Scientific Studies of Reading

ISSN: 1088-8438 (Print) 1532-799X (Online) Journal homepage: www.tandfonline.com/journals/hssr20

Peer Assisted Learning Strategies (PALS-UK) Increases Reading Attainment, Oral Fluency and Comprehension: A Cluster Randomized Controlled Trial

Helen L. Breadmore, Stephen P. Morris, Sandor Gellen, Cathy Lewin, Emma J. Vardy, Steph Ainsworth, Kate Wicker & Luisa Tarczynski-Bowles

To cite this article: Helen L. Breadmore, Stephen P. Morris, Sandor Gellen, Cathy Lewin, Emma J. Vardy, Steph Ainsworth, Kate Wicker & Luisa Tarczynski-Bowles (09 Feb 2026): Peer Assisted Learning Strategies (PALS-UK) Increases Reading Attainment, Oral Fluency and Comprehension: A Cluster Randomized Controlled Trial, Scientific Studies of Reading, DOI: 10.1080/10888438.2026.2614770

To link to this article: https://doi.org/10.1080/10888438.2026.2614770

SCIENTIFIC STUDIES OF READING
https://doi.org/10.1080/10888438.2026.2614770

Routledge
Taylor & Francis Group

Peer Assisted Learning Strategies (PALS-UK) Increases Reading Attainment, Oral Fluency and Comprehension: A Cluster Randomized Controlled Trial

Helen L. Breadmore [a], Stephen P. Morris [b], Sandor Gellen [b], Cathy Lewin [c], Emma J. Vardy [d], Steph Ainsworth [c], Kate Wicker [c], and Luisa Tarczynski-Bowles [d]

[a]School of Education, University of Birmingham, Birmingham, UK; [b]Department of Sociology, Manchester Metropolitan University, Manchester, UK; [c]School of Education, Manchester Metropolitan University, Manchester, UK; [d]Psychology Department, Nottingham Trent University, Nottingham, UK

ABSTRACT

Purpose: This study evaluates the impact of Peer Assisted Learning Strategies (PALS-UK) in developing pupils' reading attainment, reading skills (comprehension and fluency) and affective factors (reading self-efficacy and motivation).

Method: All Year 5 pupils (9–10 years old, $N = 4840$, 49% female and 51% male) in 114 schools in England, took part in a two-armed, randomized controlled efficacy trial randomizing schools. The final analyzed sample included 53 treatment schools ($N = 1907$, 51% female and 49% male) and 50 business-as-usual control schools ($N = 1721$, 49% female and 51% male). In treatment schools, class-teachers were asked to deliver PALS-UK three times per week for 20 weeks.

Results: Pupils in treatment schools demonstrated higher curriculum-aligned reading attainment than pupils in business-as-usual control schools. A moderate effect size was found for this primary outcome. Exploratory subgroup analyses suggested that no groups were disadvantaged by treatment. In addition, analyses of secondary outcomes showed significant positive treatment effects for reading comprehension and reading fluency/rate – a measure based on speed/accuracy of reading connected text. The treatment effect was not significant for multidimensional fluency (measuring qualitative differences in expressive reading), reading self-efficacy or motivation to read.

Conclusion: This study is the most rigorous evidence to date that PALS-UK is effective in improving reading outcomes. It provides strong evidence in support of the use of this structured approach to paired reading. We conclude that the approach works, when implemented with fidelity, because it supports pupils to practice reading aloud and scaffolds use of reading comprehension strategies, which improve reading outcomes.

Introduction

The prospects for pupils who do not attain expected standards in reading by middle childhood are stark (e.g., Merga, 2020). Identification of effective approaches to the teaching of reading is therefore a strategic priority for schools and governments both nationally and internationally. In a cluster-randomized controlled trial (schools randomized to conditions), we evaluate the impact of Peer

CONTACT Helen L. Breadmore h.breadmore@bham.ac.uk School of Education, University of Birmingham Edgbaston, Birmingham B15 2TT, UK

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10888438.2026.2614770

Assisted Learning Strategies (PALS-UK) on reading outcomes for Year 5 pupils in England (aged 9–10 years). We assess whether PALS-UK supports pupils' reading attainment, skills and affective factors.

The impact of any instructional approach must be understood in context. This trial took place in England, where early reading instruction is highly prescriptive and narrowly focused. Government policy emphasizes systematic synthetic phonics and provides explicit guidance on implementation (DfE, 2023a, 2023b). The Reading Framework (DfE, 2023c) articulates the approach to systematic phonics instruction required in early primary school and recommends that the focus of instruction shifts to reading fluency and comprehension once decoding is secure, in the second half of primary school. Phonics approaches have a positive impact on word reading skills for many children (Torgerson et al., 2018 review the evidence). Attainment is a measure of achievement against curriculum standards. In England, there are two statutory assessments for reading attainment – the Phonics Screening Check at the end of Year 1 focusing on decoding, and the Key Stage 2 reading assessments at the end of primary school focusing on comprehension (STA, 2012, 2015, 2023). Despite improvements in early decoding skills over time, broader reading attainment remains stubbornly inadequate. The year we conducted this trial 27% of pupils in England failed to meet expected standards in Key Stage 2 reading assessments, at the end of primary school (DfE, 2024). Reading relies on a range of cognitive and linguistic processes (see Breadmore et al., 2019; Castles et al., 2018 for reviews). Instruction in decoding, although necessary, is not sufficient to become a good reader.

By middle childhood, reading attainment encompasses multiple reading skills. Reading fluency is the ability to access the meaning of words and produce pronunciations accurately and automatically (Hudson et al., 2009; Kuhn et al., 2010; Rasinski et al., 2011; Wolf & Katzir-Cohen, 2001). Fluent oral reading is characterized by smooth reading, with speechlike intonation and expression (Miller & Schwanenflugel, 2006; Wolf & Katzir-Cohen, 2001). Reading fluency and comprehension have a bidirectional relationship (Klauda & Guthrie, 2008; Ravid & Mashraki, 2007). The meaning of most individual words in a sentence needs to be understood to achieve text comprehension, but doing so effortlessly is particularly important because it frees cognitive capacity for comprehension (Kim et al., 2021; LaBerge & Samuels, 1974). Reading with correct prosody helps the reader to parse syntactic and semantic information. Much of the information needed for correct intonation and expression is not explicitly marked in text and so comprehension is necessary to support prosodic reading (Breadmore et al., 2019; Rasinski, 2012; Schreiber, 1991).

In addition to reading skills, children's feelings about reading (affective factors), also influence reading attainment (e.g., McGeown, Duncan, et al., 2015; McGeown, Johnston, et al., 2015). Carroll and Fox (2017) found a positive relationship between reading self-efficacy and word reading ability in 8 to11-year-olds. Reading self-efficacy is an individual's judgment of their ability to complete reading tasks, while reading motivation is the drive to read, resulting from internal (intrinsic) or external (extrinsic) factors (Conradi et al., 2014). Motivation matters – children who enjoy and value reading tend to have better reading abilities compared to those that do not. Vardy et al. (in press) explored these factors in 9- to 10-year-olds. They found that motivation to read and reading self-efficacy were separate constructs, and both constructs contributed unique variance in reading attainment. Hence, to understand how an instructional approach influences reading attainment, it is important to explore the impact on pupils' reading attainment, skills and affective factors.

What is PALS and why might it be effective?

PALS Grades 2–6 is a whole class paired reading approach that was originally developed in the US (e.g., Fuchs et al., 1997, 2000; McMaster et al., 2006) and has been adapted for implementation in various international contexts (Vardy et al., 2022). This version of PALS is centered around pairs of pupils reading and providing peer support to one another through a series of structured activities. Key elements include pupils taking turns to act as reader and coach, feedback focused on reading accurately, rereading to promote fluency, and prompts to embed use of comprehension strategies. Any text can be used, but it should be at an instructional level appropriate for both pupils, and pupils

should be given choice, to support effective coaching, modeling and motivation to read. As such, PALS provides pupils with individual and differentiated reading instruction. The teachers' role is to train pupils at the beginning, and then to facilitate and monitor pupils. In each PALS lesson, pairs complete the following core activities; partner reading, retell, paragraph shrinking, and prediction relay (see Methods for more detail).

Studies of other peer mediated interventions have generated moderate to high effect sizes for reading outcomes (EEF, 2021a; Topping et al., 2011), including oral reading fluency and comprehension (Algozzine et al., 2009; Marr et al., 2011). The What Works Clearinghouse Intervention Report on PALS (WWC, 2012) summarized the evidence as "potentially positive," but only two relevant studies met their evidence standards (Fuchs et al., 1997 - with reservations; Sáenz et al., 2005). Fuchs et al. (1997) explored effects in 12 active (and 20 control) schools in the US, finding that PALS could be delivered with high fidelity and led to positive effects on reading accuracy and comprehension. Using the same reading outcome measures with Spanish-speaking English Language Learners also in the US, Sáenz et al. (2005) found significant effects on pre-/post-treatment improvement for comprehension only. While both studies show promise, they were conducted by the original developers, had small samples, targeted specific subgroups (low ability readers, English Language Learners), and had other research design constraints that limit generalizability to a mixed ability whole class context.

PALS-UK has been developed for implementation in England with evidence-informed adaptations made through a series of iterations, including through a feasibility pilot (Vardy et al., 2022; Vardy et al., in press) and previous efficacy trial (Culora et al., 2024). Together these studies provided promising evidence and confirmed that the programme can be implemented in varied whole class contexts. They did not, however, provide robust evidence about the impact on pupils' reading outcomes. For example, the largest study was impacted by the COVID-19 pandemic, resulting in delays and abandonment of outcome testing, and substantial uneven attrition (Culora et al., 2024). The present study is a partial replication of that trial, and it represents the most rigorous test of the impact of PALS in a whole class setting both in England and internationally.

PALS-UK lessons are based on evidence about instructional approaches that support development of reading fluency and comprehension. In recent systematic reviews, Shhub et al. (2023) highlighted that effective fluency instruction typically includes repeated reading with immediate feedback, listening to modeled reading and use of multiple strategies. Hudson et al. (2020) found that most of the fluency interventions which are effective for children with reading difficulties include repeated reading alongside other approaches. Meanwhile, Zimmermann et al. (2021) found no significant difference in reading outcomes from repetitive or non-repetitive reading fluency interventions. They highlight that non-repetitive reading interventions bring additional benefits from greater text exposure due to wider reading, and effect sizes on reading comprehension tend to be higher than for fluency. PALS-UK lessons strike a balance between repeated and wide reading. Across sessions, children are exposed to a large volume of text, while within each PALS-UK lesson, pupils practise oral reading with immediate feedback focused on word reading accuracy, with some activities including re-reading. These elements are expected to improve oral reading fluency.

PALS-UK activities are also designed to support reading comprehension. Central to theories of comprehension is the idea that the reader constructs a mental model to represent the characters, places and events described in a text (Perfetti, 1999; Stafura & Perfetti, 2017; van Dijk & Kintsch, 1983). Mental models are updated dynamically as more information is encountered. Instruction in comprehension strategies is intended to scaffold the reader's approach to developing a coherent mental model. There is abundant evidence that instruction in comprehension strategies can support reading (e.g., the Education Endowment Foundation Teaching and Learning Toolkit summarizes evidence from 141 studies, concluding this is a high impact approach - EEF, 2021b). Comprehension monitoring is an example of a strategy that PALS-UK supports – the ability of readers not only to notice that their understanding of the text has broken down, but to respond with behaviors that support comprehension. Comprehension monitoring is clearly associated with reading comprehension outcomes (Language and Reading Research

Consortium, Jiang and Davis, 2017; Language and Reading Research Consortium, Yeomans-Maldonado, 2017; Oakhill & Cain, 2012; Oakhill et al., 2003). Several PALS-UK activities make this explicit – for example, coaches help readers to check their understanding as they progress through the text.

Other comprehension strategies that are explicitly scaffolded in PALS activities include identifying the main idea, sequencing events, retelling and summarizing, and predicting what will happen next (Fuchs et al., 1997, 2000; McMaster et al., 2006). Peng et al. (2024) conducted a Bayesian network meta-analysis of 52 intervention studies testing the effectiveness of training reading comprehension strategies to struggling readers. They concluded that rather than there being one particularly strong active-ingredient driving the effects of reading strategy instruction, the combination of strategies taught (ingredient-interaction) is important in effectively reducing cognitive load. They also highlight how some strategies effectively encompass the functions of others. For example, identifying the main idea and retell often require the reader to draw inferences. Because of this, we expect that the combination of strategies scaffolded in PALS-UK will lead to improvements in reading comprehension.

The present study

To become skilled readers, pupils need practice and instructional support to develop reading fluency and comprehension. PALS-UK lessons engage pupils in structured peer-assisted reading activities, which align with evidence on reading fluency and reading comprehension. The primary aim of this study is to establish what impact PALS-UK has on reading attainment, by comparing outcomes of pupils exposed to PALS-UK to business-as-usual controls. We quantify reading attainment using New PiRA (Progress in Reading Assessment – Ruttle et al., 2021), a termly curriculum-aligned norm-referenced assessment with strong predictive/criterion validity due to its association with statutory assessments (further detail including reliability and validity metrics are presented in the method). The mapping to National Curriculum is indicated in the manual, and in a systematic review the test was evaluated as having good face validity as a measure of reading attainment in England (Breadmore & Carroll, 2021). This is a broad measure of reading, with test items designed to tap the following content domains; vocabulary, comprehension, summary, inference, prediction, structure, impact and comparisons. As such it is a valid measure of reading attainment as well as the constructs that PALS-UK aims to develop. Secondary aims include understanding the impact on pupils' reading skills, affective factors and whether effects are similar for various subgroups of children.

We distinguish between the primary research question and secondary, exploratory questions. Conclusions as to whether the evidence presented supports the hypothesis that PALS-UK is effective are based solely on answers to the primary research question (RQ1). Answers to secondary questions (RQ2 and RQ3) are exploratory and are used to support development of hypotheses for future research and theoretical models.

RQ1: What impact does PALS-UK have on reading attainment? Is there a difference in reading attainment among Year 5 pupils (9–10-year-olds) in schools exposed to PALS-UK, compared to Year 5 pupils in control schools exposed to business-as-usual conditions? We anticipate a positive treatment effect – higher scores on reading attainment for pupils exposed to PALS-UK.

RQ2: What impact does PALS-UK have on reading skills and affective factors? In each case, we expect to see a positive treatment effect, with more positive scores in pupils exposed to PALS-UK compared to pupils in control schools.

(a) What is the treatment's effect on reading skills, (a) oral reading fluency and (b) reading comprehension?

(b) What is the treatment's effect on affective factors, (a) reading self-efficacy and (b) motivation for reading?

RQ3: Do all pupils respond similarly to PALS-UK? Do treatment effects differ for subgroups of pupils at risk of lower reading attainment (low baseline reading attainment, special educational needs or disabilities – SEND, English as an Additional Language – EAL, or eligibility for free school meals – FSM[1])? While we expect PALS-UK to support all learners, there is not sufficient evidence to make strong predictions about these outcomes.

Methods

We conducted a two-armed, pragmatic, cluster-randomized controlled efficacy trial. Schools were assigned to either the treatment (participating in PALS-UK) or control group ("business-as-usual"). We targeted all Year 5 pupils (9 to 10-year-olds) in participating primary schools. Randomization was stratified on key school-level variables (described later). Balance tests, comparing baseline reading attainment and socio-demographic indicators measured at the pupil and school levels, at both randomization and at analysis, revealed no appreciable differences between trial-arms (see Table 1).

This research was conducted in accordance with the BPS Code of Human Research Ethics (British Psychological Society, 2021). Approval was granted by Manchester Metropolitan University Research Ethics Committee (and accepted by all other institutional ethics committees), following a detailed review of project design, protocol, and procedures. Before data were collected, schools signed a Memorandum of Understanding and Data Sharing Agreement. Parents were informed about the research and given an opportunity to withdraw their children. Children's class teacher also showed them a video (produced by the research team) explaining the evaluation. They were informed that they did not need to complete the evaluation tasks if they did not want to. Test administrators were also trained to look out for any non-verbal signs that children might not be comfortable completing the tasks, and to stop the activities where appropriate (consistent with assent procedures outlined in Dockett & Perry, 2011). The study was funded by the Education Endowment Foundation Accelerator Fund and the protocol was preregistered prior to randomization (https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/pals-uk-accelerator-fund and https://archive.org/details/osf-registrations-f2wgj-v1).

Participants

Schools

State-funded primary schools in the English Regional School Commissioner areas of the North, East Midlands & Humber, and West Midlands were eligible to take part, unless they were involved in any other similar trials targeting pupils entering Year 5 in the same academic year (funder mandated limitations on eligibility). Schools were also ineligible if they had previously used PALS-UK.

At randomization, 114 schools were assigned 1:1 to treatment and control groups. Eight schools subsequently withdrew (due to staffing issues), and another three had missing endline data for the primary outcome. The sample as analyzed included 103 schools (53 treatment, 50 control). Control schools did not receive training and did not deliver PALS-UK to their pupils; they continued using their usual approaches to reading.

Pupils

Pupils were identified and enumerated at the end of Year 4, before treatment commenced at the beginning of Year 5. Demographic information was collected. This included binary indicators for SEND, EAL, and FSM eligibility sourced from school records to form four subgroups of pupils to

Table 1. Distribution of baseline school and pupil characteristics by trial arm at randomization and analysis.

	At randomization sample (pupils 4840, classes 197, schools 114)			At analysis sample (pupils 3628, classes 174, schools 103)			
	Treatment	Control	Unadjusted difference	Treatment	Control	Unadjusted difference	Adjusted difference [95% CI]
School level							
Rural school (%):	10.5	15.8	−5.3	11.3	16.0	−4.7	−0.40 [−1.54, 0.74]
Ofsted rating (%):							
Outstanding	10.0	<5		10.9	<5	–	
Good	78.0	86.5	−8.5	76.1	88.9	−12.8	
Requires improvement	12.0	<5		13.0	<5	–	
School type[i] (%):							
Academy	42.1	50.9	−8.8	41.5	46.0	−4.5	−0.18 [−0.96, 0.60]
Maintained	57.9	49.1	8.8	58.5	54.0	4.5	
Proportion of the school roll:							
EAL	15.2	17.6	−2.4	15.9	17.8	−1.9	−0.02 [−0.11, 0.24]
FSM	27.3	28.1	−0.8	27.4	28.3	−0.9	−0.04 [−0.12, 0.04]
KS2 reading scaled score school (mean):	104.3	104.2	0.1	104.4	104.3	0.1	0.05 [−0.90, 0.99]
Pupil level							
Age at endline in months (mean. SD)	127.5 (3.5)	127.5 (3.6)	0.0	127.6 (3.5)	127.5 (3.6)	0.1	0.14 [−0.12, 0.41]
Reading attainment at baseline (mean and SD PiRA raw score, possible range 0 – 40)	16.1 (8.2)	16.3 (8.2)	−0.2	16.3 (8.0)	16.5 (8.1)	−0.2	−0.17 [−0.26, 0.79]
Gender (%):							
Male	50.9	51.0	−0.1	49.3	50.7	−1.4	−0.04 [−0.20, 0.11]
Female	49.1	49.0	0.1	50.7	49.3	1.4	
SEND (%):							
SEND pupils	17.1	17.2	−0.1	14.6	15.8	−1.2	−0.07 [−0.33, 0.19]
Non-SEND pupils	82.9	82.8	0.1	85.4	84.2	1.2	
EAL (%):							
EAL pupils	19.5	24.2	−4.7	19.6	22.4	−2.8	0.09 [−0.72, 0.91]
Non-EAL pupils	80.5	75.8	4.7	80.4	77.6	2.8	
FSM (%):							
FSM pupils	32.7	37.0	−4.3	31.8	34.7	−2.9	−0.14 [−0.53, 0.25]
Non-FSM pupils	67.3	63.0	4.3	68.2	65.3	2.9	

Note. [i]On average we have less than two classes per school. Some schools a single class schools within each year group, whereas others have multiple classes per school. [ii] There are two types of state-funded schools in England – maintained schools are owned and maintained by local authorities, academy schools receive funding directly from government and operate independently.

answer the third research question: (1) lower baseline reading attainment (operationalised as in the lowest quartile), (2) SEND, (3) EAL and (4) FSM. Pupils were excluded from baseline assessment if, by their school's judgment, they were unable to complete them. These pupils and any who joined the school after baseline assessments could participate in the treatment but are not included in this evaluation. The final analyzed sample consisted of 3628 pupils – 1907 in the treatment (PALS-UK) and 1721 in the control group.

Treatment delivery: PALS-UK

To understand treatment delivery, it is helpful to distinguish between lessons, teacher training and materials, and pupil training and materials. Beginning with lessons, pupils read aloud together in pairs and complete a series of structured activities, taking it in turns to act as reader and coach. Lessons took around 35 minutes (including setup) with activities completed in a set order: Partner Reading (10 minutes), Retell (2 minutes), Paragraph Shrinking (10 minutes), and Prediction Relay (10 minutes). Schools were asked to deliver PALS-UK three times a week for 20 school weeks, from October 2022 to May/June 2023.

Teachers designated "first" and "second" readers based on reading ability using baseline reading attainment (supplemented by professional judgment). The higher ability half were first readers, and the lower ability half were second readers. During Partner Reading, the first reader reads aloud while the second reader reads along silently, using a script to coach if they notice mistakes. The coach focuses on identifying and correcting word reading errors, encouraging the reader to re-read the whole sentence accurately and fluently if they make a mistake. They swap roles, and the second reader re-reads. In Retell, the coach uses a series of prompts to encourage the reader to summarize short sections of text. During Paragraph Shrinking, the coach uses prompts to help the reader summarize and think about sequencing. During Prediction Relay, the coach uses prompts as scaffolds to help the reader to make predictions about what will happen in the next half a page, before reading and checking predictions.

Before classroom delivery, teachers attended a one day in-person training event (repeated in seven regional locations). The day began with a summary of the skills needed to comprehend text and read fluently, and a summary of how existing empirical evidence and teacher feedback informed development of the approach. Teachers were then familiarized with what a PALS-UK lesson should look like. Finally, teachers were supported to plan implementation within their specific setting and to discuss potential barriers. They left with everything necessary to deliver pupil training – a manual, detailed scripted lesson plans, pupil worksheets. They had scheduled their PALS-UK lessons, assigned first and second readers, pairs, and considered which pupils might struggle and what adaptations/support might help them. They were invited to an online top-up training session after training their class and were provided just-in-time support throughout (described in more detail in Vardy, Lewin, et al., In prep). This package of teacher training differs somewhat from the original PALS Grade 2–6 US version, as it has been specifically developed for the UK context (by the 1st and 5th Authors). It aims to ensure that teachers understand not only know how to deliver PALS-UK effectively within their context, and also why it is important to implement the approach with fidelity. The training explicitly links theoretical perspectives on reading development, national policy/guidance on teaching reading, as well as empirical evidence of what works to support reading, to explain why each core component of PALS activities are included. In addition, the training has been modernized to include digital and online elements (such as videos and webinars).

Pupils were introduced to the PALS-UK activities gradually, through teacher-led whole-class pupil training lessons delivered three times per week for the first four weeks. Extracts and a workbook were provided with additional prompts to help to establish routines. Minor adaptations were made to these workbooks compared to the original PALS Grades 2–6, largely to modernize their appearance, present the text in British English, as well as updating extracts to include locally appropriate cultural references. Once established, pupils completed activities using two prompt cards and any text. In this trial, they continued with three lessons per week for a further 16 weeks. In previous studies, we found that some schools found it difficult to provide a variety of texts to enable each pair to read a book they could access and were motivated to read together (Vardy et al., 2022). Hence, here we provided a selection of 20 age-appropriate contemporary fiction books per class, emphasizing that these should supplement existing book supplies and were not central to the programme. Sample books varied in length, level of difficulty, and diversity of content and character. Training included time dedicated to exploring different

approaches to help children to select books, with consideration of text complexity, interest, and reading level. Classroom observations confirmed that schools appropriately combined these books with their existing supplies to provide pupils with a choice of texts of appropriate difficulty. The provision of books and training on supporting book choice was particularly important as whereas PALS Grades 2–6 includes an extrinsic points-based motivational framework, PALS-UK instead focuses on mechanisms that support intrinsic motivation supplemented using existing motivational frameworks as appropriate (such as school merit systems). The rationale for these adaptations is described further in Vardy et al. (2022).

In addition to the outcomes reported here, the trial included an Implementation and Process Evaluation which we report in detail elsewhere (Lewin et al., 2024; Vardy, Lewin, et al., In prep). Only 2/53 schools were considered noncompliant (teachers and pupils were not trained and did not deliver/receive PALS-UK). There is clear evidence from multiple sources that, after attending training, teachers were able to implement both pupil training and main delivery phases with high fidelity. Structured classroom observations from peer observers and researchers confirmed that the overwhelming majority of teachers implemented each of the core activities appropriately – completing the activities in the right order, with the correct prompts and for the correct duration. Only 1 school showed poor implementation adherence in the researcher observations, and while the number of schools that completed peer observations toward the end of programme delivery was relatively low (N = 25/53), 22/25 showed excellent implementation adherence and the remaining three were very good. Teachers' logs, surveys and interviews confirmed adherence to dosage. Monitoring data was successfully gathered from 49/51 of compliant schools, confirming that they delivered PALS-UK at least once a week for 20 weeks, with the mean total number of sessions reported to be 51. In response to surveys, 90% of teachers confirmed that all sessions for the four weeks of pupil training were delivered, and 80% reported full delivery of the remaining 16 weeks of the programme.

Outcomes and measures

The primary outcome (reading attainment) was assessed at baseline before randomization (end of Year 4, Summer 2022) and again at endline after pupils in the treatment group had finished the program (end of Year 5, Summer 2023). Secondary reading skill outcomes (reading comprehension and fluency) were assessed at endline only among a random subsample of pupils (sampling described later). Affective factors (reading self-efficacy and motivation) were assessed at both timepoints with all pupils.

Primary outcome - reading attainment: new PiRA (Progress in reading assessment)

The primary outcome, reading attainment, was derived from the New PiRA (Progress in Reading Assessment - Ruttle et al., 2021). New PiRA is a series of termly UK curriculum-aligned standardized norm-referenced assessments. We used the Summer Term Year 4 test at baseline and the Summer Term Year 5 test at endline. The manual reports excellent internal consistency (Cronbach's alpha > .9) and criterion validity, with a strong relationship with statutory national test scores (Pearson coefficient .64–.79 for Key Stage 2 - Ruttle et al., 2021).

New PiRA was administered to the whole class by the classroom teacher in accordance with standardized instructions. At baseline, the digital version was administered. At endline, the paper version was used and hand scored by trained markers (administration changed due to technical issues experienced at baseline). In both cases, the assessment took 40–50 minutes. Raw scores (sum of correct responses) could vary from 0 to 40 at baseline and 0 to 45 at endline. Although New PiRA offers both age and cohort normed standardized scores, raw score is used as the dependent/outcome variable for Reading Attainment in our analyses, to reduce any potential biases that could be introduced by differences between the norming sample and the present sample.

Secondary outcomes: reading skills and affective factors

Secondary outcomes included measures of component reading skills (comprehension and two measures of oral reading fluency) and affective factors (self-efficacy and motivation).

WIAT-III UK-T reading comprehension and oral reading fluency subtests. Reading comprehension and oral reading fluency subtests from the WIAT-III UK-T (Wechsler, 2018) were administered individually by trained testers, in line with standardized instructions, including use of chronological age-related start/stop points and performance-related reverse rules. Administration took approximately 30 minutes. Total weighted raw scores were obtained using the conversion tables in the manual, to account for differences between passages. The manual reports good to excellent internal consistency for this age range (reading comprehension .80, fluency .92).

For Reading Comprehension, pupils read a series of passages. After reading each passage, the examiner asked literal and inferential comprehension questions, and the pupil provided oral responses. This contrasts in both content and mode of response to the PiRA (a measure of reading attainment closely aligned to curriculum assessments) and has face validity as measure of reading comprehension skill. Pupils chose whether to read silently or aloud and could refer back to the passage when answering. The examiner used the manual to score responses for accuracy (0 – no credit, 1 – partial credit, 2 – full credit). The Reading Comprehension outcome was the total weighted raw score, which could range from 0 to 44.

For Oral Reading Fluency, pupils read two different passages of expository and narrative text aloud, which was audio recorded and later scored offline. They were asked to "be sure to do your best reading." The time (in seconds) spent reading each of two passages was summed – from when the pupil began to read the first word until they finished the last word. The number of correctly read words was calculated by subtracting the number of errors from the total number of words in the passages. Reading rate per minute was then calculated by dividing the number of correctly read words by time (in seconds) and converted into minutes by multiplying the result by 60. The Reading Fluency/Rate outcome was total weighted raw score, which could range from 0–121.

Multidimensional fluency scale (MDFS). Multidimensional fluency data were collected using the MDFS (Rasinski, 2004b; Zutell & Rasinski, 1991). Previous studies have shown MDFS to provide reliable assessment of the prosodic elements of reading, with good inter-rater and intra-rater reliability (Moser et al., 2014; Smith & Paige, 2019). MDFS was scored by trained test administrators using audio recordings of pupils reading the WIAT-III UK-T oral reading fluency subscales. Test administrators used the rubric to rate each dimension (expression and volume, phrasing, smoothness, and pace) on a four-point scale and summed together to form the Multidimensional Fluency outcome. Scores could range from 4 to 16.

The reading skills (reading comprehension and fluency) test administrators were all experienced in diagnostic assessment of primary age children and received extensive training to minimize inter-rater variability. This included calibration discussions, whereby testers rated multiple recordings obtained during pilot testing and discussed scoring until a consensus was reached for each recording and each dimension. Recordings from the first school visited by each tester were moderated. When occasional discrepancies arose, additional training was provided and scores amended following consensus. In such cases, further moderation was conducted for those testers.

Feelings about reading survey. A two-part questionnaire measured affective factors (i) reading self-efficacy and (ii) motivation for reading, at baseline and endline. The reading self-efficacy scale was initially developed by Carroll and Fox (2017). Vardy et al. (Under review - 2026) slightly adapted and validated that scale while also developing the motivation for reading scale. They demonstrated that both measures had good-excellent internal reliability (Cronbach's α .83 and .90). Here, responses to both scales were on a 7-point Likert scale with items reverse scored as necessary so that a low score indicated negative, and a high score indicated positive feelings about reading. The survey took about 10 minutes to complete,

digitally at baseline (using Jisc Online Surveys) and on paper at endline. Summed responses to 20 items quantified the Reading Self-efficacy outcome – scores could range from 20 to 140. Summed responses to 10 items quantified the Motivation for Reading outcome – scores could range from 10 to 70.

Randomization procedure, subsamples and baseline characteristics

The treatment effects on group administered measures of reading attainment and affective factors were assessed for *all* pupils in the study following the principle of intention-to-treat. The treatment effects on reading skills were assessed for a *subgroup* identified before randomization. One class per school was selected at random and, within that class, 10 pupils with valid baseline PiRA scores were randomly chosen. To boost response rates, an additional five pupils were selected as reserves who could be asked to complete an assessment if any pupils were absent on the day of testing.

Randomization took place after baseline assessments were complete and the subsample for secondary reading skills measures had been identified. Schools were randomly assigned to treatment or control, stratified by school size (single-form vs. multiple-form – the number of classes per year group) and percentage of pupils eligible for FSM (low-FSM proportion vs. high-FSM proportion school).

Table 1 presents the characteristics of schools and pupils in "as randomized"[2] and "as analyzed"[3] samples. Treatment and control groups were similar – they had comparable school OFSTED ratings,[4] school types, rural/urban classification, and aggregate prior reading attainment.[5] At both randomization and analysis, pupil-level characteristics (including baseline reading attainment, SEND status) were well balanced. Gender distribution was also balanced, with nearly equal numbers of male and female pupils in both groups. Some differences between trial-arms were identified, for example, the proportion of pupils with EAL and FSM. Despite this, average baseline test scores were similar across trial-arms (cohort norm-referenced equivalent standard score of 91), indicating good overall balance for primary analysis. These findings also suggest that sample characteristics remained stable from randomization to analysis, with minor differences due to school withdrawals between randomization and endline.

Sample size and statistical analysis

We used the concept of the Minimum Detectable Effect Size (MDES) to determine sample size (Dong & Maynard, 2013). The MDES is the smallest true effect (defined as a standardized difference in means) of PALS-UK on the primary outcome that a given sample size and study design can detect at a chosen level of statistical significance and sample power. The research team had capacity to recruit 120 schools, which acted as a feasible upper limit on sample size. The following assumptions were made to calculate MDESs:

- Type I and II error rates at 5% and 20%, respectively, with randomization on a 1:1 basis.
- Baseline and endline correlations for the primary outcome (reading attainment – New PiRA raw scores) at 0.70.
- A three-level clustered design (pupils nested within classes, classes nested within schools) with Intraclass Correlations Coefficients (ICCs)[6] of .10 at the school level and .05 at the class level.
- An average of 25 Year 5 pupils per class and 1.55 classes per school.
- A 10% school-level attrition from randomization to endline testing.

Using these parameters, the MDES was 0.203 at the protocol stage, 0.218 at randomization, and 0.212 at analysis. These effect sizes were considered sufficiently small such that we would have a reasonable chance of detecting a statistically significant difference between the two trial-arms that we might regard as meaningful from the perspective of informing teaching practice.

For each primary and secondary analysis, a sample estimate was derived from multi-level linear regression models. Models are described in further detail later. The uncertainty of the treatment effects was expressed through continuous p-values and 95% confidence intervals. Regression estimates for treatment effects were converted into effect sizes. All statistical analyses were performed in STATAv18 software. Stata do-files for analyses and randomization can be found in Supplementary Materials.

Estimation of effect sizes

For both the primary and secondary analyses, results from the models were reported as effect sizes, expressed as standardized mean differences consistent with Hedges' g. To calculate the effect size for each outcome, an initial empty model was estimated to obtain unconditional variance components at the pupil, class, and school levels. The adjusted treatment effect was then obtained from the impact model. This effect estimate was standardized using the square root of the sum of the variance components from the empty model. Each effect size was accompanied by a 95% confidence interval. These intervals were calculated by dividing the upper and lower bounds of the adjusted treatment effect's confidence interval by the same standard deviation used in computing the effect size.

Analysis of missing data

Missing data analysis focused on the primary outcome. Since missing data before randomization was unlikely to bias treatment effects, sensitivity analyses focused on follow-up data. The type of missingness was assessed, and sensitivity tests were carried out under an assumption of missingness being missing at random (MAR) as well as procedures agnostic to the nature of missingness. Multiple Imputation analysis, conducted under the MAR assumption, indicates no change in the overall treatment conclusions. However, further sensitivity analyses reveal some uncertainty in the intervention's effect on reading score. See supplementary materials for further details.

Results

We examine the treatment effect of PALS-UK on six outcomes to address RQ1 and RQ2. These include the primary outcome – reading attainment, and secondary outcomes – reading skills (comprising reading comprehension and two measures of oral reading fluency), and affective factors (reading self-efficacy and motivation to read). Descriptive statistics related to each primary and secondary measures at baseline (if available) and endline are presented in Table 2. See Table 1 for pupil demographics.

Table 3 presents bivariate correlations between reading measures, with separate matrices above and below the diagonal for the treatment and control groups, respectively. Significant positive correlations were observed across all reading measures, ranging from $r = .19$ to .75 (Table 3).

Estimated effects of PALS-UK

For each outcome, results from the regression model with covariates are presented in Table 4. Model estimates are obtained using restricted maximum likelihood. Statistical estimates of treatment effects are "intention-to-treat estimates" - pupils are analyzed as they were randomized, as described in the study protocol and statistical analysis plan (Ainsworth et al., 2022; Gellen & Morris, 2023). The model contains the treatment group indicator (coded to "1" if the school was randomized to PALS-UK, "0" for control), which also captures the strata used in the randomization, along with pupil month of birth[7], and a baseline measure of either reading attainment, self-efficacy, or motivation, depending on which dependent variable/outcome is considered. The model takes the following form:

Table 2. Descriptive statistics at baseline and endline for primary and secondary outcomes overall, and in each trial-arm.

	Primary outcome		Secondary outcomes						
	Attainment[i]		Comprehension[ii]	Fluency (rate)[iii]	Fluency (multi)[iv]	Self-efficacy[v]		Motivation[vi]	
	(B)	(E)	(E)	(E)	(E)	(B)	(E)	(B)	(E)
Overall									
N schools	114	103	97[vii]	106	106	112	102	112	102
N pupils	4,481	3,628	922	999	1,009	3,841	3,025	3,841	3,019
Mean	16.20	21.91	56.55	47.94	11.89	106.00	105.15	53.33	51.99
(SD)	(8.22)	(8.56)	(9.02)	(8.51)	(2.88)	(23.47)	(19.82)	(12.40)	(12.93)
Range	0–38	0–43	2–80	13–76	4–16	20–140	23–140	10–70	10–70
Skewedness	0.18	−.28	−.79	−.42	−.53	−.78	−.60	−.79	−.72
Kurtosis	2.23	2.48	6.20	3.55	2.87	3.33	3.17	3.11	2.85
PALS-UK									
N schools	57	53	49	54	54	56	52	56	52
N pupils	2,290	1,907	467	506	512	1,977	1,563	1,977	1,561
Mean	16.13	22.32	57.46	48.66	11.99	105.51	105.18	53.38	52.35
(SD)	(8.22)	(8.26)	(8.32)	(8.16)	(2.85)	(23.75)	(20.10)	(12.14)	(12.86)
Range	0–38	0–42	27–79	21–67	4–16	20–140	23–140	10–70	10–70
Control									
N schools	57	50	48	52	52	56	50	56	50
N pupils	2,191	1,721	455	493	497	1,864	1,462	1,864	1,458
Mean	16.26	21.46	55.61	47.21	11.78	106.52	105.13	53.28	51.61
(SD)	(8.17)	(8.87)	(9.61)	(8.81)	(2.91)	(23.16)	(19.52)	(12.66)	(13.00)
Range	0–37	0–43	2–80	13–76	4–16	20–140	20–140	10–70	10–70

Note. B = Baseline; E = Endline. i Reading attainment measured using PiRA raw score, ii Reading comprehension measured using WIAT-III UK-T reading comprehension subtest weighted score, iii Fluency measured using WIAT-III UK-T oral reading fluency subtest weighted score, iv MDFS measured using the multidimensional fluency scale total score, v Reading self-efficacy and vi motivation to read measured using total scores from subscales from the Feelings About Reading survey. vii Data from nine schools (75 pupils) were excluded from these analyses due to improper administration of the comprehension subtest. One test administrator did not allow pupils to refer to the text while answering, contrary to test guidelines. Please refer to Table 1 for pupil demographics.

Table 3. Correlation matrix of baseline and endline measures for PALS-UK and control participants.

Variables	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. Attainment[1] (B)	–	0.75	0.60	0.59	0.54	0.36	0.34	0.28	0.27
2. Attainment[i] (E)	0.73	–	0.64	0.61	0.53	0.34	0.38	0.31	0.27
3. Comprehension[ii] (E)	0.60	0.64	–	0.55	0.49	0.27	0.30	0.26	0.23
4. Fluency (rate)[iii] (E)	0.55	0.54	0.44	–	0.70	0.34	0.35	0.23	0.23
5. Fluency (multi)[iv] (E)	0.53	0.53	0.49	0.70	–	0.34	0.29	0.26	0.19
6. Self-efficacy[v] (B)	0.34	0.34	0.27	0.34	0.34	–	0.50	0.60	0.31
7. Self-efficacy[v] (E)	0.35	0.38	0.30	0.35	0.29	0.50	–	0.41	0.61
8. Motivation[vi] (B)	0.33	0.31	0.26	0.23	0.26	0.60	0.41	–	0.51
9. Motivation[vi] (E)	0.25	0.27	0.23	0.23	0.19	0.31	0.61	0.51	–

Note. Lower triangle = intervention group (PALS-UK); upper triangle = control group. All values are Pearson correlation coefficients. All correlations are statistically significant at $p < .001$. All significant correlations remained significant after Benjamini-Hochberg correction for multiple comparisons.

i Reading attainment measured using PiRA raw score, ii Reading comprehension measured using WIAT-III UK-T reading comprehension subtest weighted score, iii Fluency measured using WIAT-III UK-T oral reading fluency subtest weighted score, iv MDFS measured using the multidimensional fluency scale total score, v Reading self-efficacy and vi motivation to read measured using total scores from subscales from the Feelings About Reading survey. B = Baseline; E = Endline.

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

Here, Y_{ijk} denotes the unstandardzed score achieved by pupil i in class j and school k on the given outcome. The variable T_k equals one if the pupil's school was randomized to the intervention group, and zero otherwise. The sample estimate of the coefficient β_1 is the effect of the intervention on treated pupils. X_{ijk} is the raw score obtained by pupil i on the respective baseline test. Z_{ijk} denotes the pupil's month of birth, drawn from baseline demographic data. S_k represents the set of school-level

Table 4. Primary and secondary outcome analysis – estimated effect of PALS-UK on reading attainment, reading skills and affective factors.

Parameters	Estimate (B)	SE	p -value	FDR adj. p-value	Effect size [95% CI]
Reading attainment (PiRA)					
Intercept	2.767	3.204	.388		
PALS-UK vs. Control	1.025	0.459	.026		0.12 [0.014,0.223]
Pretest	0.766	0.012	<.001		
Age	0.046	0.026	.077		
High FSM/single-form school	0.211	0.737	.774		
Low FSM/multi-form school	1.227	0.676	.069		
High FSM/multi-form school	−0.171	0.605	.777		
Reading comprehension (WIAT-III UK-T)					
Intercept	62.237	8.031	<.001		
PALS-UK vs. Control	1.468	0.666	.027	.053	0.16 [0.018, 0.308]
Pretest	0.636	0.030	<.001		
Age	−0.126	0.065	.054		
High FSM/single-form school	−1.564	0.986	.113		
Low FSM/multi-form school	−0.708	0.949	.456		
High FSM/multi-form school	−2.834	0.858	.001		
Reading fluency/rate (WIAT-III UK-T)					
Intercept	52.132	7.621	<.001		
PALS-UK vs. Control	1.165	0.554	.035	.053	0.14 [0.010, 0.269]
Pretest	0.627	0.028	<.001		
Age	−0.129	0.062	.037		
High FSM/single-form school	1.126	0.822	.170		
Low FSM/multi-form school	0.777	0.792	.326		
High FSM/multi-form school	1.547	0.721	.032		
Multidimensional fluency (WIAT-III UK-T)					
Intercept	16.443	2.488	<.001		
PALS-UK vs. Control	0.103	0.244	.675	.675	0.04 [−0.131, 0.202]
Pretest	0.191	0.009	<.001		
Age	−0.060	0.020	.003		
High FSM/single-form school	−0.173	0.362	.633		
Low FSM/multi-form school	−0.458	0.349	.190		
High FSM/multi-form school	−0.628	0.316	.047		
Reading self-efficacy (FAR)					
Intercept	55.244	10.759	<.001		
PALS-UK vs. Control	0.354	1.037	.733		0.018 [−0.085, 0.120]
Pretest	0.413	0.014	<.001		
Age	0.032	0.087	.711		
High FSM/single-form school	1.270	1.819	.485		
Low FSM/multi-form school	3.309	1.527	.030		
High FSM/multi-form school	1.005	1.430	.482		
Motivation to read (FAR)					
Intercept	41.037	6.993	<.001		
PALS-UK vs. Control	0.890	0.623	.153		0.069 [−0.026, 0.163]
Pretest	0.518	0.017	<.001		
Age	−0.142	0.056	.011		
High FSM/single-form school	−0.242	1.119	.829		
Low FSM/multi-form school	0.940	0.926	.310		
High FSM/multi-form school	−0.579	0.870	.506		

Note. Linear mixed effects models with random effects for schools and classes, estimated using the command "mixed" and restricted maximum likelihood in STATA v18. Effect size and confidence intervals standardized using unconditional pooled standard deviation obtained from an empty model. School size is indicated as single-form (one class per year group)/multi-form (multiple classes per year group).

stratification variables. w_k is the school-level random effect, u_{jk} is the class-level random effect, and e_{ijk} is the pupil-level residual term.

As attainment is the primary outcome, the result from this model is decisive in terms of whether PALS-UK is ultimately viewed as effective, or not. In the secondary analyses, the model was similarly determined *a priori* to assess the impact of PALS-UK. For reading skills (reading comprehension and the two measures of oral reading fluency), baseline measures were not

collected using the same instruments. Instead, PiRA baseline scores were used to control for prior reading attainment.

Table 4 presents the treatment effect estimate for each outcome (β), its standard error, and p-value, as obtained directly from the regression model. The estimated effect size (standardized difference in means) is reported with its 95% confidence interval, alongside parameter estimates (β, standard error, and p-value) to express the effect of PALS-UK (Table 4). Throughout, statistical estimates are obtained from mixed effects linear models, with class and school level random effects. This is the case for all outcomes, except for the comprehension and fluency secondary outcomes, where models contain a random effect at the school level only.[8]

The strength of the estimated effect of PALS-UK on each outcome is assessed on the basis of the estimated effect size and in comparison to effect sizes obtained across high quality recent field trials in education, in which "small effect sizes" were those less than 0.05 (standardized difference in means), "medium effect sizes" 0.05 to less than 0.20, and "large effect sizes" 0.20 or greater (Kraft, 2020, p. 250).

In response to RQ1, the regression analysis for the primary outcome (reading attainment) indicates that pupils in schools implementing PALS-UK demonstrated greater gains in reading attainment relative to pupils in the control group. At endline, the mean difference without controlling for covariates, was B = 0.949 (SE = 0.704). After controlling for covariates, the treatment effect indicates scores were just over a point higher (B = 1.025, SE = 0.459, p = .026, Table 4). We converted this effect estimate into an effect size (standardized mean difference) by dividing by the outcome unconditional pooled standard deviation. This yielded an effect size of 0.12 (95% CI: 0.01 to 0.22), which falls within the medium range. Although the confidence interval is relatively wide, it does not include zero, suggesting a statistically meaningful positive effect. In the model with covariates, the ICC was 0.10 (95% CI: 0.06 to 0.15) at the school level and 0.04 (95% CI: 0.00 to 0.04) at the class level.[9]

Two sets of hypotheses were tested in response to RQ2. The first relate to reading skills Pupils in the treatment group scored on average one and a half points higher than control group pupils on reading comprehension at endline (B = 1.468, SE = 0.666, p = .027, Table 4), with an effect size of 0.16 (95% CI: 0.02 to 0.31). Without covariates, the estimated mean difference was β = 1.808 (SE = 0.938). On oral reading fluency (rate), pupils in treatment schools scored just over one point higher compared to their control group counterparts (B = 1.165, SE = 0.554, p = .035, Table 4), equivalent effect size 0.14 (confidence interval 0.01 to 0.27), which we interpret to be of medium size. Without covariates, the model yielded a mean difference of β = 1.440 (SE = 0.637). On the multidimensional fluency scale, the average score among pupils in treatment schools was 0.1 points greater than that of their peers in control schools (B = 0.103, SE = 0.244, p = .675, Table 4), equivalent effect size 0.04 (95% confidence interval −0.13 to 0.20), which we interpret as small and nonsignificant. Without covariates, the difference between groups was β = 0.187 (SE = 0.287). None of these effects on reading skills reach criterion for significance after adjusting for multiple comparisons (adjusted-alphas .0167, .025, .05 respectively) using the Benjamini – Hochberg procedure (Benjamini & Hochberg, 1995).

The second set of hypotheses in response to RQ2 relate to treatment effects on affective factors. First, note the negative trend observed between baseline and endline. Table 2 indicates that, overall, total scores on both the reading self-efficacy and reading motivation scales tended to decline. The decline appears somewhat less dramatic in the PALS-UK trial-arm. On the self-efficacy subscale, at endline pupils in treatment schools scored around a third of a point higher than their peers in control schools (B = 0.354, SE = 1.037, p = .733, Table 4), equivalent to a small effect size of 0.02 (95% confidence interval −0.09 to 0.12). Without covariates, the difference between groups was B = 0.228 (SE = 1.139). Similarly, in terms of motivation to read we found that pupils in treatment schools scored just under one point higher than control group pupils (B = 0.890, SE = 0.623, p = .153, Table 4), equivalent to a medium effect size of 0.07 (95% confidence interval −0.03 to 0.16;). Without covariates, the difference between groups was B = 0.228 (SE = 1.139). With or without adjustment for multiple comparisons, effects on affective factors are not significant.

Table 5. Subgroup effects.

Model	Estimate [95% CI]	Effect size [95% CI]
From subgroup model		
Low attainment subgroup	1.53 [0.37, 2.69]	0.22 [0.05, 0.39]
SEND subgroup	1.84 [0.62, 3.06]	0.21 [0.07, 0.35]
EAL subgroup	1.08 [−0.43, 2.59]	0.12 [−0.05, 0.29]
FSM subgroup	0.76 [−0.30, 1.81]	0.09 [−0.04, 0.21]
From interaction model		
Low attainment subgroup	0.22 [−0.65, 1.08]	
High attainment subgroup	0.96 [0.05, 1.88]	
SEND	0.97 [−0.02, 1.96]	
non-SEND	0.87 [−0.04, 1.78]	
EAL	0.42 [−0.69,1.52]	
non-EAL	0.95 [0.03, 1.87]	
FSM	0.022 [−0.77, 0.82]	
non-FSM	0.98 [0.056, 1.90]	

Subgroup effects and interactions

Finally, to respond to RQ3 we explore whether different subgroups respond similarly to PALS-UK by examining the effects on reading attainment. Results are presented in Table 5. In each case, estimated subgroup and interaction effects are reported. For the subgroup effect, the primary outcome model is estimated in the relevant subgroup sample only. For example, for the FSM subgroup we compare average outcomes among FSM pupils in treatment schools to FSM pupils in control schools. Interaction models explore whether the combination of a pupil being both in the subgroup and treatment group leads to a different outcome to other groups – this identifies whether the subgroup benefit disproportionately or otherwise from treatment, relative to other groups (Table 5).

These analyses reveal positive subgroup differences for all four subgroups, suggesting that all four have better reading attainment when receiving PALS-UK than their counterparts in the control groups. For the subgroup with low baseline reading attainment, the estimated effect size is 0.22. For the SEND subgroup, the estimated effect size is 0.21. Additionally, interaction effects are reported for all four subgroups, each indicating additional gains relative to their corresponding comparison groups. Whilst the additional gains are relatively small for most subgroups, the interaction effect for SEND was 0.97. This suggests that pupils with SEND in the treatment group improved more in reading attainment than their non-SEND peers.

Discussion

This randomized controlled efficacy trial is the most rigorous test of the impact of PALS on pupils' reading outcomes in whole class mixed ability settings, not only in England but also internationally. We conclude that PALS-UK led to a moderate increase in pupils' reading outcomes. We found moderate treatment effects for reading attainment, comprehension, and fluency/rate. These effects were predicted a priori and found to be significant at a 95% confidence level (although effects on comprehension and fluency/rate were not significant after applying adjustments for multiple comparisons). Our results are broadly consistent with other evidence on PALS (WWC, 2012). Further, our evidence suggests that children from certain subgroups (children with SEND, low baseline reading attainment and EAL) might particularly benefit. Our effect sizes are slightly smaller than previous studies, which is a common outcome for large trials (Zhang et al., 2013). There are several possible explanations, related to variability in sample, treatment effects or the conservative use of intention-to-treat analysis. Since previous studies tended to be smaller and focused on distinct sub-populations, this study provides the most reliable estimate of the effects of PALS-UK on a general population of pupils. Our study has a large sample, reasonably low

attrition, high statistical power and explores effectiveness in mixed ability classrooms. As such, this result provides compelling and generalizable evidence that PALS is an effective approach, which has educationally meaningful positive impacts on pupils' reading attainment, comprehension, and fluency/rate.

To summarize results, for RQ1 meaningful improvements in reading attainment are observed. To contextualize these findings, Lipsey et al. (2012) report that U.S. students in Grades 4–5 (ages 9–10) typically demonstrate an annual gain of 0.40 standard deviation units in reading. Assuming comparable developmental trajectories, the observed effect size of 0.12 standard deviations in this study represents approximately a 30% increase over the expected annual gain for this age group, underscoring the potential educational relevance of the intervention. As for RQ2, after receiving PALS-UK, pupils in treatment schools showed greater reading skills (reading comprehension and reading fluency/rate) than those in control schools, with moderate effect sizes. However, these differences were not statistically significant after adjusting for multiple comparisons. No significant treatment effects were observed for multidimensional fluency, self-efficacy, or motivation. Subgroup analyses were explored in response to RQ3, confirming that pupils with low baseline reading attainment and pupils with SEND gained more than their peers.

PALS-UK has a positive effect on reading attainment

The primary outcome, reading attainment as measured by the PiRA (Ruttle et al., 2021), is the measure by which the formal efficacy of PALS-UK is to be judged. Our hypothesis was confirmed. At endline, the reading attainment of pupils in schools delivering PALS-UK exceeded their counterparts in schools allocated to the business-as-usual control condition. This effect was maintained after controlling for a range of covariates (see Table 4). While gains in terms of raw reading scores appear small (just over one point difference between conditions), the effect size is "medium" sized by the standards of contemporary field trials in education (Kraft, 2020).

PALS-UK has a positive effect on reading skills

RQ2 explored the impact of PALS-UK on specific reading skills (reading comprehension, reading fluency/rate and multidimensional fluency) and affective factors (self-efficacy and motivation). These factors underpin reading attainment and may help to explain why PALS-UK is effective. We found that PALS-UK had positive, moderate treatment effects on reading comprehension and reading fluency/rate measured using the WIAT-III UK-T (Wechsler, 2018). At endline, pupils in treatment schools scored on average 1.47 points higher than controls for reading comprehension and 1.17 points higher for reading fluency/rate. While numerically small and non-significant after accounting for multiple comparisons, these are "moderate" effect sizes which are educationally important. These findings are particularly compelling given that, compared to the primary analysis, the power to detect significant effects is weaker due to the reduced sample size and increased random error resulting from using reading attainment as the baseline covariate (rather than measures of comprehension and fluency at both baseline and endline). This result means that we can be confident that PALS-UK is effective in supporting the very skills that pupils practise most. Repeated exposure to structured use of comprehension strategies scaffolds pupils' reading comprehension skills. The time spent reading aloud with a peer (coach) attending to and supporting reading accuracy supports development of reading efficiency – the ability to read aloud accurately and quickly.

Less positively, the treatment effect was not significant for multidimensional fluency, reading self-efficacy or motivation. To avoid making Type 4 errors, it is important not to over interpret these null results. Nothing in the pattern of results suggests that the treatment had a negative effect on these outcomes. In all cases, any differences between conditions were broadly in the hypothesized direction (toward PALS-UK having a positive impact). There are reasons to suspect that the null results may be

due to limitations in our research design rather than a failure of the treatment to impact on the construct.

The null result for multidimensional fluency contrasted with the significant treatment effect for fluency/rate. The same recordings of children reading aloud were used for both measures, and despite strong correlations between the measures (see Table 3, $r = .7$), differences emerged due to how the measures were scored and the construct(s) they attempt to quantify. Treatment effects were significant for the WIAT-III UK-T oral reading fluency subscale, which quantifies accuracy and speed of oral reading. Scoring is relatively objective and the instrument has robust psychometric properties (Wechsler, 2018). The MDFS acknowledges that skilled oral reading involves more than efficiently reading the words. Expressive reading not only includes changes in intonation but can include varying reading speed in ways which, counterintuitively, result in slower reading. Hence, the MDFS aims to capture these qualitative, prosodic elements of oral reading (Rasinski, 2004a, 2004b; Rasinski et al., 2011). However, because this approach uses subjective, qualitative judgments, it is likely less reliable. This, in turn, increases random error and reduces the likelihood of significant treatment effects. Hence, the null effect for multidimensional fluency could be symptomatic of measurement insensitivity.

Another compelling explanation for differing effects on fluency measures is that this reflects the specific components of fluency that PALS activities treat – they do indeed focus attention on reading accuracy and speed, more so than the prosodic elements of expressive oral reading. Since we know that prosodic reading skills account for additional variance in comprehension after controlling for accuracy and speed (Miller & Schwanenflugel, 2008; Rasinski et al., 2009; Veenendaal et al., 2015), future research might consider adaptations that place greater emphasis on reading with expression. Recent reviews highlight that listening to skilled modeled reading, echo, and choral reading are promising approaches to supporting expressive oral reading (Hudson et al., 2020; Shhub et al., 2023), but these approaches may be less suited to a peer assisted context. Just as it is important to identify the precise skills that an educational approach will influence, it is similarly important to identify the boundaries of the approach. Future research should explore whether these skills are supported, indirectly, from exposure to text through PALS-UK or whether explicit instruction is necessary in other lessons. Other literacy skills that might be examined in this way include vocabulary, inferencing, disciplinary literacy, and writing composition. With this knowledge, educators might consider whether and how they support those skills elsewhere in the curriculum.

It is important to note the potential threat to internal validity caused by reading skills being assessed at endline only and among a smaller, random subsample, compared to the more robust pre-/post-test design used for reading attainment and affective factors. This decision was primarily due to the costs associated with these measures, which needed 1-to-1 administration by trained assessors (other measures were group administered by classroom teachers). Nonetheless, we minimized this threat by using robust randomization procedures to address potential bias and were satisfied that applying reading attainment as a baseline covariate to address unexplained variance.

Overall, our results concur with evidence from alternative approaches to paired reading, which also found gains in reading fluency and comprehension (e.g., Algozzine et al., 2009; Marr et al., 2011). PALS-UK enables pupils to practice reading with immediate peer feedback focused on supporting word reading accuracy and giving opportunities for repeated re-reading – approaches shown to support oral reading fluency (Hudson et al., 2020; Shhub et al., 2023). Our findings confirm that peer assisted reading can be an effective context for this practice and feedback. They confirm that working with a peer is a suitable context to practise sequencing, retelling, and summarizing, and making predictions about what will happen next and that doing so supports reading comprehension skills (Fuchs et al., 1997, 2000; McMaster et al., 2006).

These findings are also consistent with evidence from meta-analyses presented elsewhere, which highlight the particular benefits of training multiple reading comprehension strategies (Peng et al., 2024), using multiple approaches to support reading fluency (Shhub et al., 2023), as well as the benefits of increasing text exposure through wide reading (Zimmermann et al., 2021). We argue that the

combined and interactive effects of these ingredients likely bring about these outcomes, rather than one specific component of PALS. Our results show that integrated instructional approaches that focus on practising specific reading skills (in this case oral reading fluency and reading comprehension strategies) lead to gains in reading attainment, as well as training those specific reading skills (discussed in further depth in Gellen et al., under review - 2026). This is consistent with most theoretical frameworks of reading comprehension, which commonly highlight the multiplicative or interactive nature of underpinning skills and processes (Breadmore et al., 2019; Hogan et al., 2011; Hoover & Gough, 1990; Perfetti, 1999).

Our findings also challenge us to extend existing theoretical frameworks. Typically, models of reading comprehension draw a distinction between processes associated with word identification and comprehension (Breadmore et al., 2019; Hogan et al., 2011; Hoover & Gough, 1990; Perfetti, 1999). Within those models, fluency is often characterized as a word level process. Here, we have shown that instructional approaches that support oral passage reading fluency and use of comprehension strategies led to improvements in reading attainment, comprehension and fluency (rate). This was found even without pupils receiving explicit instruction on word recognition within PALS-UK sessions. This is not to say that word recognition is not an important. It is – readers could not have engaged with PALS-UK unless they had sufficient word reading ability to be able to read passages with some fluency. Nonetheless, our findings suggest that, once word recognition processes are reasonably secure, readers still benefit from instructional approaches that allow them to practice oral reading fluency and use of comprehension strategies. This concurs with evidence presented elsewhere that passage reading fluency contributes more to comprehension than word recognition speed (e.g., Kim, 2015). Others have argued that theories of reading have been preoccupied with understanding how word reading accuracy is achieved, leaving a gap in understanding passage reading fluency (Share, 2021). We extend this argument to suggest that passage reading fluency is a key link between word recognition and comprehension processes. The interactive contribution of fluency and comprehension demands more thorough attention in both empirical investigation and theoretical models of reading.

Future research should examine in more detail how instructional approaches that focus more precisely on the comprehension strategies scaffolded during PALS-UK activities and impact specific comprehension skills. For example, we suspect that comprehension monitoring may be an important mechanism – 70% of teacher respondents delivering PALS-UK agreed that pupils displayed greater "self-monitoring" (Lewin et al., 2024). A more nuanced understanding of whether this approach "works" should consider what makes it work well and who it works for, including qualitative analysis to better understand teachers' experiences and other contextual factors that influence implementation and outcomes. We describe this in more depth when reporting the implementation and process evaluation elsewhere (Vardy, Lewin, et al., In prep).

Affective factors

Exploring treatment effects on affective factors revealed a concerning overall trend worthy of further consideration. In general, pupils' scores on reading self-efficacy and motivation to read decreased from baseline to endline, regardless of whether they were in treatment or control conditions. Others have also reported a downwards trend as pupils progress through school (Wigfield et al., 2016) and large-scale cross-sectional studies suggest societal decline in children's attitudes toward reading in recent years (Clark et al., 2024). Even so, relatively few previous studies have measured these constructs longitudinally using both the same instruments and the same participants, as we did in this study, or focused on this age range. The possibility that differences in mode of administration (digital at baseline, paper at endline) impacted pupil responses cannot be ruled out and is worthy of further investigation. Even so, as the effect of differences in mode of administration would be equal for both PALS-UK and control conditions it does not impact on our interpretation of the effect of PALS-UK. Our study suggests that PALS-UK partially mitigated this decline, but effect sizes were small and non-significant.

The effects on reading self-efficacy cannot be meaningfully interpreted due to the small nonsignificant effect size and only a modest, nonsignificant effect was observed on the motivation to read. There are challenges in measuring affective factors which are also relevant here. For example, the measure may not have captured the affective factors about reading that were influenced by PALS-UK. Some support for this view comes from teachers delivering PALS-UK: 55% agreed that their pupil's interest in reading had improved (Lewin et al., 2024). Future research should continue to explore the role of affective factors about reading in the impact of reading interventions, and for peer mediated approaches it also seems important to consider the role of peer relationships and social skills. It is worth noting that one of the adaptations of PALS-UK compared to the original PALS Grades 2–6 was the removal of a points-based motivational framework. Instead, teacher training explicitly focused on intrinsic motivators (book choice, experience of coaching and shared reading) supplemented by locally relevant motivational frameworks (e.g., school merit systems). This may have led to variation in the approaches teachers used to motivate their pupils. Future research should explore the impact of different motivational practices and how they interact with outcomes in more depth.

PALS-UK is at least as effective for subgroups of pupils.

To answer RQ3, about whether PALS-UK is effective for all pupils, we explored whether there were different treatment effects on reading attainment for pupils with low baseline reading attainment, SEND, EAL or FSM. Both pupils with low baseline reading attainment and those with SEND benefit from the treatment, but not differentially so compared to higher attainers and non-SEND pupils in the same settings (Table 5). The evidence for pupils with SEND is more compelling than for those with low baseline reading attainment, particularly because the lower limit of the 95% confidence interval on the interaction effect, testing whether SEND pupils in the treatment group scored higher in reading attainment than their non-SEND peers, was very close to zero. We interpret this as weak evidence that PALS-UK may be particularly effective for improving reading attainment of pupils with SEND. Note that the power to detect significant differences is reduced in these analyses compared to the primary outcomes analyses with the full sample. Future research should explore subgroup differences in an adequately powered study, to explore the impact on different pupils in more depth. Here though, we can at least confirm that pupils with low baseline reading attainment, SEND, EAL, or FSM do not appear disadvantaged by being placed in classrooms that are delivering PALS-UK.

We should note that converting baseline reading attainment raw scores to standardized scores using norm-referenced information provided in the manual (Ruttle et al., 2021) suggest that our sample had relatively low reading attainment at the beginning of the study (mean cohort standardized scores = 91). This is likely to have resulted from timing and recruitment strategy. The study was conducted in the academic year 2022/23 following substantial disruption to education caused by the response to COVID-19. In addition, funding required that recruitment specifically targeted regions of England where there was concern about reading attainment. It should also be noted that this measure can be administered digitally or on paper (and was standardized using a combination of both administration methods), and within this study all participants completed digital baseline assessment and paper-based endline assessment. This was true for all schools in the trial, and randomization means that this is unlikely to bias the results. Nonetheless, a larger scale effectiveness study is needed to test the impact of the PALS-UK under even more "real-world" conditions, across a wider range of settings, and with more relaxed control over fidelity of implementation. Such a study could not only examine the size of the treatment effect but also explore whether adaptations further improve effectiveness.

This study provides the most robust evidence to date that peer assisted learning strategies (PALS) are effective in improving reading attainment, reading comprehension, and reading fluency/rate. Treatment effects were promising for most measures (and never negative), but statistical significance and effect sizes varied. Moderate, positive treatment effects were observed on reading attainment, reading comprehension, and reading fluency/rate. These compelling findings should assure educators that they can be confident that

in choosing to use this evidence-based approach to paired reading they are likely to see positive impacts on their pupils' reading outcomes. More broadly, these findings contribute to our understanding of reading processes, by demonstrating the important role that oral reading fluency and comprehension strategies have in supporting reading development. They challenge theoretical assumptions that position fluency as a word level process and suggest that passage reading fluency is a key link between word recognition and comprehension.

Notes

1. Free School Meals eligibility is used as a proxy for socio-economic status. Children in England are eligible for FSM if their family receives certain benefits or have an income below a threshold.
2. The sample including all participants who have been assigned to different study groups (e.g., treatment or control).
3. Referring to the participants whose data were included in the final analysis of the primary outcome.
4. Publicly available ratings that assess schools in England across quality of education, behavior, personal development, and leadership. Schools are rated as Outstanding, Good, Requires Improvement, or Inadequate.
5. Using publicly available Key Stage 2 reading scores at the school level.
6. The intraclass correlation coefficients (ICC) are measures of the degree to which pupil are correlated within classes, and within schools. The larger the degree of correlation, the greater the differences in student attainment across clusters. The ICC is a key quantity used in the calculation of sample sizes for cluster randomized controlled trial designs.
7. The pupil month of birth variable is included to account for age because dependent variables are included in their unstandardized form.
8. This is because comprehension and fluency outcomes were not collected from all pupils in Year 5 – as was the case for reading attainment, self-efficacy, and motivation measures – but from a subsample of Year 5 pupils selected at random from one class per school. Thus, for these outcomes there are only two levels in the data: pupils nested within schools. For the other outcomes, there are three levels, pupils nested within classes and classes nested within schools.
9. The ICCs were slightly higher in the model without covariates: 0.12 (95% CI: 0.08 to 0.17) at the school level and 0.04 (95% CI: 0.00 to 0.08) at the class level.

Acknowledgments

Many thanks to the many pupils, teachers and schools involved in this trial as well as previous research which contributed to the development of PALS-UK. We thank the original developers of PALS, particularly Douglas Fuchs, Lynn Fuchs, and Kristen McMaster for their advice on using and adapting PALS in England. We would also like to thank AlphaPlus and FFT Education (for supporting data collection and enumeration of the sample), the many research assistants who supported delivery and data collection, and the many individuals at the EEF who supported delivery and evaluation processes, particularly Aoife Duff.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Education Endowment Foundation and the Department for Education's Accelerator Fund. The EEF Evaluation Report for the trial (https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/PALS_report_final.pdf?v=1734278684) is available under the Open Government Licence (http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3).

ORCID

Helen L. Breadmore http://orcid.org/0000-0003-3050-8908
Stephen P. Morris http://orcid.org/0000-0001-6869-8933
Sandor Gellen http://orcid.org/0000-0001-9137-8209

Cathy Lewin http://orcid.org/0000-0002-3430-4075
Emma J. Vardy http://orcid.org/0000-0002-7825-2270
Steph Ainsworth http://orcid.org/0000-0003-0739-6485
Kate Wicker http://orcid.org/0000-0002-3922-4197
Luisa Tarczynski-Bowles http://orcid.org/0000-0003-3869-1735

Data availability statement

As a condition of the grant agreement provided by the study funders (The Education Endowment Foundation (EEF)), the study data are deposited in the EEF's evaluation data archive. Due to data protection and confidentiality concerns, access to the data is restricted. Researchers wishing to access the trial data need to apply for permission to EEF.

Ethics statement

This research was conducted in accordance with the BPS Code of Human Research Ethics (British Psychological Society, 2021). Approval was granted by a University Research Ethics Committee (and accepted by all other institutional ethics committees), following a detailed review of project design, protocol and procedures. Before data were collected, schools signed a Memorandum of Understanding and Data Sharing Agreement. Parents were informed about the research and given an opportunity to withdraw their children. Children's class teacher also showed them a video (produced by the research team) explaining the evaluation. They were informed that they did not need to complete the evaluation tasks if they did not want to. Test administrators were also trained to look out for any non-verbal signs that children might not be comfortable completing the tasks, and to stop the activities where appropriate (consistent with assent procedures outlined in Dockett & Perry, 2011). The study protocol was preregistered prior to randomization. (https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/pals-uk-accelerator-fund and https://archive.org/details/osf-registrations-f2wgj-v1).

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. https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/PALSUK_Protocol_v1_20221114.pdf?v=1690191477

Algozzine, B., Marr, M. B., Kavel, R. L., & Dugan, K. K. (2009). Using peer coaches to build oral reading fluency. *Journal of Education for Students Placed at Risk (JESPAR)*, *14*(3), 256–270. https://doi.org/10.1080/10824660903375735

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Breadmore, H. L., & Carroll, J. M. (2021). *A systematic review of measures of attainment in literacy, mathematics and science* Report commissioned by the Education Endoment Foundation. London, UK:. https://educationendowmentfoundation.org.uk/public/files/Review_of_Attainment_Measures_final.pdf

Breadmore, H. L., Vardy, E. J., Cunningham, A. J., Kwok, R. K. W., & Carroll, J. M. (2019). *Literacy development: Evidence review*. Evidence Review commissioned by the Education Endowment Foundation. https://educationendowmentfoundation.org.uk/public/files/Literacy_Development_Evidence_Review.pdf

British Psychological Society. (2021). *BPS code of human research ethics*. https://www.bps.org.uk/sites/www.bps.org.uk/files/Policy/Policy%20-%20Files/BPS%20Code%20of%20Human%20Research%20Ethics.pdf

Carroll, J. M., & Fox, A. (2017). Reading self-efficacy predicts word reading but not comprehension in both girls and boys. *Frontiers in Psychology*, *7*(2056). https://doi.org/10.3389/fpsyg.2016.02056

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Clark, C., Picton, I., Cole, A., & Oram, N. (2024). *Children and young people's reading in 2024*. National Literacy Trust. https://nlt.cdn.ngo/media/documents/Children_and_young_peoples_reading_in_2024_Report.pdf

Conradi, K., Jang, B. G., & McKenna, M. C. (2014). Motivation terminology in reading research: A conceptual review. *Educational Psychology Review*, *26*(1), 127–164. https://doi.org/10.1007/s10648-013-9245-z

Culora, A., Dimova, S., Illie, S., Sutherland, A., Gilder, L., Picken, N., Groom, M., Sibuosa, M., Merewood, J., & Speciani, E. R. (2024). *Peer assistent learning strategies - UK: Evaluation report (updated 2024)*. Education Endowment Foundation. https://d2tic4wvo1iusb.cloudfront.net/documents/projects/PALS.pdf?v=1646051946

DfE. (2023a). *Guidance: Choosing a phonics teaching programme*. Department for Education, https://www.gov.uk/government/publications/choosing-a-phonics-teaching-programme/list-of-phonics-teaching-programmes.

DfE. (2023b). *ITT core content framework*. Department for Education, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/974307/ITT_core_content_framework_.pdf.
DfE. (2023c). *The reading framework*. Department for Education, https://www.gov.uk/government/publications/the-reading-framework-teaching-the-foundations-of-literacy.
DfE. (2024). *Key stage 2 attainment, academic year 2022/23*. Department for Education, https://explore-education-statistics.service.gov.uk/find-statistics/key-stage-2-attainment.
Dockett, S., & Perry, B. (2011). Researching with young children: Seeking assent. *Child Indicators Research*, *4*(2), 231–247. https://doi.org/10.1007/s12187-010-9084-0
Dong, N., & Maynard, R. (2013). Powerup!: A tool for calculating minimum detectable effect sizes and minimum required sample sizes for experimental and quasi-experimental design studies. *Journal of Research on Educational Effectiveness*, *6*(1), 24–67. https://doi.org/10.1080/19345747.2012.673143
EEF. (2021a). *The EEF teaching and learning toolkit - peer tutoring*. Education Endowment Foundation. Retrieved November, 2024, from https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/peer-tutoring
EEF. (2021b). *The EEF teaching and learning toolkit - reading comprehension strategies*. Education Endowment Foundation. Retrieved November, 2024, from https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/reading-comprehension-strategies
Fuchs, D., Fuchs, L. S., & Burish, P. (2000). Peer-assisted learning strategies: An evidence-based practice to promote reading achievement. *Learning Disabilities Research & Practice*, *15*(2), 85–91. https://doi.org/10.1207/SLDRP1502_4
Fuchs, D., Fuchs, L. S., Mathes, P., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, *34*(1), 174–206. https://doi.org/10.3102/00028312034001174
Gellen, S., Breadmore, H. L., Morris, S. P., Ainsworth, S., Vardy, E. J., Lewin, C., Wicker, K., & Tarczynski-Bowles, M. L. (under review - 2026). *Understanding how paired reading works: The mediating role of fluency and comprehension in reading attainment*.
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. https://d2tic4wvo1iusb.cloudfront.net/documents/projects/PALS-UK-Statistical-Analysis-Plan.pdf?v=1683118618
Hogan, T. P., Bridges, M. S., Justice, L. M., & Cain, K. (2011). Increasing higher level language skills to improve reading comprehension. *Focus on Exceptional Children*, *44*(3), 1–20. https://doi.org/10.17161/foec.v44i3.6688
Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, *2*(2), 127–160. https://doi.org/10.1007/BF00401799
Hudson, A., Koh, P. W., Moore, K. A., & Binks-Cantrell, E. (2020). Fluency interventions for elementary students with reading difficulties: A synthesis of research from 2000-2019. *Education Sciences*, *10*(3), 1–28. https://doi.org/10.3390/educsci10030052
Hudson, R. F., Lane, H. B., Pullen, P. C., & Torgesen, J. K. (2009). The complex nature of reading fluency: A multidimensional view. *Reading and Writing Quarterly*, *25*(1), 4–32. https://doi.org/10.1080/10573560802491208
Kim, Y. G. (2015). Developmental, component-based model of reading fluency: An investigation of predictors of word-reading fluency, text-reading fluency, and reading comprehension. *Reading Research Quarterly*, *50*(4), 459–481. https://doi.org/10.1002/rrq.107
Kim, Y.-S. G., Quinn, J. M., & Petscher, Y. (2021). What is text reading fluency and is it a predictor or an outcome of reading comprehension? A longitudinal investigation. *Developmental Psychology*, *57*(5), 718–732. https://doi.org/10.1037/dev0001167
Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology*, *100*(2), 310–321. https://doi.org/10.1037/0022-0663.100.2.310
Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189X20912798
Kuhn, M. R., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Review of research: Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly*, *45*(2), 230–251. https://doi.org/10.1598/RRQ.45.2.4
LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*(2), 293–323. https://doi.org/10.1016/0010-0285(74)90015-2
Language and Reading Research Consortium, Jiang, H., & Davis, D. (2017). Let's know! Proximal impacts on prekindergarten through grade 3 students' comprehension-related skills. *The Elementary School Journal*, *118*(2), 177–206. https://doi.org/10.1086/694220
Language and Reading Research Consortium, & Yeomans-Maldonado, G. (2017). Development of comprehension monitoring in beginner readers. *Reading and Writing*, *30*(9), 2039–2067. https://doi.org/10.1007/s11145-017-9765-x
Lewin, C., Morris, S., Ainsworth, S., Gellen, S., & Wicker, K. (2024). *Peer assisted learning strategies - UK (PALS-UK): A whole-class reading approach. Evaluation report*. Education Endowment Foundation. https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/PALS_report_final.pdf?v=1734278684

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S., & Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms*. National Center for Special Education Research.

Marr, M. B., Algozzine, B., Nicholson, K., & Keller Dugan, K. (2011). Building oral reading fluency with peer coaching. *Remedial and Special Education*, *32*(3), 256–264. https://doi.org/10.1177/0741932510362202

McGeown, S. P., Duncan, L. G., Griffiths, Y. M., & Stothard, S. E. (2015). Exploring the relationship between adolescent's reading skills, reading motivation and reading habits. *Reading and Writing*, *28*(4), 545–569. https://doi.org/10.1007/s11145-014-9537-9

McGeown, S. P., Johnston, R. S., Walker, J., Howatson, K., Stockburn, A., & Dufton, P. (2015). The relationship between young children's enjoyment of learning to read, reading attitudes, confidence and attainment. *Educational Research*, *57*(4), 389–402. https://doi.org/10.1080/00131881.2015.1091234

McMaster, K. L., Fuchs, D., & Fuchs, L. S. (2006). Research on peer-assisted learning strategies: The promise and limitations of peer-mediated instruction. *Reading & Writing Quarterly*, *22*(1), 5–25. https://doi.org/10.1080/10573560500203491

Merga, M. K. (2020). "Fallen through the cracks": Teachers' perceptions of barriers faced by struggling literacy learners in secondary school. *English in Education*, *54*(4), 371–395. https://doi.org/10.1080/04250494.2019.1672502

Miller, J., & Schwanenflugel, P. J. (2006). Prosody of syntactically complex sentences in the oral reading of young children. *Journal of Educational Psychology*, *98*(4), 839–853. https://doi.org/10.1037/0022-0663.98.4.839

Miller, J., & Schwanenflugel, P. J. (2008). A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. *Reading Research Quarterly*, *43*(4), 336–354. https://doi.org/10.1598/RRQ.43.4.2

Moser, G. P., Sudweeks, R. R., Morrison, T. G., & Wilcox, B. (2014). Reliability of ratings of children's expressive reading. *Reading Psychology*, *35*(1), 58–79. https://doi.org/10.1080/02702711.2012.675417

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*, *16*(2), 91–121. https://doi.org/10.1080/10888438.2010.529219

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes*, *18*(4), 443–468. https://doi.org/10.1080/01690960344000008

Peng, P., Wang, W., Filderman, M. J., Zhang, W., & Lin, L. (2024). The active ingredient in reading comprehension strategy intervention for struggling readers: A Bayesian network meta-analysis. *Review of Educational Research*, *94*(2), 228–267. https://doi.org/10.3102/00346543231171345

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 167–208). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198507932.003.0006

Rasinski, T., Rikli, A., & Johnston, S. (2009). Reading fluency: More than automaticity? More than a concern for the primary grades? *Literacy Research and Instruction*, *48*(4), 350–361. https://doi.org/10.1080/19388070802468715

Rasinski, T. V. (2004a). *Assessing reading fluency*. Pacific Resources for Education and Learning (PREL). https://files.eric.ed.gov/fulltext/ED483166.pdf

Rasinski, T. V. (2004b). Creating fluent readers. *Educational Leadership*, *61*(6), 46–51. https://www.ascd.org/el/articles/creating-fluent-readers

Rasinski, T. V. (2012). Why reading fluency should be hot! *The Reading Teacher*, *65*(8), 516–522. https://doi.org/10.1002/TRTR.01077

Rasinski, T. V., Reutzel, C. R., Chard, D., & Linan-Thompson, S. (2011). Reading fluency. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. Afflerbach (Eds.), *Handbook of reading research* (Vol. 4, pp. 286–319). Routledge.

Ravid, D., & Mashraki, Y. E. (2007). Prosodic reading, reading comprehension and morphological skills in Hebrew-speaking fourth graders. *Journal of Research in Reading*, *30*(2), 140–156. https://doi.org/10.1111/j.1467-9817.2007.00340.x

Ruttle, K., McCarty, C., Lallaway, M., Bennett, M., Pepper, L., Kilburn, V., Swift, J., Kiek, T., & Hiorns, C. (2021). *New PiRA primary: Progress in reading assessment year R, Key Stage 1 and Key Stage 2 test guidance (September 2021)*. RS Assessment from Hodder Education.

Sáenz, L. M., Fuchs, L. S., & Fuchs, D. (2005). Peer-assisted learning strategies for English language learners with learning disabilities. *Exceptional Children*, *71*(3), 231–247. https://doi.org/10.1177/001440290507100302

Schreiber, P. A. (1991). Understanding prosody's role in reading acquisition. *Theory Into Practice*, *30*(3), 158–164. https://doi.org/10.1080/00405849109543496

Share, D. L. (2021). Is the science of reading just the science of reading English? *Reading Research Quarterly*, *56*(S1). https://doi.org/10.1002/rrq.401

Shhub, A., Jimenez, Z., & Solis, M. (2023). A synthesis of reading prosody: Evaluating phrasing and syntax interventions. *Reading & Writing Quarterly*, *39*(6), 530–547. https://doi.org/10.1080/10573569.2022.2147464

Smith, G. S., & Paige, D. D. (2019). A study of reliability across multiple raters when using the NAEP and MDFS rubrics to measure oral reading fluency. *Reading Psychology*, *40*(1), 34–69. https://doi.org/10.1080/02702711.2018.1555361

STA. (2012). *Assessment framework for the development of the Year 1 phonics screening check*. Crown Copyright Retrieved from https://www.gov.uk/government/publications/assessment-framework-for-the-development-of-the-

year-1-phonics-screening-check/assessment-framework-for-the-development-of-the-year-1-phonics-screening-check

STA. (2015). *English reading test framework: National Curriculum tests from 2016*. For text developers. Crown Copyright. https://assets.publishing.service.gov.uk/media/5a82454eed915d74e3402a2a/2016_KS2_Englishreading_framework_PDFA_V3.pdf

STA. (2023). *English reading test mark schemes: Reading answer booklet*. Crown Copyright. https://assets.publishing.service.gov.uk/media/6464d5a0d3231e000c32db4c/2023_key_stage_2_English_reading_answer_booklet.pdf

Stafura, J., & Perfetti, C. A. (2017). Integrating word processing with text comprehension. In R. K. Parrila, K. Cain, & D. L. Compton (Eds.), *Theories of reading development*. John Benjamins Publishing Company, 9–32. https://doi.org/10.1075/swll.15.02sta

Topping, K., Miller, D., Thurston, A., McGavock, K., & Conlin, N. (2011). Peer tutoring in reading in Scotland: Thinking big. *Literacy*, *45*(1), 3–9. https://doi.org/10.1111/j.1741-4369.2011.00577.x

Torgerson, C., Brooks, G., Gascoine, L., & Higgins, S. (2018). Phonics: Reading policy and the evidence of effectiveness from a systematic 'tertiary' review. *Research Papers in Education*, *34*(2), 1–31. https://doi.org/10.1080/02671522.2017.1420816

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.

Vardy, E. J., Al Otaiba, S., Breadmore, H. L., Kung, S. H., Pétursdóttir, A. L., Zaru, M. W., & McMaster, K. L. (2022). Teacher-researcher partnership in the translation and implementing of PALS (peer-assisted learning strategies): An international perspective. *Journal of Research in Reading*, *45*(3), 517–526. https://doi.org/10.1111/1467-9817.12404

Vardy, E. J., Breadmore, H. L., & Carroll, J. M. (in press). Measuring the will and the impact on the skill of reading: Validation of new motivation and self-efficacy to read scales.

Vardy, E. J., Breadmore, H. L., & McMaster, K. (in press). *Implementation and process evaluation (IPE) guide: A case study using the reading programme Peer Assisted Learning Strategies (PALS)*.

Vardy, E. J., & Lewin, C. Ainsworth S. Wicker K. Tarczynski-Bowles L. Breadmore H.L. Morris S. Gellen S. (in press). *Translating an evidence informed reading program to classroom practices: Lessons learned from an implementation and process evaluation of PALS-UK*.

Veenendaal, N. J., Groen, M. A., & Verhoeven, L. (2015). What oral text reading fluency can reveal about reading comprehension. *Journal of Research in Reading*, *38*(3), 213–225. https://doi.org/10.1111/1467-9817.12024

Wechsler, D. (2018). *Wechsler individual achievement test - third UK edition for teachers (WIAT-IIIUK-T): Examiner's manual*. Pearson Assessment.

Wigfield, A., Gladstone, J., & Turci, L. (2016). Beyond cognition: Reading motivation and reading comprehension. *Child Development Perspectives*, *10*(3), 190–195. https://doi.org/10.1111/cdep.12184

Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading*, *5*(3), 211–239. https://doi.org/10.1207/s1532799xssr0503_2

WWC. (2012). *Peer assisted learning strategies. What Works Clearinghouse intervention report*. What Works Clearinghouse. https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_pals_060512.pdf

Zhang, Z., Xu, X., & Ni, H. (2013). Small studies may overestimate the effect sizes in critical care meta-analyses: A meta-epidemiological study. *Crit Care*, *17*(1), R2. https://doi.org/10.1186/cc11919

Zimmermann, L. M., Reed, D. K., & Aloe, A. M. (2021). A meta-analysis of non-repetitive reading fluency interventions for students with reading difficulties. *Remedial and Special Education*, *42*(2), 78–93. https://doi.org/10.1177/0741932519855058

Zutell, J., & Rasinski, T. V. (1991). Training teachers to attend to their students' oral reading fluency. *Journal of Theory into Practice*, *30*(3), 211–217. https://doi.org/10.1080/00405849109543502