical procedures, where the dependent variables 'Y' is derived from the relevant scales at follow-up. As WIAT-III UK-T III and the MDFS are not administered at baseline, the covariate 'X' is derived instead from the baseline New PiRA assessment.

Hypothesis tests for the treatment effects in each specification will be reported in the form of p-values and 95% confidence intervals. For the secondary analysis, treatment effect estimates will be reported as effect sizes (Hedges' g).

Subgroup and exploratory analysis

For three subgroups we will conduct separate analyses. The subgroups of interest will be pupils entering Year 5 in participating schools that are: 1) ever-FSM (using the variable EVERFSM_6), 2) designated SEND and 3) pupils scoring in the lowest quartile on the baseline New PIRA test. Differential effects are to be explored for each subgroup by including an interaction term in equation [1] above, comprising the relevant subgroup indicator interacted with the treatment dummy indicator.

In addition, further exploratory analysis will examine the effects of PALS-UK for EAL pupils. To do this, it is proposed that an indicator is created that combines the NPD-type measure of EAL with a pupil's score on the baseline reading assessment, where that score falls in the lower half of the sample distribution. In other words, a binary indicator is created at the pupil level and takes the value '1' if a pupil is EAL and that same pupil's score on the baseline reading tests falls below the median score for the sample. By constructing this variable, we aim to remove higher-performing EAL students from our assessment and focus instead on those EAL students that fall in the bottom half of the reading assessment score distribution. The decision to construct such an indicator, in the manner described, stems from concerns that the NPD-type EAL measure alone captures a very heterogeneous group of pupils, and that these pupils have widely varying levels of language proficiency. A binary indicator capturing membership of this group will be interacted with the treatment dummy with the resulting estimates examined.

Effect size

In order to obtained an effect size measure from the primary outcome regression model, we will divide the sample estimate of β_1, as specified in the model[15], which is the regression coefficient on the treatment dummy variable, by the unconditional pooled standard deviation, and then repeat this transformation using the lower and upper limits of the confidence interval

[15] $Y_{ijk} = \beta_0 + \beta_1 T_k + \beta_2 X_{ijk} + \beta_3 Z_{ijk} + \beta_4 S_k + w_k + u_{jk} + e_{ijk}$

from the primary regression analysis to derive the relevant confidence interval for the effect size. All effect sizes will be consistent with Hedges' g.

This approach will be used to calculate effect sizes for both primary and secondary outcomes.

Implementation and process evaluation

The IPE will gather data in relation to the PALS-UK manual and resources, teacher training, pupil training, the delivery of the intervention and teacher support from the Delivery Team. The logic model developed for the first trial has been reviewed and revised, and underpins the design of the IPE. Quantitative and qualitative data will be gathered through pre- and post-intervention surveys of headteachers and teachers, together with case studies of two schools per region (six in total), which will be visited on two occasions. The IPE will also review the documents provided to teachers to support delivery of PALS-UK and the data gathered by the delivery team via teacher surveys and other activities. Another key data source will be the observation logs completed by research assistants and school nominated peer reviewers which will inform understanding of the fidelity of delivery across all intervention schools. Finally, the delivery team will be interviewed.

Research questions

1. How was PALS-UK delivered and supported?
    - 1.1 To what extent did fidelity vary and why?
    - 1.2 What are the enablers of and barriers to success with respect to training and support, and the individual components of PALS-UK?
    - 1.3 How do different stakeholder groups (e.g. students, teachers) experience PALS-UK and what is the perceived impact?
    - 1.4 What contextual factors contribute to or inhibit PALS-UK effectiveness?
    - 1.5 Were there any unexpected outcomes?
2. What comparable initiatives were undertaken within control group schools and did they change over time?

Research methods

The primary forms of data collection will be surveys, interviews, observations and gathering relevant documents and resources. Survey and interview questions will take account of questions asked in the first trial to enable comparisons to be made if deemed appropriate. Questions will be developed in consultation with the Delivery Team and draw on in-house literacy specialists. Survey questions which do not relate directly to PALS-UK will be piloted with 2-3 headteachers or teachers (as appropriate) to ensure that they are clear and elicit the anticipated responses. Surveys will include a mixture of closed and open questions.

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 reading interventions delivered in the school

- planned changes to the way in which reading is taught for 2022/23 (in addition to PALS-UK if allocated to the intervention arm);
- future staff CPD relating to teaching reading (in addition to PALS-UK if allocated to the intervention arm);
- the school reading culture.

The survey will be administered in September 2022. The survey will be administered after randomisation due to the need to inform schools of the outcome of this process as soon as possible to inform their literacy planning. It is not considered appropriate to administer the survey during the summer term due school recruitment potentially continuing until one week before the end of the summer term and the demands of other baseline data collection.

All teachers who will be teaching Year 5 classes (or classes including Year 5 pupils) in September 2022 will be asked to complete a survey in September (no longer than 15-20 minutes). Teachers acting as a peer reviewer from single form entry schools will also be asked to complete the survey. The teacher survey will elicit data about current literacy practices and knowledge including:

- current literacy practices including other reading interventions delivered in the school;
- estimated pupil time spent on reading instruction and reading practice each week;
- formal and informal CPD relating to teaching reading undertaken in 2021/22;
- school reading culture;
- knowledge of the process of learning to read.

The survey will be administered in September 2022, after randomisation due to the need to inform schools of the outcome of randomisation as early as possible to assist in their planning requirements.

Training and support

Training and support is key to ensuring the fidelity of the implementation. We will observe 2 of the initial training events and observe 2 of the top-up training events. The key source of data will be structured field notes. The documentation (e.g. manual) and support materials (e.g. videos) provided to teachers will be reviewed. Attendance registers from the training, top-up training and other support activities will be analysed. The Delivery Team will also conduct short surveys and activities in relation to training and support in order to inform their activities. These data will be shared with the evaluation team and will provide an insight into teacher perceptions of the training and support.

Delivery

As PALS-UK is a manualised intervention, fidelity is a key implementation dimension for this evaluation. This will be evaluated in two ways. Firstly, as part of the delivery of PALS-UK four structured observations will be undertaken of all teachers in intervention schools. The structured observation will focus on the classroom set-up and the delivery of the four pupil activities (partner reading, re-tell, paragraph shrinking, prediction relay). The observers will complete a checklist developed by the Delivery Team. The Delivery Team will oversee the data collection which will be undertaken by both research assistants and intervention teachers' colleagues (peer reviewers). A research assistant will undertake 2 separate observations in all schools delivering the intervention during weeks 1-4 (the pupil training period) and weeks 10-14. Each teacher delivering the intervention will also be trained as a peer reviewer and undertake 2 further structured observations during weeks 5-9 and weeks 15-20. In single form

entry schools, another Year group teacher or teaching assistant will be invited to the initial training session to act as a peer reviewer for the Year 5 teacher.

Six intervention schools, two from each region, will be recruited as case study schools. To keep the selection process as simple as possible and given the tight timeline, all intervention schools will be invited to express their interest in participating in this additional layer of data collection. It will be made clear to all intervention schools what will be involved and the potential advantages of participating. Schools will then be selected from the group that put themselves forward so that they are representative in terms of school size, urban/rural location, and headteacher commitment (from the baseline survey). There is a potential for a slight bias towards schools which are engaging to a greater degree with the intervention as schools that are less engaged are less likely to put themselves forward. However, recruiting more engaged schools will generate the richest data and will provide a deeper understanding of the participants' experiences and impact of the implementation.

Two half-day visits to each case study school will be undertaken. The first will be mid-way through the intervention (January-February 2023) and the second will be towards the end of the intervention (March-April 2023). During each visit the following activities will be undertaken:

- observation of PALS-UK (35 minutes);
- 20-minute teacher interview ;
- 20-minute headteacher/senior leader interview;
- 20-minute student focus groups (6-8 students, a different group each time);
- 20-minute peer observer interview (single form entry schools, second visit only).

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. The teacher will be randomly selected and the group of students will be from the selected teacher's class. Teachers will be asked to identify a balanced group based on reading ability (first and second readers) and gender. As far as possible, students who have worked together in a pair will be selected. However, this will rely on parents' consent as well as students' assent to participate in the focus groups. A different group of students will be selected for the focus group at each visit so that opinions can be elicited from as many as possible. The peer observer will only be interviewed on the second visit so that both peer observations will have taken place.

Headteacher/senior leader interviews will focus on:

- how PALS-UK was integrated with current practices;
- perceived impact on:
  - teaching reading;
  - reading culture in the school;
  - staff development and knowledge of teaching reading;
- 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;
- the pairing process;
- perceived enablers and barriers;
- how PALS-UK was integrated with current practices and its relationship to guided reading (i.e. did it replace guided reading);
- perceived usefulness of elements of PALS-UK;
- required adaptations and identified gaps (what else did they cover outside PALS-UK).

Teacher interviews at the second visit will focus on:

- any updates to questions asked at the first interview;
- perceived impact on:
    - teaching reading;
    - reading culture in the school;
    - staff development and knowledge of teaching reading;
    - pupils' reading comprehension, fluency and attainment.
- planned future use;
- recommendations for other teachers.

Student focus groups will focus on:

- their experiences of PALS-UK, what they like and don't like about it;
- how it compares to other ways of learning to read that they have experienced;
- perceived impact on:
    - development of peer-support skills;
    - their ability to read;
    - attitudes to reading;
    - reading habits (including reading for pleasure);
    - school reading culture.

Post-delivery

Post-intervention surveys will be administered to headteachers and teachers in June/July 2023.

At the end of the intervention, 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 PALS-UK on the school and the costs involved:

- commitment to the trial;
- current literacy practices including other reading interventions delivered in the school;
- aside from PALS-UK what changes in literacy practices took place in 2022/23 (planned and unplanned);
- CPD offered during 2022/23 in relation to teaching reading (in addition to PALS-UK);
- perceived impact on
    - the school reading culture;
    - staff knowledge of teaching reading;
    - pupils' reading attainment;
    - pupils' preparedness for KS2 Reading SAT;

    - future literacy practices;
- costs associated with delivering PALS-UK.

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:

- commitment to the trial;
- current literacy practices including other reading interventions delivered in the school;
- what changes in literacy practices took place in 2022/23 (planned and unplanned);
- CPD offered during 2022/23 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). Teachers acting as a peer reviewer from single form entry schools will also be asked to complete the survey. The survey is designed to gather teachers' perceptions about the delivery of PALS-UK. The teacher survey will elicit data about:

- current literacy practices including other reading interventions delivered in the school;
- estimated pupil time spent on reading instruction and reading practice each week;
- experiences of PALS-UK training (initial and top-up) and resources;
- aside from PALS-UK training, formal and informal CPD relating to teaching reading undertaken in 2022/23;
- dosage (delivery of 4 weeks x 3 sessions pupil training, delivery of 16 weeks x 3 sessions of full PALS-UK sessions);
- the pairing process;
- adaptations made;
- identified gaps in the programme and how they were addressed;
- perceived enablers, challenges and benefits of PALS-UK and specifically SLT support;
- impact of PALS-UK on:
    - school reading culture;
    - their knowledge of the process of learning to read;
    - how they teach reading;
    - their pupils' reading;
    - their pupils' reading for pleasure;
    - different student groups;
    - preparation for KS2 SATs.

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 instruction and reading practice each week;
- formal and informal CPD relating to teaching reading undertaken in 2022/23;

The delivery team will also be interviewed to elicit their perceptions on the delivery of PALS-UK, the training and support offered, any adaptions made, barriers and challenges and how they were addressed, and the impact of PALS-UK on schools, teachers and pupils.

Implementation dimensions

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

Business as usual will be verified through the headteacher and teacher surveys administered at baseline to all schools in September 2022 and administered at post-test to control schools. Post-test surveys will ask about current practices which can be compared with baseline responses, and also about any specific changes that have taken place during the academic year. This offers a form of triangulation.

Fidelity will be assessed through data collected about the delivery of training and support (e.g. independent observations of training events) and from the PALS-UK observation checklist used in the four structured observations (focusing on classroom set-up and the four pupil activities). Intervention teachers will also be asked about any adaptations made to the programme in the post-test survey. Case study interviews will also elicit data in relation to fidelity given that it is believed to be key to successful delivery. Dosage is a key element relating to fidelity. In consultation with the delivery team we will ask teachers to self-report this in the post-test survey. We will also ask intervention teachers to complete light touch weekly logs.

The support of senior leaders was considered to be an important success factor in the previous efficacy trial. Therefore, responsiveness (the degree to which participants engage with the intervention) is also another key implementation dimension and will be evaluated through surveys and the case studies. Pupil responsiveness will be evaluated through observations of PALS-UK lessons and pupil focus groups.

Issues of bias will be addressed through the following means. Data will be collected from multiple sources (e.g. surveys, interviews, observations) and from different stakeholders. Structured observations, structured interview questions and survey questions will ensure that data collection is rigorous. Surveys will be administered online and reminders will be sent at least twice to minimise potential bias from non-responders. Qualitative data will be coded and analysed thematically, ensuring that a consistent approach is adopted by all those involved. If more schools volunteer to be a case study than are required, then schools will be selected according to specified criteria to ensure that schools reflecting different circumstances are represented. Pupils invited to participate in focus groups will be selected by their teacher who will be advised to identify pupils who are representative in terms of criteria such as gender and ability.

Analysis

Quantitative data from closed questions in the surveys will be analysed with SPSS generating descriptive statistics, cross-tabulations and graphs to enable changes over time to be illustrated and comparisons between intervention and control schools to be made. This will enable data from 120 participating schools to be summarised and presented visually including an overview of other reading programmes in use, the use of guided reading in schools,

perceived barriers and enablers to PALS-UK (at post-test), teachers' knowledge of reading and their perceptions of their own practice in relation to teaching reading. The patterns identified in the synthesis of quantitative data will be interpreted and discussed in relation to the IPE research questions and the logic model.

Qualitative data from open questions in the surveys, observations, interviews, focus groups will be analysed using NVivo and thematic analysis (Braun & Clarke, 2006; Braun & Clarke, 2021) with a mixed coding method. Interview data will be transcribed by a professional transcription service. Pure verbatim transcriptions (including pauses, stutters etc) are more expensive and not considered necessary for the IPE. A coding framework derived from the logic model (e.g. SLT support, teacher engagement, reading for pleasure, school reading culture) will be applied deductively and additional themes derived inductively, allowing for unexpected mediators and outcomes to be identified. Following Braun and Clarke's (2022) guidance on reflexive thematic analysis, data familiarisation will take place first, followed by systematic coding, and the development and review of the themes arising, repeating steps as considered to be necessary. The process of selecting codes and themes, which necessarily involves interpretation, will enable the IPE research questions to be answered through identifying patterns in the data reflecting participants' experiences and perceptions of the intervention (Kiger & Varpio, 2020). The approach to the analysis and interpretation will be primarily deductive, underpinned by the logic model. For example, the dataset will be analysed in relation to participants' perceptions of the impact of SLT support (or lack of support) on the delivery of PALS-UK, considering how this manifests in practice and how potential pitfalls could be avoided. By following Braun and Clarke's (2022) guidance, the analysis will go beyond one of the key criticisms of thematic analysis, simply summarising the data, to providing a nuanced interpretation which unpicks the significance of the data in relation to the IPE research questions.

These data will ensure that the research questions can be addressed drawing on evidence of the experiences of intervention and control schools (business as usual). Findings from the IPE will aid interpretation of the impact analyses; providing the opportunity to develop further hypotheses around possible mediators and sources of heterogeneity in treatment effects.

It will be particularly important to explore aspects of the logic model which are not tested explicitly in the impact analyses, in relation to SLT support, teacher engagement, peer support skills, pupil attitude and confidence, pupils' engagement in reading for pleasure, school reading culture and teachers' knowledge of reading. Both quantitative and qualitative data from the IPE will be used to do this.

Using structured approaches such as statistical and thematic analysis will ensure that the analyses are rigorous. Quality assurance procedures will be put in place to ensure that the findings from IPE analyses are reliable. This will include providing training to junior team members if required, documenting the procedures to be followed in analysis, keeping clear records, and regular meetings for those involved in IPE analyses to ensure that issues are identified promptly and resolved.

Table 4: IPE methods overview *(adapt as necessary)*

Implementation/ logic model relevance	Data collection methods	Participants/ data sources (type, number)	Data analysis methods	Research questions addressed
Usual practice	Survey: pre-intervention	All headteachers (120)	Descriptive Cross-tabulations Mixed coding Thematic analysis	RQ1.4 RQ2
Usual practice	Survey: pre-intervention	All teachers (240+)	Descriptive Cross-tabulations Mixed coding Thematic analysis	RQ1.4 RQ2
Fidelity Compliance Quality Responsiveness Logic model	Observation field notes of initial training (2 sessions) and top-up training (2 sessions)	Delivery Team (2), All teachers (intervention only, 120+)	Mixed coding Thematic analysis	RQ1.1 RQ1.3
Quality Logic model	Delivery documentation, surveys and activities designed to support delivery of PALS-UK	Documents provided to schools relating to PALS-UK; All teachers (intervention only, 120+)	Mixed coding Thematic analysis	RQ1.2
Fidelity Quality Responsiveness Logic model	Observation field notes of PALS-UK session	Case study class / field notes (12 visits)	Mixed coding Thematic analysis	RQ1.1 RQ1.2
Fidelity Dosage Quality Compliance Adaptations	Observation checklists of 4 x PALS-UK sessions (2 conducted by researcher, 2 conducted by peer review teacher)	All teachers (intervention only, 120+)	Descriptive Cross-tabulations	RQ1.1
Fidelity Quality Dosage Responsiveness Adaptations	Interviews with peer observers in case study schools (1 visit)	Sample of teachers (6)	Mixed coding Thematic analysis	RQ1.1
Fidelity Quality Dosage Responsiveness Adaptations Logic model	Interviews with teachers and headteachers in case study schools (2 visits)	Sample of teachers (6); Sample of headteachers (6)	Mixed coding Thematic analysis	RQ1.1 RQ1.2 RQ1.3 RQ1.4 RQ1.5
Fidelity Logic model	Interviews with a group of pupils in case study schools (2 visits)	Case study school pupil sample (96)	Mixed coding Thematic analysis	RQ1.1 RQ1.2 RQ1.3 RQ1.4 RQ1.5

Cost **Responsiveness** **Logic model**	Survey: post-intervention	All headteachers (intervention only, 60)	Descriptive Cross-tabulations Mixed coding Thematic analysis	RQ1.2 RQ1.4 RQ1.5
Usual practice	Survey: post-intervention	All headteachers (control only, 60)	Descriptive Cross-tabulations Mixed coding Thematic analysis	RQ2
Fidelity **Dosage** **Quality** **Responsiveness** **Adaptations** **Logic model**	Survey: post-intervention	All teachers (intervention only, 120+)	Descriptive Cross-tabulations Mixed coding Thematic analysis	RQ1.1 RQ1.2 RQ1.3 RQ1.4 RQ1.5 RQ2
Usual practice	Survey: post-intervention	All teachers (control only, 120+)	Descriptive Cross-tabulations Mixed coding Thematic analysis	RQ2

Cost evaluation

The overriding aim of the cost evaluation will be to ascertain the cost of the resources needed to deliver the intervention during the trial. From this we derive the following research questions:

1. What are the estimated delivery costs of the PALS UK trial per school?
2. What are the estimated delivery costs of the PALS UK trial per pupil?
3. What would be the estimated cost per school and per pupil of implementing PALS over three years?

As such the cost-evaluation will take the form of a Cost Feasibility analysis, representing a guide to the affordability of PALS, rather than a comparison between PALS and an alternative intervention. We provisionally expect the ingredients to be categorised as follows:

- Programme fees—for school access to training and materials based on the market value
- Prerequisite costs
- Staff time for teacher training, preparation, and delivery of PALS-UK, identifying separately the cost of new hires, supply staff
- Any additional (unpaid) staff time supporting the delivery of PALS-UK as reported by headteachers

Costs will be divided into pre-requisites, start-up costs and recurring costs as recommended by the EEF's cost evaluation guidance (EEF 2019). To calculate programme fees, the analysis

will rely on information provided by the Delivery Team. We will further collect cost data through the post-intervention headteacher and teacher surveys, and case study interviews with headteachers and teachers in the implementation and process evaluation.

Ethics and registration

Ethical approval has been obtained through Manchester Metropolitan University. The original submission was made on January 14th 2022 through a fast track route and approval was granted following revisions on February 1st 2022. We originally thought that the Delivery Team (from Nottingham Trent University and the University of Birmingham) would need to submit the documents for ethical approval at their institutions. However, both institutions accepted the documentation provided by Manchester Metropolitan University. Amendments to consent forms and participant information sheets were subsequently required to ensure that participants understood that only Manchester Metropolitan University undertook a full ethical review. An amendment was submitted on February 10th 2022 and approval granted on the same day. FFT Education, a sub-contractor working for Manchester Metropolitan University, requested that further information be added to the Memorandum of Understanding. This amendment was submitted on February 25th 2022 and approved on the same day. Finally, on 9th March we submitted two additional documents. It was not possible to submit these with the original application due to the need to obtain ethical approval before recruiting schools and the tight timeline we were working to. These documents outlined the content of a video to share with pupils and a teacher guidance sheet. All participating teachers were asked to at least share the video or go through the statements on the guidance sheet with their pupils. Our preference was that they do both. The video and teacher guidance is a more accessible format for presenting the information normally included in the participant information sheet to pupils.

The process for ethical approval includes providing details about the project design, information about the ethical procedures that will be adopted, and copies of participant information sheets and consent/withdrawal forms. We also included the Memorandum of Understanding and privacy notices.

The school recruitment process is as follows. The Delivery Team will identify and approach schools which meet the selection criteria in the three regions, and collect initial data including the school name, address, telephone number and URN, and the names and contact details of Year 5 teachers in 2022/23. Schools will be asked to sign a Memorandum of Understanding which provides information about the project and its aims, potential benefits for participating schools, a timetable of activities, data protection issues and responsibilities of all parties involved. In addition they are required to sign a separate Data Sharing Acknowledgement which outlines how personal data will be collected and shared between the Delivery Team, the Evaluation Team and the School. Schools will issue a withdrawal notice to all parents of students in Year 4. Parents will have 2 weeks to respond to this although they have the right to withdraw their child at any time. FFT will then collect baseline data from each school.

This trial is registered at the ISRCTN registry, registration number <to be provided>. The entry can be viewed here: <to be provided>.

Data protection

Manchester Met will process the personal data of pupils and school staff for the purposes of this study and will act as evaluators. This processing is regulated by the General Data Protection Regulation (GDPR) and the Data Protection Act 2018 (DPA). This Education Endowment Foundation (EEF) project is part of a wider Department for Education (DfE) funded programme called the 'Accelerator Fund'. The DfE and EEF are joint data controllers who have overarching responsibility for the programme.

- The Delivery team lead, Nottingham Trent University, is an independent Data Controller in respect of any personal data of pupils/and or school staff which they process for the purposes of the project;
- The Delivery team partner, University of Birmingham, is a Data Processor for the Delivery team lead;
- The Evaluation Team, Manchester Metropolitan University (Manchester Met) is an independent Data Controller in respect of any personal data of pupils/and or school staff which they process for the purposes of the project.
- The Education Endowment Foundation (EEF) becomes the Data Controller at the end of the project once the data is submitted to the EEF Data Archive, currently managed by FFT Education (Data Processor for the archive).

Manchester Met will ensure that all personal data collected and processed by Manchester Met and the Delivery Team for this research project are:

- Processed in a manner that is fair, transparent and lawful;
- Adequate and relevant to the study, and are processed solely for the purposes set out in this document;
- Accurate, and where necessary, kept up to date;
- Kept in a form which permits identification of data subjects for no longer than is necessary and;
- Processed in a manner that ensures appropriate security of the personal data.

This evaluation will be assessed for data protection and ethics as part of the research ethics approval process in place at Manchester Met. All personal data will be treated with strictest confidence by the evaluators in accordance with the requirements of the GDPR 2018.

Manchester Met 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 Manchester Met and Nottingham Trent University.

The Delivery Team shall ensure that all participating schools sign a Data Sharing Acknowledgement outlining what data schools will share with the Delivery Team and the Evaluation Team.

Data will be processed by Manchester Met in order to ascertain the impact of the intervention on the pupil outcomes above, and to make judgements about compliance and fidelity. 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 agreed with the University's Data Protection Officer for parents and teachers. Pupils may withdraw from data processing at any time during the study.

As a public authority conducting research and analysis in the public interest which has undergone ethical approval, use the following lawful bases for the processing of:

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

Any information identifying students will be given a unique code immediately after collection and prior to analysis in order to reduce risk. Archived data will include pupil UPNs and matching to the NPD and other administrative data may take place by the Data Archive Manager. However, data will only be released subsequently to interested parties in an anonymised format. 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 pupil personal data collected for the impact evaluation and held by Manchester Met and Nottingham Trent University will be destroyed in accordance with the GDPR when it is no longer required, and no later than July 31st 2024. Data from interviews (headteachers, teachers, pupils) will be archived in an open access repository and the consent forms will be retained if participants indicate on their consent forms that they are happy for this to happen.

Personnel

DELIVERY TEAM: PALS-UK (NOTTINGHAM TRENT UNIVERSITY AND THE UNIVERSITY OF BIRMINGHAM)

Project Leader(s): Dr Emma Vardy (Nottingham Trent University) and Dr Helen Breadmore (University of Birmingham), supported by Dr Luisa Tarczynski-Bowles (Nottingham Trent University)

EVALUATION TEAM: (MANCHESTER METROPOLITAN UNIVERSITY)

PRINCIPAL INVESTIGATOR(S): Prof Cathy Lewin (Manchester Metropolitan University) and Prof Stephen Morris (Manchester Metropolitan University)

Core fieldwork and analysis team: Dr Steph Ainsworth (Manchester Metropolitan University), Sandor Gellen (Manchester Metropolitan University), Dr Kate Wicker (Manchester Metropolitan University)

Risks

Risk	Probability (1=low, 2=medium & 3=high)	Impact (1-3)	Severity (probability * impact, max = 9)	Mitigation
Recruiting sufficient numbers of schools	2	2	4	Recruiting from 1st trial control schools if necessary. Including schools with mixed age group literacy instruction.
Covid-19 affecting delivery and evaluation	2	2	4	Online data collection where possible. Wider window for post-test on primary outcome.
Missing data and sample attrition	3	2	6	Schools recruited to the trial may decide to withdraw, and this sample loss might both reduce precision of statistical estimates and introduce bias. Drawing on our experience and that of the developer, we will devise a strategy to limit attrition. Where attrition occurs, steps can be taken in analysis to test various assumptions regarding missingness and assess consequences for bias and precision (described above). Other sources of missingness can result from mis-recording of identifying data for pupils, mitigated through carrying out extensive checks on student records prior to randomisation.
Addressing and measuring compliance	2	2	4	In many studies, schools assigned to an intervention fail to engage with it and those in the control group take part. To limit this, a strong communications strategy is required that emphasises the importance of schools engaging with the intervention and stressing the benefits. It will be important to communicate expectations clearly to control schools given that they will have signed up to the project because of an identified need. Identifying what control schools do in relation to teaching reading during the evaluation will be essential. It has been important for us to work closely with the developers to arrive at a meaningful definition of compliance.
Staff shortages and retention in the evaluation team	2	2	4	Studies that run over extended periods will experience research staff turnover. We have a large number of appropriately qualified and experienced staff, and flexible workload and staff management systems that will enable us to ensure this project is always appropriately staffed and managed. Processes such as maintaining a variable library and log for each data source will also be put in place, alongside appropriate handover processes if required.
Poor communications between evaluators and developers	1	3	3	As part of our project planning, we will hold regular meetings with the developers with greater frequency around important milestones (e.g. school recruitment, enumeration, etc.). We will also share our project plans and risk management documents with the developer regularly and ensure our management processes align where possible with those of the project team ensuring collective and appropriate responses to emerging challenges.

Timeline

Table 5: Timeline

Dates	Activity	Staff responsible/ leading
Dec 2021 – Jan 2022	Start-up meetings/review theory of change	Cathy Lewin/Stephen Morris
Dec 2021 – Feb 2022	Data governance/MoU, parental withdraw & data processing notices drafted and agreed	Cathy Lewin
Feb 2022	Ethical approval	Cathy Lewin
Feb 2022	Finalise trial/IPE design	Cathy Lewin/Stephen Morris
Feb-Jun 2022	School recruitment	Emma Vardy
Apr-Jun 2022	Parent withdrawal process	Emma Vardy
Apr-Jun 2022	Data collection team requests data from schools	FFT Education
Jun-Jul 2022	Baseline assessment	Steph Ainsworth/Kate Wicker
Sep 2022	Randomisation	Sandor Gellen
Sep 2022	Inform schools of outcome	Emma Vardy/Helen Breadmore
Sep 2022	Headteacher and teacher baseline surveys	Steph Ainsworth/Kate Wicker
Oct 2022	Initial training for intervention	Emma Vardy/Helen Breadmore
Oct 2022	Observe initial training session	Cathy Lewin
Nov 2022	Protocol	Stephen Morris/Cathy Lewin
Nov 2022	SAP	Stephen Morris
Nov 2022	Intervention commences	Emma Vardy/Helen Breadmore/Schools
Nov 2022	Pupil training	Schools
Nov 2022	Observations of pupil training	Emma Vardy
Nov 2022	Top-up training	Emma Vardy/Helen Breadmore
Nov 2022	Observe top-up training	Cathy Lewin
Dec 2022	Pupils commence full intervention period	Schools
Dec 2022/May 2023	Observations of intervention	Emma Vardy/Schools
Jan/Feb 2022	First case study visits	Cathy Lewin

Dates	Activity	Staff responsible/ leading
Mar/Apr 2023	Second case study visits	Cathy Lewin
April 2023	Recruit and train post-test data collectors	Steph Ainsworth
May 2023	Mandatory period of intervention ends; schools may continue until end of academic year.	
Jun/Jul 2023	Post-test headteacher and teacher surveys	Steph Ainsworth/Kate Wicker
Jun/Jul 2023	Interviews with delivery team	Cathy Lewin
Jun/Jul 2023	Post-test student data collection (New PiRA, WIAT-III UK)	Steph Ainsworth
Jul 2023	IPE data collection completed	Cathy Lewin
Jul 2023	Impact data collection completed	Steph Ainsworth
Jul-Sep 2023	Data linking, cleaning and structuring	Sandor Gellen
Oct 2023-Jan 2024	Data analysis	Cathy Lewin/Stephen Morris
Jan 2024	Draft report	Cathy Lewin/Stephen Morris
|June 2024	Final report	Cathy Lewin/Stephen Morris

References

Allen, R., Jerrim, J., Parameshwaran, M., & Thompson, D. (2018). *Properties of commercial tests in the EEF database.* London: EEF Research Paper, (001).

Boylan, M., Demack, S., Wolstenholme, C., Reidy, J., & Reaney, S. (2018). *ScratchMaths: evaluation report and executive summary*. London: Education Endowment Foundation.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.

Braun, V., & Clarke, V. (2021). Braun, V., & Clarke, V. (2021). *Thematic Analysis: A Practical Guide*. London: Sage.

Breadmore, H.L., Vardy, E.J., Cunningham, A.J., Kwok, R.K.W., & Carroll, J.M. (2019). Literacy Development: Evidence Review. London: Education Endowment Foundation. The report is available from: https://educationendowmentfoundation.org.uk/public/files/Literacy_Development_Evidence_Review.pdf

Carroll, J.M., & Fox, A.C. (2017). Reading Self-Efficacy Predicts Word Reading But Not Comprehension in Both Girls and Boys. *Frontiers in Psychology, 7*.

Clark, C., & Rumbold, K. (2006). *Reading for pleasure: A research overview*. National Literacy Trust.

Culora, A., Dimova, S., Ilie, S, Sutherland, A., & Gilder, L. (2022). *Peer Assisted Learning Strategies – UK Evaluation Report.* London: Education Endowment Foundation.

Demack, S. (2017) *Statistical Analysis Plan for ScratchMaths.* London: Education Endowment Foundation.

Demack, S. (2019). *Does the classroom level matter in the design of educational trials? A theoretical & empirical review.* London: Education Endowment Foundation.

Dong, N., & Maynard, R. A. (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. https://doi.org/10.1080/19345747.2012.673143

Dunst, C., Simkus, A., & Hamby, D. (2012). *Children's Story Retelling as a Literacy and Language Enhancement Strategy*. *CELLreviews* 5(4). Asheville, NC: Orelena Hawks Puckett Institute, Center for Early Literacy Learning.

Education Endowment Foundation (2018). *The EEF Teaching and Evidence Toolkit.* London: Education Endowment Foundation. Retrieved from: https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/peer-tutoring/

Education Endowment Foundation (2019). *Cost evaluation guidance for EEF evaluations*. London: Education Endowment Foundation. Retrieved from: https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/Cost_Evaluation_Guidance_2019.12.11.pdf

Fuchs, D., Fuchs, L.S., Mathes, P.G., & Simmons, D.C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal, 34*(1), 174-206.

Gorard, S., Siddiqui, N., See, B.H., Smith, E., & White, P. (2017). *Children's University: Evaluation Report and Executive Summary.* London: Education Endowment Foundation

Humphrey, N., Hennessey, A., Ashworth, E., Frearson, K., Black, L., Petersen, K., ... & Pampaka, M. (2020). *Good Behaviour Game. Evaluation Report and Executive Summary.* London: Education Endowment Foundation.

Jay, T., Willis, B., Thomas, P., Taylor, R., Moore, N., Burnett, C., ... & Stevens, A. (2017). *Dialogic teaching: Evaluation report and executive summary.* London: Education Endowment Foundation.

Kiger, M.E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical Teacher, 42*(8), 846-854.

Klauda, S.L., & Guthrie, J.T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational psychology, 100*(2), 310.

Lloyd, C., Edovald, T., Kiss, Z., Morris, S., Skipp, A., & Ahmed, H. (2015). *Paired Reading: Evaluation Report and Executive Summary*. London: Education Endowment Foundation. Retrieved from: https://files.eric.ed.gov/fulltext/ED581127.pdf

Morris, S.P., Edovald, T., Lloyd, C., & Kiss, Z. (2016). The importance of specifying and studying causal mechanisms in school-based randomised controlled trials: lessons from two studies of cross-age peer tutoring. *Educational Research and Evaluation, 22*, 422-439.

O'Hare, L., Stark, P., Cockerill, M., Lloyd, K., McConnellogue, S., Gildea, A., ... & Bower, C. (2019). *Reciprocal Reading: Evaluation Report.* London: Education Endowment Foundation

Paige, D.D., Rasinski, T., Magpuri-Lavell, T., & Smith, G.S. (2014). Interpreting the relationships among prosody, automaticity, accuracy, and silent reading comprehension in secondary students. *Journal of Literacy Research, 46*(2), 123-156.

Peura, P., Aro, T., Viholainen, H., Räikkönen, E., Usher, E.L., Sorvo, R., & Aro, M. (2019). Reading self-efficacy and reading fluency development among primary school children: Does specificity of self-efficacy matter? *Learning and Individual Differences, 73*, 67-78

Pikulaki, J.J., & Chard, D.J. (2011). Fluency: Bridge Between Decoding and Reading Comprehension. *The Reading Teacher, 58*(6), 510-519.

Rasinski, T.V. (2003). *The fluent reader: Oral reading strategies for building word recognition, fluency, and comprehension*. Scholastic Inc.

Rasinski, T.V. (2004). Creating fluent readers. *Educational Leadership, 61*, 46–51

Rasinski, T.V., Padak, N. D., McKeon, C. A., Wilfong, L. G., Friedauer, J. A., & Heim, P. (2005). Is reading fluency a key for successful high school reading? *Journal of Adolescent & Adult Literacy, 49*(1), 22-27.

Reis, S.M., Eckert, R.D., McCoach, D. B., Jacobs, J. K., & Coyne, M. (2008). Using Enrichment Reading Practices to Increase Reading Fluency, Comprehension, and Attitudes, *The Journal of Educational Research, 101*(5), 299-315.

Rudd, P., Aguilera, A.B.V., Elliott, L., & Chambers, B. (2017). MathsFlip: *Flipped Learning. Evaluation Report and Executive Summary.* London: Education Endowment Foundation.

Topping, K., Millder, D., Thurston, A., McGavock, K., & Conlin, N. (2011). Peer tutoring in reading in Scotland: thinking big. *Literacy*, 45(1), 3–9.

Vardy, E.J., Breadmore, H.L., & Carroll, J.M. (under review). Measuring the will and the skill of reading: Validation of the self-efficacy and motivation to read scale.

WWC. (2012). *Peer-Assisted Learning Strategies.* Retrieved from: https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_pals_060512.pdf

Yildirim, K., Rasinski, T., Ates, S., Fitzgerald, S., Zimmerman, B., & Yildiz, M. (2014) The Relationship Between Reading Fluency and Vocabulary in Fifth Grade Turkish Students, *Literacy Research and Instruction*, 53(1), 72-89.

Zutell, J., & Rasinski, T. V. (1991). Training teachers to attend to their students' oral reading fluency. *Theory Into Practice, 30*(3), 211-217.

Appendices

Appendix 1: Changes since the previous EEF evaluation

	Feature	20/21 EEF Efficacy Trial (impacted by Covid-19) to 22/23 EEF Efficacy Trial
Intervention	Intervention content	Manual and initial top-up training materials updated following feedback from previous trial; book list updated; more extensive 'just-in-time' support for teachers; teachers will be encouraged to use baseline data to inform initial pairings
	Delivery model	'Just-in-time' support for teachers to include more components that can be accessed asynchronously online (e.g. support videos, FAQs, discussion boards)
	Intervention duration	No changes
Evaluation	Eligibility criteria	Exclusion criteria revised to focus on the three required Regional School Commissioner Regions; schools must not be involved in another Accelerator Fund efficacy trial or another EEF literacy project; schools must provide access to technology for online reading assessment; control schools from previous trial will be recruited.
	Level of randomisation	No changes
	Stratification variables	Unlike in the previous trial, randomisation won't be stratified by region. Instead, we introduce FSM eligibility as the second stratifying variable alongside school size.
	Outcomes and baseline	Baseline testing will be in the summer of year 4 (rather than autumn year 5) using PiRA Summer 4; the secondary outcomes, reading comprehension and oral reading fluency, will be measured using WIAT-III UK-T face to face after the intervention (online administration of this was done with only a few students in the previous trial due to covid-19 disruption); a qualitative multidimensional fluency scale will be used alongside WIAT-III

	UK-T; motivation for reading will also be included as an additional secondary outcome measure; the post-test primary outcome measure will be taken in the summer term following the intervention (rather than the 6 months delay in the previous trial). Post-tests to be administered by evaluation team (not teachers);
Control condition	No changes

Appendix 2. Harmonic mean calculations

Number of classes at Y5[16]	Number of schools
1	8004
2	4319
3	1210
4	241
5	31
6	10
8	1
Number of Y5 pupils per class	**Number of schools**
less than 10	839
10-15	1064
15-20	862
20-25	2261
more than 25	8790
Mean calculations	
Average no of classes (Arithmetic mean)	1.55
Average no of classes (Harmonic mean)	1.30
Average no of pupils per class (Arithmetic mean)	24.51
Average no of pupils per class (Harmonic mean)	20.19

[16] All state-funded primary schools are included where the assumed number of Y5 pupils is more than 1