Manchester Metropolitan University

e-space

Manchester Metropolitan University's Research Repository

Please cite the Published Version

Gellen, Sandor, Wicker, Kate, Ainsworth, Steph, Morris, Stephen and Lewin, Cathy (2024) Using the DreamBox Reading Plus adaptive literacy intervention to improve reading attainment, a two-armed cluster randomised trial evaluation protocol. Project Report. Education Endowment Foundation, London, UK.

Publisher: Education Endowment Foundation

Version: Published Version

Downloaded from: https://e-space.mmu.ac.uk/635153/

Usage rights: © In Copyright

Enquiries:
If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines)

Using the DreamBox Reading Plus adaptive literacy intervention to improve reading attainment, a two-armed cluster randomised trial
Evaluation Protocol

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

Evaluation summary

Project title	Using the DreamBox Reading Plus adaptive literacy intervention to improve reading attainment, a two-armed cluster randomised trial
Developer *(Institution)*	Reading Solutions UK
Evaluator *(Institution)*	Manchester Metropolitan University (MMU)
Principal investigator(s)	Steph Ainsworth and Stephen Morris
Protocol author(s)	Sandor Gellen, Kate Wicker, Steph Ainsworth, Stephen Morris, Cathy Lewin
Trial design	Two-arm cluster randomised controlled trial with random allocation at the school level
Trial type	Efficacy
Pupil age range and Key stage	9-10 years old, KS2
Number of schools *(at design stage)*	126
Number of pupils *(at design stage)*	6,426
Primary outcome measure and source	Reading attainment (New PiRA Summer 5 Test)
Secondary outcome measure and source	Reading comprehension (KTEA-3 comprehension subscale) Reading fluency (KTEA-3 Silent Reading fluency subscale) Reading vocabulary (KTEA-3 vocabulary subscale) Reading self-efficacy (Feeling about Reading self-efficacy subscale) Reading motivation (Feeling about Reading motivation subscale)

Protocol version history

Version	Date	Reason for revision
1.2 [*latest*]		
1.1		
1.0 [*original*]	05/07/2024	N/A

Table of Contents

Tables

Figures

Study rationale and background

DreamBox Reading Plus (known as Reading Plus hereafter) is an online (EdTech) adaptive silent reading programme designed to improve reading and language comprehension skills (Spichtig, A. N. et al., 2019). The programme supports fluency, comprehension (e.g. inference) and vocabulary growth, It is designed to support all readers including pupils with EAL, SEND, as well as the most able students. The programme incorporates a visual skills element which scaffolds pupils' reading via a Guided Window text display to support eye movement control (Radach and Kennedy, 2013). The programme also includes additional visual skills activities to support struggling readers. Another unique feature of the programme is that children self-select reading tasks based on age-appropriate texts. These tasks are designed to be 'high interest' for all children, including lower attaining readers. Overall, the programme promises to have an important socio-emotional impact by building stamina and motivation to read. In addition to promoting reading proficiency directly (by improving pupils' fluency, vocabulary and reading comprehension), the programme also promises to have an indirect effect on reading attainment, through improvements in teachers' knowledge of the role of silent reading fluency in developing reading stamina and comprehension.

A small non-randomised efficacy study of Reading Plus in the UK (Reading Solutions UK, 2022) provided promising evidence of Reading Plus's potential. Within this study, involving 470 year 6 pupils, the intervention group engaged with Reading Plus for a minimum of 30 minutes, three times per week over nine months. Students eligible for pupil premium were found to make, on average, 97% more progress in reading attainment with Reading Plus than in the comparison schools. In this study, reading attainment was measured at three timepoints (beginning, middle and end of the intervention), using past SATS papers for the first two timepoints and the actual SATS paper for the final timepoint. In the US, a small randomised controlled trial (Spichtig, et al., 2019) of Reading Plus with fourth- and fifth-grade students (n=426) produced gains in reading achievement relative to business as usual. In this trial pupils in the intervention group engaged with the programme in 25-minute supplemental literacy blocks and reading proficiency was measured at the beginning and end of the school year using the Group Reading Assessment Diagnostic Evaluation (GRADE)[1]. While the two existing trials suggest that Reading Plus may be effective for accelerating pupil reading attainment, a larger independent evaluation is needed. Reading Plus is growing in popularity with over 1,400 schools currently using it within the UK.

The digital nature of Reading Plus provides several potential benefits. At the core of the programme is automated differentiated reading instruction underpinned by continued formative assessment, which informs tailored self-paced reading lessons (Spichtig et al., 2019). This automated intertwined process of assessment and teaching is designed to provide learning experiences closely adapted to pupils' current level of development. It also has the benefit of reducing teacher workload, as the programme is designed to largely run itself once pupils get started.

The proposed benefits of Reading Plus as an online reading intervention align with previous research into the advantages of Ed Teach interventions. Recent studies in England suggest that headteachers and teachers believe EdTech interventions have the potential to reduce

[1] GRADE Group Reading Assessment & Diagnostic Evaluation (pearsonassessments.com)

workload through, for example, supporting assessment (Cooper Gibson Research, 2022, 2021). More specifically, See et al. (2022) conclude that EdTech formative assessment can support improvements in reading at the primary phase, and a meta-analysis of comprehension interventions conducted by Silverman et al. (2020) suggest that technology interventions can have a positive impact on reading comprehension. These positive indications are broadly supported by a review commissioned by the Education Endowment Foundation (EEF) (Lewin et al., 2019), although the evidence in Lewin's systematic review does not consistently point toward a positive impact.

In relation to groups who might benefit, Silverman et al. (2020) reported that disadvantaged students did not benefit as much from such interventions. In contrast, some of the studies reviewed by Lewin et al. (2019) suggested that socio-economically disadvantaged students benefit more. Silverman et al. (2020) also noted that English language learners benefitted more from EdTech reading comprehension interventions than native language speakers. These apparent contradictions suggest a need to gather further rigorous evidence about which groups benefit and why.

The evaluation design set out within this protocol aims to extend this emerging evidence in relation to the potential efficacy of Reading Plus. The impact evaluation (IE) takes the form of a two arm, parallel, cluster randomised controlled efficacy trial, with whole primary schools randomised to intervention and control groups. This will be integrated with an implementation and process evaluation (IPE), which will explore the relationship between the delivery of the programme and any impact of Reading Plus on pupil outcomes (or not) within the IE. The IPE will be grounded in the theory of change (Humphrey et al., 2016).

Primary schools in the intervention arm of the trial will be invited to use Reading Plus as part of their timetabled reading instruction. Within intervention schools, pupils entering Year 5 in September 2024 and their teachers will be eligible to use Reading Plus. The primary outcome for the study will be a measure of reading attainment. Pupils in both intervention and control schools will complete reading assessments at the end of the intervention in the summer term of 2025. The effect of Reading Plus on reading attainment will be calculated by computing the difference in average pupil reading attainment scores for the intervention and control arms of the trial. The study will run in England and not be restricted to any areas of the country.

The decision to employ a cluster randomised controlled trial (RCT) design was made after careful consideration of various other options. These options included:

- A single year group individual pupil-level randomisation with school blocking.
- A single year group class-level randomisation with school blocking.
- A two-year group randomisation in which intact Year 5 and 6 groups are randomised to either treatment or control conditions blocked by school.
- A cluster randomised design in which whole schools are randomised to intervention and control conditions.

Option 1 was considered impractical due to worries about the burden of restructuring classes. It was felt that imposing such a condition on schools for trial participation would likely deter many from taking part. Option 2 was rejected because it was anticipated that control class teachers might take compensatory measures if intervention and control groups were in the same school year. Additionally, participation in the study would be limited to schools with two

or more classes per year group, thereby excluding smaller schools with single-class year groups or mixed year-group teaching.

Option 3 was considered but ultimately dismissed in favour of Option 4. The decision to proceed with Option 4 was influenced by practical factors. Practically, concerns focused on the delivery of Reading Plus to Year 6 pupils, given the high stakes associated with Year 6 national curriculum tests. Randomisation occurring at the end of the summer term or the beginning of the autumn term would have imposed undue pressure on schools and teachers to adjust their teaching plans promptly. Therefore, focusing solely on Year 5 pupils was deemed more feasible. Additionally, informal discussions with teachers and headteachers indicated a preference for introducing Reading Plus in Year 5 rather than Year 6.

Intervention

The subsequent segment provides comprehensive details of the intervention, adhering to the TIDieR framework (Hoffmann et al., 2014).

As outlined above, Reading Plus is an online (EdTech) adaptive silent reading programme designed to improve pupils' reading and language comprehension skills (Spichtig et al., 2019). Its key features include: an emphasis on silent independent reading, a Guided Window (see below) to support eye movement control as pupils read the text on the screen (Radach and Kennedy, 2013), visual skills activities to support struggling readers as needed, explicit vocabulary instruction, and 'high interest' texts for readers of all abilities which pupils self-select. Each of these aspects of the intervention will be explored in detail below. Within the UK, Reading Plus is advertised as being suitable for children in year two and above (to beyond GCSE). The current trial aims to measure the potential effect of Reading Plus on pupils in Year 5, when there is a strong focus on developing their reading comprehension skills in preparation for the end of KS2 reading assessment.

Training

The intervention will be delivered to all year 5 pupils in intervention schools. Year 5 teachers will attend two sessions of online training delivered by Reading Solutions UK (each lasting 45 minutes): one session at the beginning of the intervention (October 2024) and one three weeks into the implementation period. This training will equip teachers with the knowledge they need to deliver Reading Plus, including explanation of the pedagogy underpinning the programme and how to use the assessment data generated to provide targeted support. The delivery team will set up all pupils on the Reading Plus system, minimising the burden on teachers. About six weeks into the programme teachers will also receive a call from a Reading Development Consultant, who will check-in with the Year 5 teachers to address any queries and ensure that they are fully utilising the capabilities of the programme. As part of the intervention, Year 5 teachers can access optional unlimited supplementary online training and gain access to resources (ongoing, refresher, or training for new staff) throughout the intervention period.

Dosage

Year 5 teachers in intervention schools will be instructed to deliver Reading Plus to the whole class for at least 90 minutes a week over a minimum of two sessions e.g., two 45 minute sessions or three 30 minute sessions per week for the duration of the intervention period (October 2024 to May 2025). The Reading Plus lessons will replace other reading activities that would have taken place in school, e.g., guided reading sessions or other reading/literacy activities. The intervention will mainly take place online within the Reading Plus platform, apart from some optional supplementary offline activities that teachers can use with pupils who might require further support. Schools may continue to use Reading Plus for the rest of the academic year if they wish. There is no requirement for schools to encourage pupils to access Reading Plus at home outside of lessons in school, but it is acceptable for them to do this.

The logic model within Figure 1 sets out the mechanisms, which are proposed to underpin the programme. In the rest of this section, the key elements of the programme are described along with how each element is designed to support a specific aspect of reading development.

In-built assessment

At the beginning of the intervention, when pupils first log onto the Reading Plus platform, they will be assessed on their reading speed, comprehension, vocabulary and motivation via the InSight reading assessment. This is a proprietary online reading assessment, which is built into the Reading Plus platform[2]. Scores on this assessment have been shown to have strong correlations with a range of other normed assessments including the Group Reading Assessment Diagnostic Evaluation (GRADE, r = .84) (Gehsmann et al., 2017). A review of the InSight assessment found 'convincing evidence' of its classification accuracy, reliability and validity (National Center on Intensive Intervention, 2019). This review of InSight also reports a bias analysis, stating that all items with large Differential Item Functioning were removed following field testing. There are norms published for the InSight assessment, generated within the US; there are no UK norms available, but the assessment within the UK version of Reading Plus is the same as the US version apart from anglicisation of spelling, etc.

The InSight assessment has four components:

- a motivation inventory, which asks pupils to rate items (on a six-point Likert scale) relating to their reading confidence, reading interest and self-inprovement belief
- an adaptive vocabulary assessment, which presents a target word and asks students to identify a synonym from a choice of four words
- an adaptive comprehension assessment, where students read passages of text and are asked literal and inferential questions about each text (in multiple choice format)
- a measure of comprehension-based silent reading rate. This is calculated based on how long it takes pupils to read the text on each screen. It is defined as a comprehension-based reading rate because only reading rates from those screens where pupils understood the text (as evidenced by them correctly answering the question) are used to calculate the reading rate.

The InSight assessment takes around 45 minutes to complete (including time spent watching a short video at the beginning of each task to explain what they will need to do). The different scores are combined to create a composite reading proficiency year-level score.

Pupils' scores on their initial InSight assessment will influence which reading activities are then available to them on the platform. This aspect of the programme is based on the assumption that effective reading comprehension requires texts to be matched to pupils' developmental level along the dimensions of vocabulary, syntactic and semantic complexity, required background knowledge, and length' (Spichtig et al., 2019, p. 43). Teachers will be able to access a Class Screening Report which indicates which students are below, at, or above year-level expectations. The screening report also identifies students who may need additional support with reading comprehension/vocabulary (which is reported as a composite score) and/or reading fluency.

[2] Further detail about the InSight assessment and the broader Reading Plus programme can be found here https://www.readingsolutionsuk.co.uk/wp-content/uploads/2024/04/DreamBox-Reading-Plus-Theoretical-Framework-and-Foundational-Research.pdf

Formative assessment is built into the programme throughout and continues to influence the content that pupils engage with. In this way the programme is designed to provide adaptive scaffolding for pupils' reading development, which may be particularly important for struggling readers, including pupils experiencing social disadvantage, who are more likely to experience difficulties with reading (e.g., EEF, 2017; Jehangir, Glas & van den Berg, 2015).

Year 5 teachers will oversee pupils' use of Reading Plus. They will be able to monitor pupils' live use of the platform through the teacher dashboard. The reports available to teachers will allow them to identify gaps in proficiency/skills. They can then use this data to intervene with small group/one-to-one sessions using the supplementary lesson materials provided. The logic model also suggests that these reports may be useful to the school/trust, as they provide assessment data which will allow them to track pupils' reading progress throughout the year, including monitoring the progress of disadvantaged pupils.

Guided Window

Within the online lessons, pupils self-select from the range of online texts that are available to them based on their current ability level. As they read the texts, they are supported by a (patented) 'Guided Window' on the screen. The Guided Window (described in detail below) is designed to, help to build fluency and stamina. Fluency is a foundational skill that is needed for pupils to become skilled readers (e.g., Ehri, 1997; Breadmore et al., 2019) as it supports students to focus on reading for meaning, allowing them to develop their proficiency in reading comprehension (Paige, Silverman, Speece, Harring & Ritchley, 2013). The association between fluency and reading comprehension becomes increasingly important as pupils reach the stage where they are confident at decoding print (Language and Reading Consortium, 2015), and direct teaching of fluency has been shown to drive improvements in reading attainment (e.g., Klaudia & Guthrie, 2008). Reading Plus is designed to be used with children who have good decoding skills to support them with the transition for 'learning to read' to 'reading to learn' (Spichtig et al., 2019, p. 461).

The 'Guided Window' presents the page through a filter such that its layout is generally discernable but only specific pieces of text are legible. The Window is 25 characters wide. It moves left to right over the page displaying different pieces of text as it moves. The design of the Guided Window aligns with natural reading behaviours. As stated within the Reading Plus Theoeretical Framework and Foundational Research document (Reading Solutions UK, no date, p. 2):

> *…it accommodates the right-oriented perceptual span asymmetry of readers of English (MConkie & Rayner, 1975; Rayner, 1986), as well as typical saccade spans, which commonly average eight to ten characters during silent reading, and over 90 percent of regressive saccades (Rayner, 1998; Vitu & McConkie, 2000).*

IThe speed at which the window moves is determined by a student's reading rate. Actual letters and words can only be recognised and read when they appear within the Window. This feature of the platform is designed to model effective reading behaviours, helping to build fluency and stamina.

The Guided Window is a novel tool for supporting fluency development. It guides the pupil's eyes across the text on the screen at a speed that is adapted to the current ability level of the reader. Specifically, if a pupil scores 80% on the comprehension questions across two or three texts in a row, the speed of the guided window increases. Pupils are able to choose Pupils can

choose by how many words per minute it will increase by: 3, 5, or 7. The Guided Window is designed to focus the pupil's visual attention so that they can effectively keep track of where they are on the page (cf. the traditional method of asking pupils to follow along the page with their finger). By focusing the pupils' eyes on one small section of text at a time, the Guided Window aims to support sequential reading of the words on the screen, which in turn, supports fluent reading. It aims to minimise obstacles that emergent readers can face, such as frequent re-reading of words/phrases (because the reader have lost their place) and decoding words that are already familiar. While some back and forth eye-movements have been found to be a normal part of the reading process in skilled readers in emergent readers too many regressions (times when the reader moves their eyes backward) can make the reading process inefficient (Spichtig et al., 2016). Similarly, inefficient reading is also associated with a large number of fixations (eye stops – the number of times that the eyes rest on a particular word) and a smaller perceptual span (the number of letters a reader can perceive and process during a fixation). In other words efficient readers, do not move their eyes backwards as often when reading, their eyes do not pause on a particular word/section of text as often, and when they do, they are able to process more of the text at once (ibid). In a small-scale US trial, the ReadingPlus programme led to significant reductions in fixations ('eye stops' when reading a particular word) and regressions (going back to 'refixate' on a word) (Spichtig et al., 2019, p. 447). The adaptive nature of the Guided Window increases the rate at which it moves gradually to help pupils' incrementally improve their stamina and fluency levels. The pupil is not able to pause the Guided Window as it moves across the text in order to re-read a section that they perhaps have not understood. However, if a pupil is not certain of the answer to a comprehension question once they have read the text, they can press the 'ReRead' button, which then allows them to look at the portion of text, which relates to that specific question.The pupil is able to use up to two 'ReReads' per text selection.

The way that the system calculates this rate takes into account the fact that it is not just reading speed that is important but being able to understand what is being read as you read. After a child has read a text, they are asked some questions about the text's meaning. The questions are designed to both develop and monitor pupils' reading comprehension skills. The system is intended to gradually increase pupils' 'comprehension-based silent reading rate' – the speed with which they can read and effectively understand the text. A core assumption of the logic model is that by gently ramping up pupils' fluency, the programme supports pupils to focus more on reading for meaning (e.g., see Paige et al., 2013). The Guided Window becomes optional when a pupil reaches the expected level of reading proficiency for their age. They are then able to focus on reading independently. At this stage the system continues to scaffold pupils' development of stamina and fluency by gradually increasing the length of the text segments that are displayed on the page. Previous studies have provided promising evidence of the potential for use of the Guided Window to accelerate reading achievement (Rasinski et al., 2011; Reutzel, Petscher & Spichtig, 2012).

Reading tasks

Pupils using Reading Plus will generally complete five reading tasks per week (plus one vocabulary task – see below), which are designed to support both silent reading fluency and comprehension development (although this can be adapted to suit individual pupil needs).

Comprehension tasks

Once pupils have finished reading a text, the system asks them ten comprehension questions. The pupils' performance on these questions influences how they progress through the programme. The questions are designed to gradually increase in complexity, scope and depth. They provide scaffolds to support the pupils in engaging meaningfully with the text e.g., pupils may be provided with a relevant excerpt of the text alongside the questions. Pupils move onto the next level when they consistently achieve 80% of more on the questions. The adaptive nature of the programme, which is designed to provide pupils with the most appropriate text for their current proficiency level, aims to build reading self-efficacy and confidence.

The comprehension questions are designed to develop three categories of skills:

- Core – These questions involve pupils recalling details, making inferences, drawing conclusions, determining main ideas and themes, summarising, paraphrasing, predicting, and analysing plots, characters and cause and effect.
- Craft – Pupils are asked questions about the way that the language is crafted and the structure of the text. This includes questions about the author's use of language, examination of text structures and the sequencing of ideas, and considering the author's purpose, point of view, use of persuasive devices etc.
- Critical – Pupils are prompted to consider the arguments presented within a text, to consider the use of scaffolds within the text, e.g. imagery, and to compare and contrast across texts.

The questions are presented in a range of formats designed to align with the kinds of question styles that pupils will be asked in the national assessment for reading at the end of year 6. Pupils' responses to the questions allow their development across a range of comprehension subskills to be monitored (teachers can pull down reports from the platform in relation to this). Teachers can use this data to assign specific lessons to pupils through the platform if they have specific gaps in their skills profile.

Vocabulary tasks

Reading Plus is also designed to support pupils' vocabulary development. As pupils move through the programme, they will be exposed to texts which are increasingly sophisticated in terms of the vocabulary, and they will engage with one vocabulary task per week (in addition to the five reading tasks). These tasks are designed to broaden pupils' academic and literary vocabulary with a focus on words that they will encounter in the non-fiction and fiction texts that they will engage with through the ReadingPlus platform. Vocabulary development, including acquisition of academic vocabulary, is a crucial supporting factor for the development of reading comprehension (Quinn, Wagner, Petscher & Lopez, 2015) and learning across the curriculum more broadly (Schuth, Köhne & Weinert, 2017). The broad range of texts available to pupils within Reading Plus has the potential to develop pupils' cultural capital by providing 'windows' (Bishop, 1990, p. ix) into different worlds. A recent study suggested that pupils' engagement with reading-related activities was found to be a significant predictor of pupils' GCSE grades (Stopforth & Gayle, 2022). The logic model predicts that students' cross-curricular knowledge will be developed by engaging with a diverse range of non-fiction texts, allowing them to better access the broader curriculum.

The vocabulary tasks provide explicit instruction 'based on a proprietary ageneral academic vocabulary list that features a core vocabulary derived from 2,451 morphological word families'

(Reading Solutions, no date p. 8). The words within the vocabulary component of Reading Plus are grouped into 'year levels' according to age of acquisition, word frequency, morphological profile and other relevant criteria (further details can be found in Hiebert, Scott, Castaneda & Spichtig, 2019). The vocabulary tasks are designed to draw pupils' attention to high-utility root words and their different inflectional forms (e.g. use, useful, useless). In this way the programme is designed to show children how 'a root word can have multiple forms and shades of meaning' (Reading Solutions UK, no date, p. 8). In the reading comprehension component of the programme (where pupils answer questions about texts) they are provided with opportunities to generalise this knowledge to additional morphological forms (e.g. overused, resuable, etc). The focus on 'high utility morphological word families' – word families that pupils are likely to encounter frequently when reading age-appropriate fiction/non-fiction and/or when learning across the curriculum – aims to support pupils' in inferring the meaning of unfamiliar words from their existing knowledge of words with a shared root (Bowers & Kirby, 2010).

Pupils' engagement with words within the vocabulary component of the programme depends on their performance. There are four types of vocabulary activities. All pupils engage with the first two activities for a given word. If they provide an incorrect answer to one or both of these, they then move on to the second two activities.The four vocabulary activities are:

- Instant recognition – pupils are shown a target word and are asked to choose the word or phrase that most closely aligns with its meaning.
- Word in context – pupils judge which sentences correctly use the target word. 'Hint sentences' as needed.
- Teaching page – pupils are provided with explanations of word meaning, examples of the word being used in sentences and an introduction to the morphological family that the word belongs to.
- Sentence completion – pupils complete a sentence using the correct form of the target word with the morphological family displayed for support.

The Reading Plus platform also contains optional writing activities designed to provide a bridge between reading and writing. However, these are not part of the core programme and are not generally used as part of 'business as usual'. Teachers will be made during the initial training that these writing prompts exist within the platform, but that they should not form part of the minimum of 90 minutes a week of the intervention.

Elements to support motivation to read

The system is also designed to promote motivation to read. While the development of fluency is itself associated with motivation (pupils are less likely to feel motivated to read if it feels too difficult, e.g., Mehigan, 2020), Reading Plus also aims to promote motivation through 'high interest' texts, gamification features and instant feedback. Here the term 'high interest' refers to texts that are likely to be of interest to pupils within that age range, while also being pitched at an appropriate level in terms of vocabulary, text complexity etc. As described in the logic model, Reading Plus aims to achieve this by drawing upon an extensive range of fiction and non-fiction texts, which have been selected to cover a broad range of attainment levels while remaining engaging and age appropriate. The texts available for pupils to read are further tailored to pupils' interests, by the provision of recommendations based on an algorithm which

considers which texts pupils have chosen before and how highly they have rated them. A key design feature of Reading Plus is that pupils are able to select which text they read within a given lesson from a range of texts, which are made available to them depending on their current reading performance. Prior research has highlighted the importance of learners having access to a wide range of high-quality texts (e.g. Cremin, Hendry, Rodriguez Leon & Kucirkova, 2022) and having have some autonomy over which texts they read to promote motivation to read (e.g., Guthrie & Klauda, 2014). The logic model predicts, therefore, that teachers using Reading Plus will find that their classes are generally more engaged and motivated during reading lessons.

Additional teacher-directed activities

The platform provides additional online Visual Skills activities, which teachers can allocate to weaker readers as necessary, based on their baseline assessment data. These tasks have been designed to strengthen pupils' eye muscles, which Samuels, Hiebert and Rasinski (2015) suggest play an important role in the process of efficiently reading written text.

There are two types of visual skills tasks:

- Scan – Pupils' eyes are guided across lines of 'text' via a moving window, where all the letters have been replaced with circles. Every time the pupil sees a circle with a break in it (a Landolt ring), they press the space bar.The speed of the task increases as pupils performance improves. The task is designed to develop the same skills relating to visual discrimination, attention and tracking that are used when reading, but with the additional cognitive load associated with decoding and interpreting meaning removed (Radach & Kennedy, 2013; Radach & Spichtig, 2013).

- Flash – Trigrams (three-letter strings which are present within commonly occurring words) are flashed on the screen to the left or right of a central plus sign (+). The pupil is then asked to type the letters that they just saw in order, e.g. 'igh'. This is designed to support pupils automatic recall of these common letter strings, which in turn is suggested to support more efficient reading.

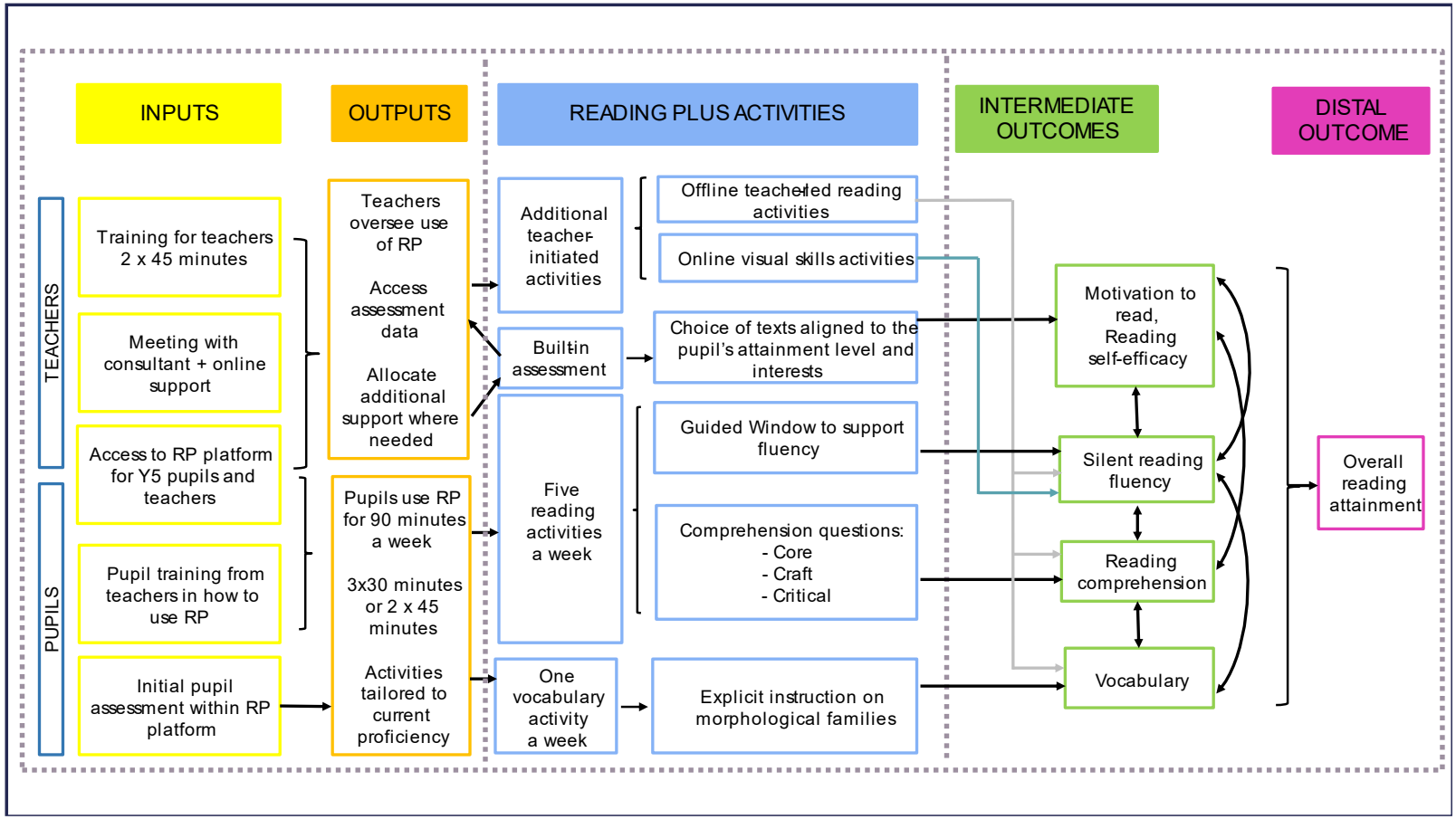
To address pupils' skills gaps, Reading Plus provides teachers with lesson plans and skills worksheets to use outside of the programme in a one-to-one or small group setting. The Skills Summary Report accessed through the teacher dashboard groups pupils who are struggling with the same comprehension sub-skill (for example, inference). Each lesson plan includes a fully scripted, ready-to-use guide for a 25-30 minute session. Printable worksheets explain a specific comprehension skill with examples from real texts. Students then apply these skills through independent practice activities to solidify their understanding.

As standard, all students are allocated five reading and one vocabulary tasks per week to be completed within 90 minutes of scheduled Reading Plus time. Teachers can reallocate the weekly online tasks for students who need additional support with vocabulary and fluency. For example, a student who needs support with vocabulary may be assigned three reading and five vocabulary tasks. Students who struggle with fluency may be assigned 'Visual Skills' tasks to train their eyes to read more efficiently and accurately. Once the student improves, they can be allocated more reading tasks.

Summary of key rationale for design of Reading Plus

In summary, the logic model for Reading Plus suggests that by simultaneously working to improve reading fluency, comprehension, vocabulary, motivation to read and reading self-efficacy, Reading Plus will improve pupils' overall reading proficiency. It is predicted that the programme will help 'close the reading gap' for disadvantaged pupils, due to the adaptive nature of the technology, and the emphasis on fluency and vocabulary development, which are areas where socially disadvantaged children may require further input (e.g., Pace, Luo, Hirsh-Pasek, Golinkoff, 2017; Martins, Reis, Castro & Gaser, 2021). By becoming more confident and experienced readers, pupils are also expected to develop their content knowledge and access to the broader curriculum.

Figure 1: Logic model



Impact evaluation design

Research questions

The purpose of the impact evaluation is to evaluate the effects of Reading Plus on: (1) reading attainment, (2) silent reading fluency, (3) vocabulary, (4) reading comprehension; (5) reading motivation; and (6) reading self-efficacy.

The primary and secondary research questions are as follows:

Primary research question

1. What is the difference in the average score for reading attainment among Year 5 pupils in schools exposed to Reading Plus, 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 silent reading fluency among Year 5 pupils in schools exposed to Reading Plus, compared to Year 5 pupils in control schools exposed to business-as-usual conditions?
2. What is the difference in the average score for vocabulary among Year 5 pupils in schools exposed to Reading Plus, 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 Reading Plus, 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 Reading Plus, 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 to read among Year 5 pupils in schools exposed to Reading Plus, compared to Year 5 pupils in control schools exposed to business-as-usual conditions?

Design

Table 1: Trial design

Trial design, including number of arms		Two-arm, cluster randomised
Unit of randomisation		School
Stratification variables (if applicable)		School size (one-form per year group versus two or more forms per year group) Average score in KS2 Reading NCTs over the last three May tests completed by the school (divided into terciles)
Primary outcome	**Variable**	Reading Attainment

	Measure (instrument, scale, source)	Reading attainment raw score, 0-45, New PIRA Summer 5 Test[3]
Secondary outcome(s)	**Variable(s)**	Reading comprehension, Reading fluency Reading vocabulary Reading self-efficacy Reading motivation
	Measure(s) (instrument, scale, source)	Reading comprehension subscale, Kaufman Test of Educational Achievement (Third Edition) (KTEA-3) [4], 0-67 Silent reading fluency subscale, KTEA-3, 0-110 Reading vocabulary subscale, KTEA-3, 0-46 Reading self-efficacy subscale, 20-140, Feelings about Reading (FAR) Reading Motivation subscale, 10-70, FAR
Baseline for primary outcome	**Variable**	Reading attainment
	Measure (instrument, scale, source)	KS1 reading test scores (scaled), 85-115, school records
Baseline for secondary outcome	**Variable**	Reading attainment Reading self-efficacy Reading motivation
	Measure (instrument, scale, source)	KS1 reading test scores (scaled), 85-115, school records Reading self-efficacy subscale, 20-140, FAR Reading motivation subscale, 10-70, FAR

The impact evaluation is a parallel cluster-randomised controlled trial designed to assess efficacy, involving two arms, with schools randomly assigned to intervention and control groups on a 1:1 basis. All Year 5 pupils in schools randomised to intervention will be encouraged to use Reading Plus. Pupils in schools randomised to control will not be able to access the programme. Randomisation will be stratified to ensure balance on key school-level covariates. Specifically, schools will be divided into terciles based on their average score in KS2 Reading National Curriculum Tests (NCTs) over the last three May tests completed by the school, and within these groups by whether they are single form entry or two or more form entry.

The primary outcome will be a measure of reading attainment derived from the New PiRA reading test, administered online after exposure to Reading Plus in all schools both intervention and control.

Secondary outcomes will include measures of reading fluency, comprehension, and vocabulary from the KTEA-3 standardised assessment tool. Additionally, measures of reading self-efficacy and reading motivation will be obtained from the Feelings about Reading (FAR) questionnaire. The secondary analysis will estimate the effects on reading self-efficacy and motivation outcomes for the full Year 5 cohort. Effects on outcomes measured through the KTEA-3 instrument will be estimated for a randomly selected subset of 10 pupils per school to

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

[4] Kaufman Test of Educational Achievement, Third Edition (KTEA-3) (pearsonclinical.co.uk)

manage costs, as these assessments require face-to-face one-to-one delivery by trained assessors.

Participant selection

The focal cohort is pupils entering Year 5 in September 2024 in primary schools recruited to the trial. Pupils within this cohort will be identified during Year 4, prior to their entry into Year 5.

Schools

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

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

- State-maintained primary, junior, middle or all through school in England.
- Not have held a KS2 Reading Plus programme licence in the 12 months prior to delivery (academic year 2023/24).
- Not involved in an EEF funded trial targeting pupils and teachers in year 5 (in 2024/25).
- Able to provide one-to-one access to a device (PC, laptop, tablet, etc.) so that a whole class of pupils can participate in Reading Plus and can complete online assessments (reading test, questionnaire) at the end of the intervention (June/July 2025)
- Able to timetable whole-class-use of Reading Plus for 90 minutes per week, broken down as 3 x 30-minute sessions or 2 x 45-minute sessions.

In addition, schools are given preference for inclusion in the trial where they do not have mixed age-group classes for literacy instruction. If schools that do have mixed age group classes remain interested in participating, they can be added to a recruitment waitlist. Also, there was a preference for schools not to be participating in any other EEF trial in 2024/25. This was due to concerns that this may affect their ability to engage with this project. If school are participating in another EEF trial and are interested in this project, the possibility will be held open for them to also be added to a recruitment waitlist.

Schools that meet the eligibility criteria and agree to take part in the trial are asked to sign a Memorandum of Understanding (MoU) which sets out the obligations of each stakeholder to the trial as well as the conditions schools are required to meet.

Pupils

All Year 4 pupils enrolled at the time of school recruitment and due to enter Year 5 in September 2024 are included in the study sample. Once The MoU is signed, schools circulate withdrawal letters and information sheets to the parents of these pupils. Parents are given an initial window of two weeks to withdraw their child from the study. During this period and ongoing for these pupils, it is the responsibility of the school to maintain a record of pupils that have withdrawn and ensure that they do not release any personal information regarding these

pupils to FFT[5], MMU or AlphaPlus[6]. If a pupil withdraws subsequent to the two-week period, the school will need to notify MMU using a secure link to ensure that any existing data held on pupils that wish to withdraw is destroyed. These pupils, together with those who join the schools after baseline assessment will not be included in the evaluation. While students entering schools in the treatment group later will participate in the intervention, they will not complete assessments at endline. Parents retain the right to withdraw their child from the evaluation at any point during the study period.

Recruitment

It is expected that recruitment will take place over the period February to May 2024.

The process for recruiting schools is as follows: Reading Solutions will identify and approach schools that meet the selection criteria. Initial data, including: school name, address, telephone number, URN, and the names and contact details of Year 5 teachers in 2024/25, will be collected. Schools will be invited to sign a Memorandum of Understanding (MoU). The MoU will provide information about the project, its aims, potential benefits for participating schools, a timetable of activities, data protection considerations, and the responsibilities of all parties involved. Once a school signs the MoU, the school is considered to have been 'recruited' to the study and Reading Solutions will share school information with FFT and MMU. Additionally, schools are required to sign a separate Data Sharing Agreement, detailing how personal data will be collected and shared among Reading Solutions, the Evaluation Team, and the school.

As soon as the two-week opt-out period has passed, FFT will collect pupil level information from all pupils expected to enter Year 5 in September 2024 excluding those pupils that have withdrawn. Once records are collected for each eligible pupil in the school, they will be verified and checked for completeness and passed to MMU.

Outcome measures

Primary outcome

The primary outcome will be a measure of reading attainment in the form of a score obtained from the Summer-Term Year 5 PiRA assessment administered online in June/July 2025. New PiRA has 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 curriculum test scores. The PiRA test is an overall measure of reading attainment, which relies on pupils having proficiency in fluency (so that they can get through the test in the allotted time), reading comprehension (so that they can effectively extract meaning from the text in order to answer the questions), and vocabulary knowledge (so that pupils are able to understand the text and the questions asked within the test). Given that the logic model

[5] FFT Education (FFT) – a partner organisation responsible for the enumeration of the pupil sample and collection of basic pupil-level information, and the generation of the initial record in the trial database

[6] AlphaPlus Consultancy Limited is an education service business responsible for overseeing primary data collection carried out by subcontracted Test Administrators

proposes reading attainment to be developed through the development of these three skills, PiRA represents an appropriate primary outcome measure.

The assessments will be administered online by schools with the process overseen by AlphaPlus. The instrument will be administered by teachers. All Year 5 students, that have not withdrawn and who were enumerated in the initial recruitment process, will be required to complete a New PiRA reading assessment and FAR (the latter will also be administered at baseline) at the same time. AlphaPlus will be blind to the intervention/control status of the pupils and schools; however, the teachers will clearly be aware of which group the school has been assigned to.

The decision for teachers to administer the PiRA was made in order to contain trial costs as well as keeping disruption within schools to a minimum. If independent Test Administrators were to be sent into schools, this would increase costs significantly. It would also present logistical challenges for teachers in terms of allocating set times for them to carry out the PiRA test. Because the test is conducted online, it is envisaged that most classes will need to go into a computer suite to take this test. Often such suites are in demand within school and so there may only be one or two slots a week when Year 5 classes would be able to use this space to take the test. It is likely that it would prove difficult to coordinate availability of Test Administrators with times when Year 5 classes are able to take the test without disrupting schools' usual timetables, especially in multi-form entry schools.

The risks involved with teacher administration, however, need to be considered carefully and actions taken to mitigate them introduced as far as possible (EEF, 2019). There are two key risks: non-standardisation of instructions/delivery of the test; the risk of teachers influencing pupil outcomes on the test either purposefully or inadvertently. To protect against the first risk, a number of measures will be taken to ensure that teachers are very clear on how to administer the PiRA test. We will develop clear step-by-step instructions for teachers (to be shared with schools via AlphaPlus) about how to support pupils to access and complete the test. As well as written guidance, we will also provide teachers with a short video that they will show to pupils just before they complete the test. This video will include pointers for pupils in relation to how to avoid any potential practical issues when completing the test, e.g., accidently skipping a page (on the screen). The video will be informed by our experience of using the online PiRA test within the PALS-UK evaluation (Lewin, et al, forthcoming 2024). It is also important to note that teachers routinely deliver assessments to children as part of their everyday work, and are therefore likely to be experienced at administering tests carefully, and will have a good understanding of the need for consistency in administration processes. The need for all schools to follow the detailed guidance in relation to testing will be communicated carefully and reinforced by AlphaPlus when they communicate with schools about the baseline and endline assessments.

To minimise the likelihood that teachers influence pupil outcomes, we will make the importance of pupils completing the test independently clear within the teacher guidance. We will make it clear in the written guidance for teachers and in the video for pupils when it is (and is not) appropriate for the teacher to assist a pupil when doing the test. To quality assure the administration of the PiRA test, members of the evaluation team will ask the six case study schools (see implementation and process evaluation for further details) to conduct the PiRA test at a time that coincides with our second case study visit (end of May 25). This way we will be able to pilot the use of the teacher guidance/pupil video with these six schools and make

any tweaks to the guidance/video based on how effectively they support standardisation of the delivery of the tests.

The PiRA tests are scored automatically using the online tool, so there is no potential for bias to influence scoring. 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 info/identify key details;
- Summary – summarise main ideas from 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.

Using the KS2 Year 5 New PiRA assessment for the Summer term, a student can achieve a maximum score of 45 and a minimum score of zero.

The effect of Reading Plus on reading attainment will be obtained from a multiple regression model where the pupil score from the PiRA Year 5 summer term test will be the dependent variable. The multiple regression will contain covariates including a measure of prior attainment. The prior attainment measure will be KS1 reading scaled score obtained directly from schools at enumeration. In detail, schools will be asked to provide KS1 reading raw and scaled scores as well as teacher assessed grade for reading at KS1. It is proposed that KS1 reading score is used as a covariate in the impact analysis. We prefer using KS1 scores over the teacher-assessed grades corresponding to KS1 reading outcomes, as KS1 scores offer finer distinctions and greater precision in measurement, as well as provide greater statistical power compared to the categorical variable held in the National Pupil Database. However, at this stage, we want to leave open the possibility of collecting KS1 teacher assessed grades from the NPD or directly from schools. This is as a fail-safe, if it turns out that schools cannot supply either the raw, scaled or grade to us directly in sufficient numbers. If this turns out to be the case we can revert to the NPD. Whilst collecting scores directly from schools is convenient and speeds up the reporting process, avoiding the need to access the ONS SRS, there is a risk of missing information. Collecting scores from via NPD ONS is more time-consuming and less flexible but does offer complete information with no missing values (or close to no missing values). For this reason, we leave open the possibility of accessing the NPD data for pupils in the trial via the ONS SRS and have ensured that the wording of all data protection statements and information sheets reflect this.

Secondary outcomes

Silent reading fluency, reading comprehension and reading vocabulary

Secondary outcome measures of silent reading fluency, reading comprehension and reading vocabulary will be taken using the Kaufman Test of Educational Achievement (Third Edition)

(KTEA-3)[7]. These measures will be used as secondary outcomes and mediators. KTEA-3 is a reliable, valid measure (Breaux & Lichtenberger, 2016), whose subtests provide robust observations on three secondary outcomes of central interest within the logic model.

The three KTEA-3 subtests take approximately 30 minutes to administer in total. These assessments will be conducted with 10 preselected students per school (selected at random prior to randomisation), scheduled for June/July 2025. The KTEA-3 measure will be administered by trained assessors provided by AlphaPlus.

All Test Administrators will receive in-depth face-to-face training on how to deliver the three KTEA-3 subscales. This training will include guidance on safeguarding considerations. We will ensure that all administrators complete a DBS check and are aware of the recently published guidance in relation to conducting research with children and young people[8]. This guidance recommends that researchers should not be alone with children even if they have a DBS check. This new requirement will be communicated to schools by AlphaPlus when arranging visits to collect KTEA-3 data. Alpha-Plus will make suggestions to schools about how one-to-one testing might be managed in light of this guidance, e.g., testing might take place in a separate room with a TA or teacher present (perhaps a teacher who is marking in the staff room, etc) or in a corridor/office in view of other staff. The training will also include clear directions in terms of what to do if Test Administrators have any safeguarding concerns while in schools, e.g., if a pupil discloses something to them that is of concern.

Reading self-efficacy and reading motivation

The final secondary outcome measures consist of reading self-efficacy and motivation for reading, which will be assessed jointly using a modified version of the Feelings about Reading questionnaire. Administering this questionnaire to the entire class will take approximately ten minutes. The questionnaire is divided into two parts: the first part, which measures reading self-efficacy, comprises 20 items, while the second part, measuring motivation for reading, 10 items. The Likert scale structure (seven points) of the questionnaire yields a possible score ranging from 20 to 140 for self-efficacy and from 10 to 70 for the motivation subscale. The motivation to read scale, developed by Vardy et al. (in prep), is grounded in self-determination theory (Deci and Ryan, 1985) and has demonstrated high reliability (Cronbach's α = .83). The reading self-efficacy scale, based on self-efficacy theory (Bandura, 1993), is adapted from Carroll and Fox's (2017) original version with minor phrasing revisions. This scale has a Cronbach's α value of .90 (Vardy et al., in prep).

This instrument will be administered online by teachers, and overseen by AlphaPlus at both baseline and endline to all pupils. Similarly to the PiRA administration, a step-by-step guidance will be provided for teachers in relation to how to deliver the FAR questionnaire, including training items and a script to be read out to pupils (for more details, see the primary outcome section). This guidance will be informed by our experiences of using the FAR for the PALS-UK trial (Lewin, et al, forthcoming 2024).

[7] Kaufman Test of Educational Achievement, Third Edition (KTEA-3) (pearsonclinical.co.uk)

[8] Research with children and young people - User Research Manual - Department for Education

Sample size

To arrive at a required sample size, we first need to determine an effect size that we wish the trial to be 'powered' to detect. To arrive at such an effects size, we consider results obtained from other studies of Reading Plus as well as studies of other similar reading programmes targeting primary school children. These studies give us an indication of what effect sizes it might be reasonable to expect and thus a means of judging the adequacy of our proposed sample.A randomised trial of Reading Plus in the US in the fifth grade (equivalent to Year 6), yielded effect size ES= 0.18 ($p < 0.001$) (Spichtig et al., 2019). An evaluation of Lexia Reading Core, in England, but with younger pupils, obtained an effect size on the primary outcome of ES=0.08 ($p=0.15$) for all pupils, and ES=0.18 ($p=0.04$) for FSM pupils (Tracey, L et al., 2022). There are a number of US studies of PALS that found quite large effect sizes, ranging from 0.23 to 0.71, but samples contained only small numbers of fifth grade pupils and focused on students with learning disabilities (U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse, 2012). The Reciprocal Reading programme looked at reading attainment in KS2 (O'Hare et al., 2019). The authors obtained ES=0.14 ($p<0.01$) and ES=0.18 ($p<0.001$) for reading comprehension. Together, these are the most relevant studies for which high quality evidence of effect sizes are available and based on this we judge that our trial sample should be capable of detecting an effect size of around ES=0.18 with 80 per cent power. According to (Kraft, 2020 Table 2, p. 250) this would rank as medium sized effects in contemporary education trials.

Table 2 below presents a range of minimum detectable effect sizes (MDES) for the sample as a whole and for the FSM subgroup, assuming different level of sample attrition. The MDESs are the smallest effect sizes which if true will yield tests for statistical significance at the 95 per cent level in which the null is rejected in 80 per cent of trials given the sample sizes proposed (and other assumptions).

Table 2: Sample size calculations

		Overall	FSM	Overall	FSM	Overall	FSM
School level attrition rate		**No attrition**		**5% attrition**		**10% attrition**	
Minimum Detectable Effect Size (MDES)		**0.18**	**0.20**	**0.18**	**0.20**	**0.19**	**0.21**
Pre-test/ post-test correlations	level 1 (pupil)	0.36	0.36	0.36	0.36	0.36	0.36
	level 2 (class)	0.10	0.10	0.10	0.10	0.10	0.10
	level 3 (school)	0.10	0.10	0.10	0.10	0.10	0.10
Intracluster correlations (ICCs)	level 2 (class)	0.10	0.10	0.10	0.10	0.10	0.10
	level 3 (school)	0.04	0.04	0.04	0.04	0.04	0.04
Alpha		0.05	0.05	0.05	0.05	0.05	0.05
Power		0.8	0.8	0.8	0.8	0.8	0.8
One-sided or two-sided?		Two	Two	Two	Two	Two	Two
Average cluster size for level 1 (per level 2 unit)		30	8	30	8	30	8
Average cluster size for level 2 (per level 3 unit)		1.7	1.7	1.7	1.7	1.7	1.7
Number of schools	Intervention	63	63	60	60	57	57
	Control	63	63	60	60	57	57
	Total	126	126	120	120	114	114
Number of pupils	Intervention	3,213	857	3,060	816	2,907	775
	Control	3,213	857	3,060	816	2,907	775
	Total	6,426	1,714	6,120	1,632	5,814	1,550

The MDESs are calculated using the PowerUp software (Dong & Maynard, 2013).

The following assumptions are made in completing our calculations:

- Statistical tests of the nil null hypothesis will be performed at the 95 per cent level of statistical significance with statistical power of 80 per cent.
- Randomisation of schools 1:1 to intervention and control condition
- Total proposed sample at recruitment of 126 schools, with an average of 1.7 classes per school and 30 pupils per class (as per estimates from the PALS-UK trial; project report forthcoming[9]).
- Proportion of variance explained in the outcome by covariates of 0.36. We do not have reliable evidence on the correlation between KS1 reading attainment (the main pre-test covariate) and PiRA Year 5 summer test (the primary outcome). We do know that the correlation between PiRA and KS2 reading attainment is about 0.75[10]. Thus, we assume it is lower for KS1. We also allow for some improvement in the outcome variance explained at class and school levels of 10%, assuming that the regression model from which impact estimates

[9] See Ainsworth, S et al. (2022), Gellen, S and Morris, S.(2023) and (Lewin, et al, forthcoming 2024)

[10] See https://www.risingstars-uk.com/subjects/assessment/rising-stars-pira-tests/correlation-study

will be obtained will be a three-level model accounting for clustering of pupils by school and class.

- Based on evidence from the recently completed PALS-UK trial we have assumed intra-cluster correlations (ICC) of 0.10 at the school level and 0.04 at the class level (Lewin, et al, forthcoming 2024)
- Given the trial's relatively short duration and the Delivery Team's confidence in retaining schools within the intervention, we structured the study with the expectation of minimal school attrition. With this in mind, we aimed for 126 schools, projecting the retention of 120 schools, and powered the study to detect an effect size of 0.18.

For the full sample assuming no attrition we obtain MDESs of 0.18 and for the FSM subgroup ES=0.20. Table 2 also presents two alternative scenarios, accounting for a 5 and 10 percent school attrition rates. In the first alternative scenario, considering 5 percent school attrition, the MDES for the total sample remains ES=0.18, and for the FSM subgroup it is ES=0.20. Accounting for 10 percent attrition rate, MDES increases to ES=0.19 and ES=0.21 respectively.

Based on our assessment of the existing literature, the proposed sample sizes described here should be adequate to detect a standardised difference in means between the two groups of around ES=0.18 with close to 80 per cent power, for the primary outcome.

Randomisation

By the end of the summer term in 2024, we will have verified the identities of every participating school and the pupils within them. Additionally, we will have gathered baseline KS1 reading test scores, along with reading self-efficacy and motivation for reading measures for each pupil.

Prior to randomisation, a Research Associate at MMU will draw at random a named-pupil-sample once the enumeration period has ended to whom the KTEA-III will be administered at endline. A set sample of 15 pupils per school will be selected with Test Administrators asked to complete 10 assessments working through the list of names arranged randomly in strict order. The sample will be chosen as follows. First, a list of retained eligible students will be arranged by class and by school. Then for single form entry schools:

- Each child in the year group will be given a random number from a uniform distribution (a random number seed will be set in advance and store separately); and

  Students will be arranged in ascending order by this number, with the top 15 students sampled for assessment.

Two or more form entry schools:

- A single class will be sampled at random from a list of all Year 5 classes;
- From the sampled class, each child will be given a random number from a uniform distribution; and
- Students will be arranged in ascending order by this number, with the top 15 students sampled for assessment.

Randomising schools to treatment and control conditions will only take place once these tasks are complete.

Turning to randomisation itself, recruited schools will be divided into terciles based on their average score in KS2 Reading NCTs over the last three May tests completed by the school, and within these groups by whether they are single form entry or two or more form entry. This means that our sample will likely be better balanced on school average prior attainment and school size than would be the case under complete or simple randomisation

Randomisation will be carried out using the STATA user-written command `randtreat` (Carril, 2017). The programme enables us to stratify the randomisation as described and deal with the problem of 'misfits'. Misfits occur when the number of schools in a stratum is not a multiple of the number of groups. The program enables us to balance across the sample as a whole using a 'global' approach to dealing with misfits.

The KTEA-III sampling and randomisation process will be carried out by the same Research Associate at MMU, using STATAv18 statistical software. The Research Associate will be blinded to the school identities during the randomisation process. The Research Associate will communicate the result of randomisation to Reading Solutions, who will in turn inform schools of the outcome.

Statistical analysis

Primary analysis

The primary outcome or response variable is continuous and normally distributed. Thus, an intention to treat sample estimate of the effect of Reading Plus on reading attainment can be obtained from a linear mixed model of the following form:

$$Y_{ijk} = \beta_0 + \beta_1 T_k + \beta_2 X_{ijk} + \beta_3 birth_{ijk} + \beta_4 S_k + v_k + \varsigma_{jk} + \varepsilon_{ijk}$$

Y_{ijk} is the raw reading attainment (PiRA) score for pupil i in class j and school k. T_k is a binary variable coded to '1' if school k is assigned to the intervention '0' otherwise. The sample estimate of β_1 is the estimated treatment effect of Reading Plus. X_{ijk} is the KS1 reading test scaled score for pupil i in class j, from school k (this covariate is entered as a pupil level covariate and will reduce variance explained at all three levels – school, class and pupil); $birth_{ijk}$ the month of birth for pupil i in class j and school k; and S_k captures the variables used for stratification. There are random effects at the school v_k and class levels ς_{jk} as well as a pupil level residual ε_{ijk}. The three random effects are assumed to be uncorrelated, normally distributed, with mean zero and conditional variances σ_v^2, σ_ς^2 and σ_ε^2. The total conditional variance of the response is $\sigma_T^2 = \sigma_v^2 + \sigma_\varsigma^2 + \sigma_\varepsilon^2$, and therefore the school level ICC (level 3) is $\rho_3 = \sigma_v^2/\sigma_T^2$ and class level (level 2) ICC $\rho_2 = \sigma_\varsigma^2/\sigma_T^2$.

We opted to incorporate the age covariate because, unlike many other EEF trials, we use outcome measures that are not standardised for age. This choice aligns with the methodology employed in the PALS trial (Gellen & Morris, 2023), where, in agreement with EEF, age was controlled for in the primary and secondary analyses.

Estimates of the model parameters will be obtained using restricted maximum likelihood and the command 'mixed' in STATA v18. The sample estimate of the intervention effect from the

model above will be converted into an effect size, consistent with Hedge g, and is described below.

Secondary analysis

Secondary analysis will involve estimating the effects of Reading Plus on the secondary outcomes for silent reading fluency, comprehension, and vocabulary. Response variables for each of the three models are derived from the KTEA-3 instrument that is administered to a subsample of pupils in each school chosen at random prior to randomisation. Thus, sample estimates of the effect of Reading Plus on pupils for these three outcomes will be obtained from a model of the following form, where the usual assumptions are made:

$$Y_{ij} = \beta_0 + \beta_1 T_j + \beta_2 X_{ij} + \beta_3 S_j + \varsigma_j + \varepsilon_{ij}$$

Here the variable Y_{ij} is the response – either pupil $i's$ silent fluency, comprehension or vocabulary score in school j – depending on which of the three models is considered. X_{ij} is the KS1 reading score for pupil i in school j.

Secondary analysis will also involve estimating the effects of Reading Plus on reading motivation and reading self-efficacy. Measures of these two constructs will be obtained from the Feelings about Reading questionnaire, administered to the whole sample at both baseline and endline. Sample estimates for the effect of Reading Plus on both motivation and self-efficacy will be obtained from estimating two regression models (one for each outcome) similar to that used in the primary analysis, except where Y_{ijk} is either the self-efficacy score for pupil i in class j and school k or similarly a pupil's motivation score – depending on which of the two models is considered. In these models X_{ijk} is the baseline score for either self-efficacy or motivation depending on the response.

Estimation of effect sizes

The sample estimate of the intervention effect from the primary outcome regression model above will be converted into an effect size. This will be achieved through first fitting a simple variance components model:

$$Y_{ijk} = \beta_0 + v_k + \varsigma_{jk} + \varepsilon_{ijk}$$

From this model we will obtain sample estimates of the unconditional variances for v_k, ς_{jk} and ε_{ijk}, namely, $\hat{\sigma}_v^2$, $\hat{\sigma}_\varsigma^2$ and $\hat{\sigma}_\varepsilon^2$. The sample estimate of the effect size is then calculated as:

$$ES = \frac{\hat{\beta}_1}{\sqrt{\hat{\sigma}_v^2 + \hat{\sigma}_\varsigma^2 + \hat{\sigma}_\varepsilon^2}}$$

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

The same approach will be used to compute effect sizes for all secondary outcomes. All effect sizes will be consistent with Hedges' g.

Sub-group analyses

For two subgroups - pupils entering Year 5 in participating schools that are ever-FSM, and pupils designated SEND at baseline - we will conduct separate analyses.

In accordance with the EEF statistical guidelines, we will conduct subgroup analysis using both the restricted sub-sample and interaction approaches. In the case of the former, this entails applying the primary analysis model to restricted subsets of pupils categorised as 'SEND', 'non-SEND', 'EverFSM', or 'not EverFSM'. As for the latter, we will employ a model similar to the one used in the primary analysis. This model will include the subgroup binary indicator along with an interaction term that interacts the subgroup indicator and treatment allocation dummy variable. The results will be presented in terms of effect sizes along with their corresponding 95% confidence intervals.

Analysis in the presence of non-compliance

In this trial, due to the nature of the intervention, we can discount the possibility of control group members accessing Reading Plus. Thus, non-compliance in this case refers to pupils in schools randomised to the intervention failing to use Reading Plus in any meaningful way.

There are a number of compliance measures that it is possible to derive for Reading Plus. These range from quite conservative measures to measures that are less stringent. The compliance measures considered were as follows. An indicator for each of these measures can be derived from the Reading Plus management information for each pupil and school:

1. The school fails to attend training in Reading Plus;
2. Definition 1 or the school attends training but there is no evidence that the school subsequently uses the Reading Plus platform;
3. Definition 2 or a school attends training and there is evidence that Reading Plus is used in the school but there are pupils who have engaged with fewer than 15 texts over the three terms the programme runs for.
4. Definition 3 or pupils engage with 15 or fewer texts over the first two terms and no texts in the third term.

For the purposes of our compliance analysis, we will used both definitions 3 and 4. The first two measures are seen as overly restrictive in the sense that although schools might engage with training and use the platform, some pupils may, even so, minimally engage with little consequence for their reading engagement. The last two measures were developed following detailed conversations with Reading Solutions. It was judged that engaging with fewer than 15 texts would have little influence on pupils' reading. There was some concern that pupils might engage with 15 texts in the final term of implementation just before they complete reading assessments, and that such minimal engagement in close proximity to assessment could possibly influence their subsequent performance in tests. Thus this final definition will be used as a sensitivity check.

The Reading Plus software will track usage, enabling the Delivery Team to share compliance data. No records are anticipated from students in schools that did not participate in the training (definition 1) or have yet to implement the intervention (definition 2). Consequently, criteria 3 and 4 should encompass the initial two criteria. However, to ensure data accuracy, we will request Reading Solutions to gather details on training attendance, and these logs will be cross-referenced with actual pupil records.

A complier average causal effect (CACE) will be estimated though the construction of a binary compliance variable at the pupil level (1=complier, 0=non-complier). The approach to compliance analysis will be specified in detail in the SAP.

Additional analyses and robustness checks

In additional for fitting models above, the primary analysis will also include the reporting of results from fitting the following additional models. For each of these models, reading attainment will form the response variable:

- **Simple model containing the intervention dummy variable only.** By comparing the results of this model with the results from the primary analysis we can examine the effects of covariates on sample estimates.
- **A model containing intervention dummy, month of birth and stratifiers**. This model contains all 'design' related covariates so that when compared to the primary model, the influence of design features can be assessed.
- **A model as immediately above but including additional covariates**: KS1 reading test score, FAR self-efficacy and motivation test scores, FSM, gender and EAL. This final model contains a full set of covariates available to us. We fit this model to explore whether any further gains in power can be achieved through adding these additional covariates to the model.

Analysis that explores causal processes through mediation analysis is also proposed. In particular, we wish to examine 'fluency' as a key mediator of the intervention effect. Gains in silent reading fluency are hypothesised as a key process or mechanism through which Reading Plus leads to improvements in reading attainment. In other words, gains in reading attainment for intervention group pupils over the control group, after exposure to Reading Plus, are either partially or fully mediated through changes in reading fluency.

In counterfactual language we are interested in the natural indirect effect (NIE). Using the definition of NIE from Imai et al. (2013) we can define the NIE in our case as representing the causal effect of Reading Plus on reading attainment transmitted through changes in a student's reading fluency following their receipt of Reading Plus. We will obtain a sample estimate of the NIE from the data using the command mediate in STATA v18. Given that the mediator – Fluency obtained from the KTEA-3 instrument – and the outcome – reading attainment obtained from the Year 5 Summer Term PiRA - are both continuous normally distributed variables both regressions will be estimated using linear models (further details in the SAP).

Missing data analysis

For the primary analysis, sensitivity tests will be carried out to assess whether missing data at endline leads to biased or imprecise estimates of β_1. Missingness that occurred before randomisation is unlikely to cause bias in estimated treatment effects but can result in diminished sample sizes. For the primary analysis potential sources of missingness after randomisation are likely to include:

- Parents requesting that their children be removed from the study, and their data deleted
- Pupils leaving the school prior to the completion of the New PiRA Summer 5 Test

- Schools withdrawing from the evaluation and requesting all data supplied by them to be deleted
- Pupils not present on the day of the New PiRA Summer 5 Test and unable to supply outcome data
- Schools excluding pupils from testing for idiosyncratic reasons

In the first screening stage, we will examine the type of missingness: i.e., whether data is missing completely at random (MCAR), missing at random (MAR) or potentially missing not at random (MNAR). This includes calculating and comparing the rate of missing data in the trial arms. If we find the level of missingness to be problematic – i.e., missingness exceeds five per cent in both control and treatment groups – we will assess if available baseline covariates explain missingness. We will fit a logistic regression model where the dependent variable captures in binary form whether a pupil provides a valid PiRA score at endline. This model will contain the full set of covariates available to us at baseline and is referred to as a 'drop out' model. Covariates found to be significantly associated with missing PiRA scores (with a 95 per cent confidence interval) will be considered explanators of the presence (or absence) of the endline observation on the primary outcome.

If missing data on the PiRA test at endline appear to exceed 5 per cent in anyone arm of the trial and evidence from the drop out model appears to indicate missingness associated with included covariates, further sensitivity tests will examine the consequences of missing data in the primary outcome for the sample estimates of intention to treat in the primary analysis using multiple imputation.

We will use multiple imputation using chained equation (mice) to impute missing values for each variable so affected in our analysis using a fully conditional specification. To perform multiple imputation we will use the `mi impute chained` and `mi estimate` commands in STATA v18.

STATA v18 only permits single-level imputation for missing values. Multilevel imputation is available in R statistical software but only for two-levels and the procedure is very slow taking many hours to run. For reasons of practical necessity and to ensure consistency, we will first re-esimate the primary outcome model described in the primary analysis section above obtaining the model parameters using ordinary least squares (OLS) with cluster robust standard errors[11]. This is equivalent to estimation of a population average treatment effect as opposed to a cluster-specific effect obtained in the more typical linear mixed (or multilevel) model applications. The linear OLS model will act as a benchmark against which we will compare results from fitting a similar model on the data sets created through the multiple imputation procedure.

The `mi impute chained` procedure will create a number of data sets with missing values filled-in. Using `mi estimate` we will then re-estimate the effect of the intervention through running an linear ordinary least squares (OLS) regression in the same form as the primary analysis discussed in the proceeding paragraph but on the imputed data sets. The two sets of results - the analysis based on imputed data sets and the re-run of the primary analysis - when compared will enable us to determine whether missingness follows a MAR process. Our

[11] We will use CV2 standard errors with degrees of freedom adjustment. The STATA command is `vce(hc2 [cluster_name], dfadjust)`. This model will take account of clustering at the school level.

assumption is that what ever conclusion these analyses lead to would also hold if multiple imputation had be conducting in a multilevel setting, which based on our experience in previous studies is plausible (i.e., population average and cluster specific intervention effect estimates tend to be similar and relying on one rather than the other does not lead to marked changes in interpretation).

There is a possibility that attrition does not follow MAR or MCAR; instead, missingness is MNAR. Assessing the extent to which missingness might be MNAR and making some assessment is extremely difficult. We can, however, undertake sensitivity analysis that is agnostic toward the precise processes of missingness. These analyses take the form of placing bounds on the treatment effect estimate derived from certain assumptions about missing responses.

Lee (2009) introduces an estimator for treatment effect bounds that can be used to assess the possible consequence of non-random attrition by treatment groups. Lee's approach relies on minimal assumptions; namely that treatment is randomised and a monotonicity assumption. The monotonicity assumption holds that no subjects are more likely to respond to data collect under control conditions than they would be under intervention conditions (Gerber & Green, 2012). In the presence of a notable imbalance in attrition levels, we can used 'Lee bounds' to provide upper and lower bounds on our estimates. The one slight draw back in using Lee's approach is that the bounds are calculated for the always-responding subsample and not for the full as analysed sample. That is, those subjects that will always response to requests to complete assessments regardless of whether they are assigned to intervention and control conditions. Bounds following Lee's approach will be estimated using the package `leebounds` in STATA v18 (Tauchmann, 2014).

Implementation and process evaluation (IPE) design

Research questions

The IPE aims to explore the relationship between the delivery of the programme and any impact of Reading Plus on pupil outcomes (or not) within the IE and is grounded in the theory of change (Humphrey et al., 2016). The IPE will be focused around the research questions below. Insights from the IPE will complement and be integrated with the findings from the IE as outlined under the associated research questions.

RQ1. What reading practices/programmes/interventions are:

a) Happening in control schools?

b) Happening alongside Reading Plus in intervention schools?

c) Being replaced by Reading Plus within intervention schools?

Exploration of RQ1 will provide information about the extent to which practices in control/intervention schools were differentiated, which might help to explain any significant difference (or lack of difference) in attainment between the two groups.

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

a) How often pupils use the Reading Plus platform?

b) The way in which additional visual skills activities are used to support struggling readers?

c) How often and for what purposes teachers use the additional activities to support individuals/small groups?

d) Use of data reported within Reading Plus by teachers to identify and address skills gaps?

e) e) optional use of Reading Plus at home by pupils?

RQ3. What changes or adjustments, if any, were made to the programme during its implementation?

a) Why were these adjustments made?

RQ4. What are the barriers and facilitators to the implementation of Reading Plus?

a) Which barriers and facilitators, if any, were specifically or disproportionately experienced by FSM-eligible pupils? SEND pupils?

Exploration of RQ2, 3, and 4 will provide information about the extent to which the intervention was delivered as intended and any facilitators/barriers affecting delivery. Again, this information will be helpful in explaining the potential impacts of Reading Plus, in particular any variance found across schools or across different pupil groups.

RQ5. To what extent did the support available within the Reading Plus programme enable practitioners to deliver the programme effectively? Including:

a) Initial Reading Plus training

b) Support from the Reading Consultant

c) Optional additional training

d) Access to the resources hub

RQ6. What impacts (if any) do headteachers, teachers and pupils perceive Reading Plus to have on:

a) Fluency?

b) Comprehension?

c) Reading proficiency?

d) Feelings about reading?

e) Closing the 'reading gap'?

f) Teacher workload?

g) Teachers' understanding of how silent reading fluency impacts on fluency, stamina and comprehension?

Investigation of RQs 5 and 6 will allow in-depth exploration of how school staff, pupils and Reading Plus consultants perceive the different elements of the programme. This will allow us to evaluate the different predictions made by the logic model in relation to how the programme works. This data will help us to also explain the IE results by suggesting mechanisms for any observed gains (or lack of) in the various constructs within the logic model (e.g., silent reading fluency, reading motivation, etc.).

RQ7. To what extent can formative assessment data collected within the Reading Plus system be used to predict:

a) Reading proficiency (KS2 SATs)?

b) Fluency?

c) Comprehension?

d) Vocabulary?

RQ8. What evidence is there to support the causal assumption that any observed gains in reading proficiency in the intervention group are mediated by gains in:

a) Fluency?

b) Vocabulary?

c) Comprehension?

d) Feelings about reading?

RQ7 and 8 will use quantitative data from the Reading Plus system to directly explore whether the causal mechanisms proposed to operate between the different reading-related competencies can be observed.

Research methods

The IPE research questions will be investigated using a range of qualitative and quantitative methods as detailed below and summarised within Table 3. Together these questions are designed to cover 11 of the 12 IPE dimensions set out in the EEF guidance (EEF, 2022). The final dimension, cost, is addressed separately within the Cost Evaluation Design.

Key IPE Activities

Before the intervention, survey data will be gathered from headteachers and Year 5 teachers. The surveys will be designed to provide information about current reading practices and other relevant interventions that are already taking place in school. This will include questions about what methods teachers use to monitor reading attainment and how frequently reading is assessed. Headteachers and teachers will also be asked about any existing use of EdTech to support reading and other areas of the curriculum. The teacher survey will include some additional questions about teachers' knowledge about the teaching of reading. The evaluation team will also observe the two online training sessions and review the delivery documentation to become familiar with the expectations and guidance relayed to schools.

During the intervention, Steph Ainsworth (Co-PI), Cathy Lewin (Co-I) and Kate Wicker (RA) will interview headteachers, Year 5 teachers and any other relevant teaching staff (e.g. teaching assistants in Year 5, literacy subject leaders) within six case study (intervention) schools (see below for details in relation to case study selection). We will visit each of the case study schools twice: once in the middle of the intervention and once towards the end. We will ask headteachers and Year 5 teachers about what reading practices are happening alongside Reading Plus, and which practices have been replaced by the programme; their experiences of using Reading Plus; and any impacts which they perceive the programme to have.

We will also ask pupils about their experiences of engaging with Reading Plus. In addition to these activities, at the first case study visit we will also observe a Reading Plus lesson and conduct an interview with the reading consultant allocated to the school (this will happen online but around the time of the first visit). Throughout the intervention the Reading Plus platform will automatically collect formative assessment data in relation to pupils' reading comprehension, vocabulary and fluency. This longitudinal data will be collated at the end of the trial.

After the intervention, we will send out endline surveys to teachers and headteachers in both intervention and controls school. This will include questions about any changes in reading practices since the start of the intervention. In the intervention schools we will also ask staff about their experiences of using Reading Plus; any enabling and constraining factors which influenced how it was implemented; and their perceptions of any impacts that the intervention has had.

Development of IPE instruments

The surveys, observation and interview schedules will be developed in consultation with the delivery team. The survey questions and interview questions will be adapted from those used within the EEF evaluation of PALS-UK (which is also a reading intervention) (Lewin, et al, forthcoming 2024). Adaptations will be made to ensure that the instruments used align with the logic model for Reading Plus. For example, while PALS-UK involved pupils developing oral reading fluency by reading whole books with support from a partner, Reading Plus involves pupils reading online texts silently and independently. The instruments will therefore include questions about pupils' percpetions of the advantages and limitations of the online individualised silent reading approach adopted by Reading Plus, as well as the other specific characteristics of this intervention (e.g., the use of the Guided Window, personalised recommendations, etc.). The adaptation/refinement of the instruements used for PALS-UK will also allow us to take into account 'lessons learned' from the previous trial, e.g., in terms of clarity of wording, etc. Surveys will be administered online and take around 15 minutes to complete.

In addition to the surveys, observation and interview schedules, we will also use the Reading Plus platform itself as an instrument for gathering data about usage and pupil progress. We will download data from the platform, which will provide an indication of how frequently each pupil used the platform (i.e., number of times per week), the average amount of time spent on the platform each week, and the total amount of time that each pupil spent engaging with the platform over the intervention period. As a measure of pupil progress, we will collect the following data from the platform at regular intervals over the intervention period (October, December, February, April, June):

- Comprehension level
- Vocabulary level
- Reading rate
- Reading self-improvement belief
- Reading confidence
- Reading interest

We will also collect pupils' summative scores for from the platform at the beginning and end of the intervention period (October and June). These will be cores relating to the constructs listed above with the addition of an overall reading proficiency measure (which the platform only provides within the summative assessment).

Case study selection

Six schools will be selected as case studies. A sample of six schools should achieve a balance in terms of sampling a sufficient range of contexts while keeping data collection feasible. Case studies will be selected at random stratified by: school size (one-form per year group versus two or more forms per year group) and average score in KS2 Reading NCTs over the last three May tests completed by the school (divided into terciles). This will ensure representation across diverse contexts. Since engagement as a case study schools is more involved than as a non-case study school, it is possible that some schools who are invited to be a case study

may choose not to take part in this aspect of the trial. If this is the case, schools will be replaced until a sample size of six is met .

Analysis

Quantitative data from the surveys will be analysed in SPSS statistical software generating descriptive statistics. Qualitative data from the surveys and interviews will be analysed using NVivo and thematic analysis (Braun & Clarke, 2006; Braun & Clarke, 2021) with a mixed coding method. A coding framework derived from the logic model will be applied deductively and additional themes derived inductively.

Quantitative analysis of the platform usage data and pupil progress measures generated within Reading Plus will be conducted to investigate statistical associations (using multilevel multiple regression analyses) between frequency of use on PiRA scores and FAR/KTEA-3 subscale scores for intervention group pupils. These analyses will help us understand mechanisms of change (RQ8) and therefore will support us in exploring explanations for the findings of the IE. They will also support us in testing the prediction within the logic model that the formative assessment data collected within the Reading Plus will provide useful monitoring information for schools (i.e., we will test whether it predicts relevant summative assessments) (RQ7). Analysis of the reliability of the pupil progress measures will also be conducted.

Findings from the case study interviews and observations, and the surveys will also aid interpretation of the impact analyses, providing the opportunity to develop further hypotheses around possible mediators and sources of heterogeneity in treatment effects. The integration of the quantitative and qualitative data will take a number of forms, including exploration of the qualitative data in search of possible explanations for the quantitative findings. For example, if the IE reveals that Reading Plus leads to greater gains in reading attainment relative to the control condition, the qualitative interview and survey data (in relation to RQ6) will potentially provide data in relation to why this might be the case, i.e., stakeholders' perspectives on why it is effective. Teachers', headteachers' and pupils' perspectives on specific mechanisms that might have driven any observed change will then be compared with the quantitative progress data collected from the Reading Plus platform as part of the IPE (in relation to RQ8 – data about children's progress in fluency, vocabulary, comprehension and feelings about reading) and the secondary outcome data collected endline (summative measures of these same constructs). This will allow us to see if there is convergence between what staff and pupils believe the mediating mechanisms of Reading Plus are and what the quantitative data suggests is happening. There are numerous other ways in which quantitative and qualitative data will be compared to look for evidence of triangulation and points of divergence. For example, qualitative interview data will provide rich data about how staff experienced Reading Plus; while quantitative data from the headteacher/teacher surveys will allow us to gain a sense of how representative individual views expressed within the case studies were across the sample as a whole, e.g., in relation to the differential impact of Reading Plus on particular groups of children.

Table 3: IPE methods overview

RQ addressed	IPE dimension	Research methods	Data collection methods	Sample size and sampling criteria	Data analysis methods
RQ1	Programme differentiation; Monitoring of control/ intervention groups	**Survey** (pre/post)	Online questionnaires	126 headteachers (all) Approx. 250 Year 5 teachers/teaching assistants (all)	Descriptive statistics (quantitative responses); Thematic analysis (free-text responses)
		Interview	Semi-structured interviews	6 headteachers Approx. 10 Year 5 teachers and teaching assistants Approx. 36 Year 5 pupils (6 case study schools)	Thematic analysis; Deductive coding (based on logic model); Inductive coding (to identify additional themes)
RQ2	Fidelity; Dosage; Responsiveness	**Document analysis**	Review of the Reading Plus training materials	n/a	Basic review of the key content to become familiar with teacher guidance
		Observation	Observation of Reading Plus lessons	6 lessons (case study schools)	Deductive coding
		Assessment data	Collection of formative assessment data and usage data from RP platform	6426 Year 5 pupils (all)	Descriptive statistics; Correlation; Regression
RQ3	Adaptation; Fidelity;	**Interview**	Semi-structured interviews	6 headteachers	Thematic analysis; Deductive coding (based on logic model);

	Reach			Approx. 6 to 10 Year 5 teachers and teaching assistants (case study schools)	Inductive coding (to identify additional themes)
		Survey (post)	Online questionnaires	Approx. 125 Year 5 teachers/teaching assistants (intervention schools)	Descriptive statistics (quantitative responses); Thematic analysis (free-text responses)
RQ4	Context/ moderators; Mediators	**Interview**	Semi-structured interviews	6 headteachers Approx. 6 to 10 Year 5 teachers and teaching assistants Approx. 6 consultants (case study schools)	Thematic analysis; Deductive coding (based on logic model); Inductive coding (to identify additional themes)
		Survey (post)	Online questionnaires	63 headteachers Approx. 125 Year 5 teachers/teaching assistants (intervention schools)	Descriptive statistics (quantitative responses); Thematic analysis (free-text responses)
RQ5	Quality; Moderators; Mediators	**Observation**	Observation of Reading Plus training	Observation of online training before intervention and three months in	Deductive coding
		Interview	Semi-structured interviews	6 headteachers Approx. 6 to 10 Year 5 teachers and teaching assistants Approx. 6 consultants (case study schools)	Thematic analysis; Deductive coding (based on logic model); Inductive coding (to identify additional themes)

		Survey (pre/post for teachers; post for headteachers)	Online questionnaires	63 headteachers Approx. 125 Year 5 teachers/teaching assistants (intervention schools)	Descriptive statistics (quantitative responses); Thematic analysis (free-text responses)
RQ6	Responsiveness; Quality	**Interview**	Semi-structured interviews	6 headteachers Approx. 6 to 10 Year 5 teachers and teaching assistants Approx. 6 consultants (case study schools)	Thematic analysis; Deductive coding (based on logic model); Inductive coding (to identify additional themes)
		Survey (post)	Online questionnaires	63 headteachers Approx. 125 Year 5 teachers/teaching assistants (intervention schools)	Descriptive statistics (quantitative responses); Thematic analysis (free-text responses)
RQ7	Quality; Mediators	**Assessment data**	Collection of formative assessment data from RP platform	6426 Year 5 pupils (all)	Descriptive statistics; Correlation; Regression; Mediation analysis
RQ8	Quality; Mediation	**Assessment data**	Collection of formative assessment and usage data from RP platform	6426 Year 5 pupils (all)	Descriptive statistics; Correlation; Regression; Mediation analysis

Cost evaluation design

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

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

Consequently, the cost evaluation takes the form of a Cost Feasibility analysis, serving as a guide to the affordability of Reading Plus, rather than a comparison between Reading Plus and an alternative intervention. The anticipated categorisation of ingredients included:

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

Costs will be divided into pre-requisites, start-up costs, and recurring costs in accordance with the EEF's cost evaluation guidance (EEF 2023). Program fees calculation will rely on information provided by the Delivery Team. Additionally, cost data will be collected through post-intervention headteacher and teacher surveys participating in the implementation and process evaluation.

Ethics and registration

Ethical clearance has been granted by Manchester Metropolitan University, with the initial submission made on 02.02.2024 via a fast-track route. The approval process involved providing comprehensive information on project design, ethical procedures, participant information sheets, consent/withdrawal forms, Memorandum of Understanding, and privacy notices.

The school recruitment process is structured as follows: the Delivery Team identifies and approaches schools meeting the selection criteria, gathering initial data such as school details, contact information, and Year 5 teachers' names for the 2024/25 academic year. Schools are requested to sign a Memorandum of Understanding, outlining project information, objectives, potential benefits, activity timelines, data protection aspects, and the responsibilities of all parties involved. Additionally, schools must sign a separate Data Sharing Acknowledgement specifying the collection and sharing of personal data among the Delivery Team, Evaluation Team, and the School. A withdrawal notice is issued to parents of Year 4 students, allowing a 2-week response period, with the right to withdraw their child at any time. Subsequently, FFT gathers baseline data from each school. The trial will be registered on the Open Science Framework registry following this protocol being finalised.

Data protection

Manchester Met and Reading Solutions UK are independent data controllers for this project. They make decisions about how and what personal data is used in accordance with the purposes set by the EEF. FFT and AlphaPlus are data processors. The EEF will become data controller for the data once it is archived at the end of the project.

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

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

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

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

All assessment data will be accessed and analysed by the Evaluation team (Manchester Met). For the purpose of this project, some personal data (staff names/contact details) will be shared between Manchester Met and Reading Solutions UK. Manchester Met may also need to access the National Pupil Database as part of the evaluation to access pupils' KS1 assessment data (only in the case of substantial missing data collected from schools).

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

Manchester Met shall ensure that a data sharing agreement is in place as required by the GDPR and DPA.

We will not use pupils' names or the name of the school in any report arising from the research. Pupils will be free to withdraw at any time. So that the processing of personal data relating to the pupils is fair, lawful and transparent we will use a parent information sheet, parental withdrawal form, and a privacy notice for parents agreed with the University's Data Protection Officer. Parents are free to withdraw their child from the evaluation and/or ask to have any or their child's information deleted until August 31st 2025, without giving a reason. Schools can also withdraw from the evaluation at any time, without giving a reason, and ask the school's data to be deleted before August 31st 2026.

The information collected will be used for research purposes only and no information that can identify individuals will be used for any other purpose. Any personal data collected and held by Manchester Met, Reading Solutions UK, FFT and AlphaPlus will be destroyed in accordance with the GDPR when it is no longer required, and no later than 31st July 2026.

Personnel

Delivery team – Reading Solutions UK

Ian Fitzpatrick is the Managing Director of Reading Solutions UK. With over 15 years of educational technology experience, he plays a pivotal role in the EEF trial. Tasked with overseeing trial delivery, he ensures adherence to protocols and trial conditions, guaranteeing schools effectively implemented Reading Plus with fidelity.

Briony Cragg is the PR and Communications Manager at Reading Solutions UK. She will oversees all communication aspects of the trial from a delivery perspective, ensuring that participating schools are fully informed, the trial is conducted with precision, and Reading Plus implemented with fidelity.

Evaluators – Manchester Metropolitan University

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

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

Prof Cathy Lewin, Education and Social Research Institute, Metropolitan University: Cathy Lewin is Professor of Education. She has extensive experience of mixed-method evaluation of school-based interventions, including educational technology and inclusive education. Cathy is the IPE design lead.

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

Dr Kate Wicker, Education and Social Research Unit, Manchester Metropolitan University: Kate is a research assistant with expertise in mixed methods evaluation research. She is responsible for coordinating the case studies, and supporting the implementation and process evaluation.

Risks

Risk	Likelihood	Impact (1-3)	Detail/preventive measure
Recruiting sufficient numbers of schools	Medium	2	• The delivery team will send prospective emails to around 14000, whose details are contained within an existing database (https://www.educationcompany.co.uk/spirit/). This approach will be complemented by adverts on social media and emails to contacts within the delivery and

			evaluation team's existing networks, e.g. partnership schools, local authorities, research schools, etc.EEF will also promote the trial through its website, social channels and networks. • Intervention schools are asked to contribute £500 per school plus VAT towards the costs of the training and access to the platform. This represents a small contribution towards the usual cost of the 1-year licence package, which is on average £4290 + VAT for a single year group. • Control Schools are offered payment of £750 per school if all required data are provided and the Evaluation team is given access to the school in June/July 2025 to conduct reading assessments with Year 5 pupils. • Interested schools that do have mixed age group classes will be added to a recruitment waitlist. Similarly, if schools are participating in another EEF trial and are interested in this project, the possibility will be held open for them to also be added to a recruitment waitlist. • All participating schools will have the year 5 PiRA data shared with them at the end of the trial, which may of use to teachers as the pupils enter year 6
Missing data and sample attrition	High	2	Schools recruited to the trial may decide to withdraw, and this sample loss might both reduce precision of statistical estimates and introduce bias. • The randomised controlled trial model will be explained to school during recruitment. • The value of control schools will be explained in initial discussions. • Financial incentives will be provided to control schools that choose to remain in the trial (£750) • Aim to over recruit in the main trial to allow for some attrition (see sample size scenarios above). • Schools participating in the trial will be offered Reading Plus at a discounted rate once the trial has ended. • Pupils may leave the study independently of their school's decisions on participation, or pupil data maybe of poor quality • Mis-recording of identifying data for pupils, mitigated through carrying out extensive checks on student records prior to randomisation. • Recruitment documentation stresses the importance of the study encouraging children and their parents to remain in the trial. More broadly, where attrition occurs, steps will be taken in analysis to test various assumptions regarding missingness and assess consequences for bias and precision using approaches such as multiple imputation of the estimation of bounds.
Staff shortages and retention in the evaluation team	Medium	2	Research projects spanning extended durations often encounter turnover among research staff. With a substantial pool of suitably qualified and experienced personnel, coupled with adaptable workload and staff management systems, we can consistently maintain adequate staffing and effective project management. Additionally, we will establish procedures such as maintaining a comprehensive variable

			library and log for each data source, as well as implementing appropriate handover processes if necessary.
Poor communications between stakeholders	Low	3	As an integral aspect of our project planning, we intend to conduct frequent meetings with the Delivery Team, especially as we approach significant milestones such as school recruitment and enumeration. Additionally, we will consistently provide our project plans and risk management documents to the developers, fostering alignment between our management processes and those of the project team. This synchronization will enable us to collectively and effectively address any emerging challenges that may arise.
Schools being unable to support supervised testing in schools	Medium	3	The new DfE guidelines in relation to Research with Children and Young People state that researchers should not be left alone with a pupil 1:1 even if they have a DBS. This might cause difficulties as schools are currently experiencing staff shortages and so it is predicted that they will find it difficult to release staff to supervise testing. To mitigate this risk, AlphaPlus test administrators will be flexible in their approach to this, suggesting with schools in advance a number of options for how these guidelines might be met in ways that minimise the burden to schools, e.g., testing could take in a corridor in sight of other staff and pupils; in the staff room, PPA room, or another space where there are other people present (teachers engaged in their PPA activities in the background, office staff, teaching assistants, etc.).

Timeline

Table 4: Timeline

Dates	Activity	Staff responsible/ leading
Dec 2023 – Jan 2024	Start-up meetings/review theory of change	EEF, evaluation team, delivery team
Jan – March 2024	Data governance/MoU, parental withdraw & data processing notices drafted and agreed	Evaluation team, delivery team
Feb 2024	Ethical approval	Evaluation team
Feb – May 2024	School recruitment	Delivery team
Feb – May 2024	Parental withdrawal process	Delivery team
Mar – Apr 2024	Protocol	Evaluation team
Mar – Jun 2024	Sample enumeration	FFT
Jun – Jul 2024	Baseline assessments	Alpha Plus
Jun – Jul 2024	Review of delivery documentation	Evaluation team

Jun – Jul 2024	Headteacher surveys	Evaluation team
Aug 2024	Endline sampling	Evaluation team
Complete by beginning of September 2024	Randomisation	Evaluation team
Jul 2024	Teacher surveys	Evaluation team
Sep 2024	Initial training sessions (intervention schools), with observations	Delivery team (observations by evaluation team)
Sep 2024	SAP finished	Evaluation team
Oct 2024 – May 2025	Delivery of Reading Plus	Delivery team
Dec 2024 – Jan 2025	Final training sessions (intervention schools), with observations	Delivery team (observations by evaluation team)
Jan – Feb 2025	First case study visit (intervention schools)	Evaluation team
May 2025	Second case study visit (intervention schools) QA of PiRA administration in case study schools Training for Test Administrators	Evaluation team
Jun – Jul 2025	Endline assessments	Alpha Plus
Jun 2025	Teacher and headteacher surveys	Evaluation team
Jun 2025	Interviews with the delivery team	Evaluation team
Jul 2025	Collection of usage and progress data from delivery team	Evaluation team
Aug 2025 - Jan 2026	Data analysis and report writing	Evaluation team
Jan 2026	Submit impact and IPE report	Evaluation team
Estimated late spring/ early summer 2026	Publication of evaluation report	EEF

References

Ainsworth, S., Gellen, S., Lewin, C., & Morris, S. (2022) *Evaluation of the Peer Assisted Learning Strategies for Reading UK (PALS-UK) intervention, a two-armed cluster randomised trial. Evaluation Protocol.* Education Endowment Foundation.

Bishop, R. S.(1990). Mirror, Windows and Sliding Glass Doors. *Perspectives: Choosing and Using Books for the Classrooms, 6(*3).

Breadmore, H. L., Vardy, E. J., Cunningham, A. J., Kwok, R. K. W. (2019). *Literacy Development: Evidence Review*. London: Education Endowment Foundation. Retrieved from https://educationendowmentfoundation.org.uk/public/files/Literacy_Development_Evidence_review.pdf

Breaux, K. C., & Lichtenberger, E. O. (2016). *Essentials of KTEA-3 and WIAT-III assessment*. John Wiley & Sons.

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.

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

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.

Cooper Gibson Research, (2022). *Education technology: Exploring digital maturity in Schools (Research report).* London: Department for Education.

Cooper Gibson Research, (2021). *Education Technology (EdTech) Survey 2020-21 (Research report).* London: Department for Education.

Cremin, T., Hendry, H., Leon, L. R., & Kucirkova, N. (2022). *Reading Teachers: Nurturing Reading for Pleasure*. Abingdon-on-Thames: Routledge.

Cubillos, M., & and Rousseau, R. (2024). High-interest books, choice, and independent reading: Piloting a reading program with male adolescents in Chile. Journal of Adolescent & Adult Literacy (Early View). https://doi.org/10.1002/jaal.1339

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

Education Endowment Foundation (2017). *Closing the Attainment Gap*. London: Education Endowment Foundation. Retrieved from: https://educationendowmentfoundation.org.uk/support-for-schools/bitesize-support/closing-the-attainment-gap

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). Classification of the security of findings from EEF evaluations. London: Education Endowment Foundation. Retrieved from: Classifying_the_security_of_EEF_findings_2019.pdf (d2tic4wvo1iusb.cloudfront.net)

Education Endowment Foundation (2023). *Implementation and process evaluation guidance for EEF evaluations*. London: Education Endowment Foundation. Retrieved from: EEF-IPE-Guidance-August-2022.pdf (d2tic4wvo1iusb.cloudfront.net)

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

Ehri, L. C. (1997). 'Learning to read and learning to spell are one and the same, almost'. In C.A. Perfetti, L. Rieben, & M. Fayol (Eds*.), Learning to spell: Research, theory, and practice across languages* (pp. 237–268). Mahwah, NJ: Erlbaum

Gellen, S. & Morris, S. (2023) *Evaluation of the Peer Assisted Learning Strategies for Reading UK (PALS-UK) intervention, a two-armed cluster randomised trial. Statistical Analysis Plan.* Education Endowment Foundation.

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

Guthrie, J. T., & Klauda, S. L. (2014). Effects of Classroom Practices on Reading Comprehension, Engagement, and Motivations for Adolescents. *Reading Research Quarterly*, *49*(4), 387–416. http://www.jstor.org/stable/43497193

Hess, S., (2017). Randomization inference with Stata: A guide and software. *Stata Journal 17*, 630–651.

Hiebert, E., Scott, J., Castaneda, R., & Spichtig, A. (2019). An analysis of the features of words that influence vocabulary difficulty. *Education Sciences, 9*(1), 8.

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., ... & Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *Bmj, 348*. https://doi.org/10.1136/bmj.g1687

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.

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R., & Kerr, K. (2016*). Implementation and process evaluation (IPE) for interventions in education settings: An introductory handbook*. London: The Education Endowment Foundation.

Imai, K., Tingley, D., Yamamoto, T., (2013). Experimental designs for identifying causal mechanisms. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 176,* 5–51. https://doi.org/10.1111/j.1467-985X.2012.01032.x

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

Jehangir, K., Glas, C. A., & van den Berg, S. (2015). Exploring the relation between socio-economic status and reading achievement in PISA 2009 through an intercepts-and-slopes-as-outcomes paradigm. *International Journal of Educational Research, 71*, 1-15.

Klaudia, S. L., & Guthrie, J. T. (2008). Relationships of Three Components of Reading Fluency to Reading Comprehension. *Journal of Educational Psychology, 100*(2), 310-321.

Kraft, M.A., (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher 49*, 241–253. https://doi.org/10.3102/0013189X20912798

Language and Reading Research Consortium. (2015). Learning to read: Should we keep things simple? *Reading Research Quarterly, 50*(2), 151–169.

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

Lewin, C., Smith, A., Morris, S., Craig, E., (2019). *Using Digital Technology to Improve Learning: Evidence Review*. Education Endowment Foundation, London.

Lewin, C., Morris, S., Ainsworth, S., Gellen, S., Wicker, K. (forthcoming) *Evaluation of the Peer Assisted Learning Strategies for Reading UK (PALS-UK) intervention, a two-armed cluster randomised trial.* London: Education Endowment Foundation.

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

Martins, M., Reis, A.M., Castro, S.L. et al. (2021). Gray matter correlates of reading fluency deficits: SES matters, IQ does not. *Brain Struct Funct 226*, 2585–2601. https://doi.org/10.1007/s00429-021-02353-1

Mehigan, G. Effects of Fluency Oriented Instruction on Motivation for Reading of Struggling Readers. *Educ. Sci.* **2020**, *10*, 56. https://doi.org/10.3390/educsci10030056

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

Pace, A., Luo, R., Hirsh-Pasek, K., Golinkoff, R. M (2017). Identifying Pathways Between Socioeconomic Status and Language Development. *Annual Review of Linguistics. 3,* 285-308. https://doi.org/10.1146/annurev-linguistics-011516-034226

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.

Quinn, J. M., Wagner, R. K., Petscher, Y., & Lope, D. (2015). Developmental Relations Between Vocabulary Knowledge and Reading Comprehension: A Latent Change Score Modelling Study. *Child Development, 86*(1), 159–175

Radach, R., Kennedy, A., (2013). Eye movements in reading: Some theoretical context. *The Quarterly journal of experimental psychology 66*, 429–452.

Radach, R. & Spichtig, A. (2013). *Visual Information Processing and Eye Movement Control in Reading: The "visual skills" module in Reading Plus*. Winooski, VT: Reading Plus, Inc.

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.

Rasinski, T., Samuels, S. J., Hiebert, E., Petscher, Y., & Feller, K. (2011). The relationship between a silent reading fluency instructional protocol on students' reading comprehension and achievement in an urban school setting. *Reading Psychology, 32*(1), 75–97. doi:10.1080/02702710903346873

Reading Solutions UK, (2022). *Reading Plus efficacy study in partnership with Derby Research School*. Gateshead: Reading Solutions UK.

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.

Reutzel, D. R., Petscher, Y., & Spichtig, A. N. (2012). Exploring the value added of a guided, silent reading intervention: Effects on struggling third-grade readers' achievement. *Journal of Educational Research, 105*(6), 404–415. doi:10.1080/00220671.2011.629693

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

Samuels, S. J., Hiebert, E. H., & Rasinski, T. (2015). 'Eye Movements Make Reading Possible.' In E. H. Hiebert, K. M. Wilson, & G. Trainin (Eds*.). Teaching Stamina & Silent Reading in the Digital-Global Age* (pp. 32-57). Santa Cruz: Text Project, Inc.

Schuth, E., Köhne, J., & Weinert, S. (2017). The influence of academic vocabulary knowledge on school performance. *Learning and Instruction, 49*, 157-165. https://doi.org/10.1016/j.learninstruc.2017.01.005.

Silverman, R.D., Speece, D.L., Harring, J.R., & Ritchey, K.D. (2013). Fluency has a role in the simple view of reading. *Scientific Studies of Reading, 17*(2) (2013), pp. 108-133.

Spichtig, A. N., Hiebert, E. H., Vorstius, C., Pascoe, J. P., Pearson, P. D., & Radach, R. (2016). The decline of comprehension-based silent reading efficiency in the United States. A comparison of current data with performance in 1960. *Reading Research Quarterly, 51*(2), 239-259.

efficiency in the United States: A comparison of current data with performance in 1960. Reading Research Quarterly, 51(2), 239–259Spichtig, A. N., Gehsmann, K. M., Pascoe, J. P., & Ferrara, J. D. (2019). The impact of adaptive, web-based, scaffolded silent reading instruction on the reading achievement of students in grades 4 and 5. *The Elementary School Journal, 119*(3), 443-467.

Stopforth, S., & Gayle, V. (2022). Parental social class and GCSE attainment: Re-reading the role of 'cultural capital'. *British Journal of Sociology of Education, 43*(5), 680-699.

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

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

Tracey, L, Elliot, L., Fairhurst, C, Mandefield, L, Fountain, I, Ellison, S, (2022). *Lexia Reading Core5®: Evaluation Report.* London, Education Endowment Foundation, London, Education Endowment Foundation.

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse, (2012). *Students with Learning Disabilities intervention report: Peer-Assisted Learning Strategies (WWC Intervention Report).* US Department of Education, IES.

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.