Peer Assisted Learning Strategies – UK (PALS-UK): A whole-class reading approach

Evaluation report

October 2024

Cathy Lewin, Stephen Morris, Steph Ainsworth, Sandor Gellen, and Kate Wicker

The Education Endowment Foundation is an independent charity dedicated to breaking the link between family income and educational achievement. We support schools, nurseries and colleges to improve teaching and learning for 2 – 19-year-olds through better use of evidence.

We do this by:

- **Summarising evidence.** Reviewing the best available evidence on teaching and learning and presenting in an accessible way.
- **Finding new evidence.** Funding independent evaluations of programmes and approaches that aim to raise the attainment of children and young people from socio-economically disadvantaged backgrounds.
- **Putting evidence to use.** Supporting education practitioners, as well as policymakers and other organisations, to use evidence in ways that improve teaching and learning.

We were set-up in 2011 by the Sutton Trust partnership with Impetus with a founding of £125m grant from the Department for Education. In 2022, we were re-endowed with an additional £137m, allowing u to continue our work until at least 2032.

For more information about the EEF or this report please contact:

Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP

info@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents

About the evaluator

The project was independently evaluated by a team from Manchester Metropolitan University:

Cathy Lewin, Stephen Morris, Steph Ainsworth, Sandor Gellen, and Kate Wicker. The co-principal investigators were Professor Cathy Lewin and Professor Stephen Morris.

Contact details:

Dr Steph Ainsworth,

School of Education,

Faculty of Health and Education,

Manchester Metropolitan University

Email: s.ainsworth@mmu.ac.uk

Acknowledgements

The evaluation team would like to thank all the schools, teachers, and pupils that were involved in this evaluation for engaging with Peer Assisted Learning Strategies – UK (PALS-UK) fully and sharing their experiences with us.

We would also like to thank FFT Education and AlphaPlus for working professionally and efficiently on our behalf to collect data from the schools and the pupils.

Thanks also to staff at Rising Stars who provided support and guidance for the online and paper-based administration of the New Progress in Reading Assessment (PiRA) reading test.

We would further like to thank the teams from Nottingham Trent University and the University of Birmingham who were responsible for recruiting schools and delivering PALS-UK.

Executive summary

The project

Peer Assisted Learning Strategies – UK (PALS-UK) is a structured peer reading programme that aims to raise reading attainment by developing pupils' oral reading fluency and reading comprehension, as well as reading motivation and self-efficacy (pupil confidence in their own reading abilities). PALS-UK is delivered to the whole class and is centred on pupils working in pairs to engage in four activities: partner reading; retell; paragraph shrinking; and prediction relay.

The programme was delivered to Year 5 primary school pupils (age 9–10). The intervention was delivered three times per week for 20 weeks, with each session lasting around 35 minutes. The first four weeks focused on training pupils on the PALS-UK activities and on working together in pairs. Teachers received one day of initial training followed by a 90-minute top-up training session, a programme manual, a set of 20 children's books, and access to online support and twilight webinars. PALS was developed by Douglas Fuchs, Lynn Fuchs, and other researchers from Vanderbilt University. The intervention was adapted for the UK context and delivered by Dr Emma Vardy (Nottingham Trent University) and Professor Helen Breadmore (University of Birmingham).

This efficacy trial was a two-armed, cluster randomised controlled trial. A total of 114 schools, including 4,840 pupils, participated in the evaluation. The trial evaluated the impact of PALS-UK on reading attainment, measured using the New Progress in Reading Assessment (PiRA) Summer Term Year 5 Test. It also evaluated the impact of the programme on oral reading fluency, comprehension, motivation, and self-efficacy. An implementation and process evaluation (IPE) was conducted to explore participants' experiences of PALS-UK. This included teacher and headteacher surveys, interviews, and focus groups with pupils. The study was funded by the Education Endowment Foundation (EEF) through the Department for Education's Accelerator Fund. The trial started in September 2022 and ended in summer 2023.

Table 1: Key conclusions

	Conclusions
1.	Children in PALS-UK schools made the equivalent of two months' additional progress in reading, on average, compared to children in other schools. This result has a moderate to high security rating
2.	Children in PALS-UK schools made the equivalent of two months' additional progress in reading comprehension, on average, compared to children in other schools
3.	There were mixed results in relation to the impact of the programme on oral reading fluency, with children making additional progress on one measure of fluency (oral reading fluency rate) but not on another broader measure
4.	Pupils eligible for free school meals (FSM) in PALS-UK schools made the equivalent of one months' additional progress in reading, on average, compared to children eligible for FSM in the controls schools. This result is less secure than the findings based on all children in PALS-UK schools. Further analysis suggests that the impact of PALS-UK on children eligible for FSM was very similar to the impact for other children
5.	Overall, the programme was delivered with high fidelity, although a minority of teachers reported making minor adaptations or considering adapting the programme. In some cases, this was due to concerns around the accessibility of PALS-UK for certain groups, such as children with lower attainment

EEF security rating

These findings have a moderate to high security rating. This was an efficacy trial, which tested whether the intervention worked under developer-led conditions in a number of schools. The trial was a well-designed, two-armed, randomised controlled trial. The trial was not as well-powered as originally intended because some recruited schools withdrew before the trial started. Around 25% of pupils who started the trial were not included in the final analysis. Reasons for this included schools withdrawing from the evaluation, pupils being absent on assessment days, and pupil assessments from two schools lost in transit before they could be marked. Pupils in PALS-UK schools were similar to those in the comparison schools in terms of prior attainment.

Additional findings

Pupils in PALS-UK schools made, on average, two additional months' progress compared to pupils in control schools. This is our best estimate of impact, which has a moderate to high security rating. As with any study, there is some

uncertainty around the result: the possible impact of this programme ranges from no additional progress to positive effects of up to three months' additional progress.

The results suggest PALS-UK had a positive impact on children's reading attainment and reading comprehension. In addition, the evaluation found evidence of a small positive impact of PALS-UK on pupils' motivation to read, but no evidence that children in PALS-UK schools had higher reading self-efficacy (confidence in their own reading abilities) than pupils in the comparison schools. There were mixed findings in relation to the impact of the programme on oral reading fluency, with a small positive impact found on reading fluency rate but no evidence of impact on a broader measure that assesses expression and intonation, phrasing, smoothness, and pace of reading. These two measures focus on different elements of oral reading fluency, so the discrepancy could suggest that PALS-UK impacts some elements but not others. It is also possible that the broader fluency measure was less sensitive than the measure of reading fluency rate.

Pupils in PALS-UK schools who qualified for FSM, pupils with special educational needs and disabilities (SEND), and low attainers all saw improvements in their reading attainment compared to their counterparts in schools that did not receive PALS-UK. There is tentative evidence to suggest that pupils with SEND may have benefited disproportionately from PALS-UK when compared to other children in schools that received PALS-UK, perhaps because the one to one support provided by PALS-UK may be particularly beneficial for pupils with SEND.

The perceptions of programme participants were mostly consistent with the findings of the impact evaluation, with staff and pupils generally feeling that PALS-UK had positively impacted reading attainment, reading comprehension, and oral reading fluency. Teachers and pupils also perceived improvements in enjoyment of reading, reading for pleasure, self-monitoring during reading, and peer support. There were concerns, however, that PALS-UK may not fully prepare pupils for Key Stage 2 national curriculum tests. PALS-UK is not designed to address all aspects of reading development or all aspects of the reading curriculum, so schools taking part need to ensure that other important areas of learning are being addressed outside of PALS-UK sessions.

The findings of this evaluation align with evidence from previous evaluations of PALS conducted in the United States and evaluations of other peer-assisted reading programmes where gains in fluency and comprehension have been reported (Algozzine *et al.,* 2009; Marr *et al.,* 2011). There is also wider evidence suggesting that some of the approaches used in PALS-UK, including peer tutoring and the explicit teaching of reading comprehension strategies, can be effective for improving children's attainment (EEF, 2021). The EEF previously commissioned a trial of PALS-UK, but the evaluation was disrupted by the COVID-19 pandemic. This second trial aimed to provide more robust evidence of the impact of PALS-UK.

Cost

The average cost of PALS-UK for one school was around £630, or £12.69 per pupil per year when averaged over three years. This estimate assumes an average of 25 pupils per class, and that no additional top-up training is needed after the first year. Training typically required 15 hours of staff time per teacher. Outside of the time spent delivering the programme to children, teachers also reported spending an average of three hours observing peers delivering the programme and an average of 29 hours over the course of one academic year preparing for sessions, although 74% of teachers suggested that this was not extra time above what they would usually spend preparing for lessons.

Impact

Table 2: Summary of impact on primary outcome(s)

Outcome/group	Effect size (95% confidence interval)	Estimated months' progress	EEF security rating	No. pupils	P-value	EEF cost rating
Reading attainment	0.12 (0.01, 0.23)	+2	🔒🔒🔒🔓🔓	3,628	0.03	£ £ £ £ £
Reading attainment, pupils eligible for FSM	0.09 (-0.04, 0.22)	+1	N/A	1,203	0.16	N/A

N/A, not applicable.

Introduction

Background

Raising standards in reading has been a longstanding priority in England (DfE, 2023). Reading attainment is fundamental to learning across the curriculum, yet 27% of pupils failed to attain the expected standard in reading at the end of Key Stage 2 in the summer of 2023 (DfE, 2024). Pupils who are from socioeconomically disadvantaged backgrounds are less likely to read at the expected standard (or above) and schools cite reading as one of the main barriers to attainment more broadly (alongside attendance) within this group (EEF, 2023). Following the disruption to children's learning as a result of the COVID-19 pandemic, the need to 'get pupils back on track' with their reading is arguably more urgent than ever (Ofsted, 2023), especially for disadvantaged pupils, whose learning is likely to have been differentially impacted by school closures, relative to their peers (Cullinane and Montacute, 2020).

The Department for Education (DfE) is committed to raising reading standards, and to narrow the associated 'disadvantage gap'; however, schools report the challenges in achieving this vision, when children's needs are so diverse (Ofsted, 2023), especially in the face of budgetary pressures and problems recruiting sufficient numbers of teaching assistants (Lucas *et al.,* 2023). Within this study, we evaluate the effectiveness of an intervention designed to support progress in reading: Peer Assisted Learning Strategies – UK (PALS-UK)[1] (Fuchs *et al.,* 1997). PALS-UK is a relatively low-cost option for schools who are seeking to accelerate pupil progress in reading as it can be delivered with a whole class using minimal resources. A key element of the design is that it is peer-led so that pupils receive support from their partner throughout the lesson rather than being reliant solely on the teacher (and/or teaching assistant).

PALS-UK is a whole-class intervention designed to improve children's oral reading fluency, reading comprehension, and confidence in reading. Pupils work in pairs, taking it in turn as coach and reader as they engage with four activities: partner reading; retell; paragraph shrinking; and prediction relay. Sessions last 35 minutes, conducted three times a week over 20 weeks. The logic model predicts that repeated reading with peer feedback will support all aspects of fluency: accuracy; automaticity; and prosody, while the tasks of retell, paragraph shrinking, and prediction relay will support reading comprehension. Taken together, the intervention has the potential to develop pupils' fluency, self-efficacy in reading, reading comprehension, and reading attainment.

Peer interventions are low cost and have been found to generate moderate to high effect sizes (EEF, 2021; Topping *et al.,* 2011). While PALS has been studied widely in the United States, mixed effects have been reported (WWC, 2012), and there is a lack of independent evaluations investigating PALS. UK research is limited and results of same-age peer tutoring programmes disappointing (Lloyd *et al.,* 2015). PALS-UK has been identified as having potential for scale-up due to its structured whole-class, same-age approach.

An evaluation of PALS-UK was undertaken to strengthen the evidence base around peer-assisted reading programmes. We evaluated the causal mechanisms proposed within the logic model through careful selection of primary and secondary measures and a comprehensive mixed-methods approach (Morris *et al.,* 2016). The evaluation was designed as a two-armed, cluster randomised controlled trial across three English Regional School Commissioner regions: the North, the East Midlands and the Humber, and the West Midlands. While 114 schools were recruited, baseline and endline data were only available for 103 schools of which 53 were assigned to the intervention arm and 50 were assigned to the 'business as usual' control arm of the trial. Pupils entering Year 5 in September 2022 in schools allocated to receive the intervention had the opportunity to participate in PALS-UK.

This evaluation of PALS-UK was comprised of three strands: i) an impact evaluation, to assess whether PALS-UK led to an observable improvement in outcomes for Year 5 pupils; ii) an implementation and process evaluation (IPE), which provided contextual information around programme delivery and the 'business as usual' scenario during the trial; and iii) a cost evaluation, which provided information on the delivery costs incurred by intervention schools due to delivering PALS-UK.

The evaluation was funded by the Education Endowment Foundation (EEF). The programme was delivered by the PALS-UK delivery team (Nottingham Trent University and the University of Birmingham) and the evaluation was carried out by Manchester Metropolitan University—the independent evaluator appointed by the EEF.

[1] There are two versions of PALS: one designed to support reading; and one designed to support mathematics. This evaluation focuses on the support to reading.

This trial is informed by a previous funded efficacy trial by the EEF of PALS-UK (Culora *et al.,* 2022), which was conducted recently but was impacted by the COVID-19 pandemic. The current trial addressed two important shortcomings of the previous trial relating to COVID-19 disruption: the delay of post-test measures due to lockdown restriction, leading to a likely diluted impact of PALS-UK on the intervention; and failure to complete the measurement of reading comprehension and oral reading fluency measures. As stated in the previous evaluation report, 'a key missing link here is whether exposure to the programme first results in improvements in reading fluency and comprehension' (Culora *et al.,* 2022, p. 100). The current design ensured that post-test measures were taken close to the end of the intervention and that robust measures of the intermediate outcomes, reading comprehension, and oral reading fluency, were collected, thus facilitating evaluation of key aspects of the logic model. The logic model itself has been augmented to include elements, which were observed to be important in the previous trial.

The intervention for the current trial was largely the same as for the previous evaluation, but with a few minor changes to the support provided to teachers. For example, the manual and initial top-up training materials were updated following feedback from the first trial, and the book list was updated. As in the previous trial, ongoing support was provided for teachers in the intervention schools in the form of 'just-in-time' support. This support was designed to be more extensive and include more components, which could be accessed asynchronously online, such as additional support videos, frequently asked questions (FAQs), discussion boards, as well as twilight support webinars. Unlike the previous trial, in this trial schools were encouraged to use the baseline reading attainment data to inform how children were paired in the intervention for the initial four weeks. When pairing the children, first readers remained first readers and second readers remained second readers throughout the 20 weeks.

Recruitment procedures were similar to the first trial except the exclusion criteria were revised to focus on the three required English Regional School Commissioner regions, schools must not have been taking part in another Accelerator Fund efficacy trial or another literacy project by the EEF, schools must not have received PALS training previously, and schools had to be able to provide access to technology for online assessment of reading. The full exclusion criteria are listed below.

There were a number of changes made to the evaluation from the previous trial. In the previous trial the New Progress in Reading Assessment (PiRA) Autumn Term Year 5 Test was administered at baseline in the Autumn Term. In the current trial, baseline testing occurred in the Summer Term of Year 4 and the New PiRA Summer Term Year 4 Test was used. This is because New PiRA is designed to be administered in the second half of a term (after a certain amount of content has been delivered). So, the New PiRA Autumn Year 5 Test was not deemed appropriate for administration in September 2022 as much of the content covered in the test would not yet have been taught. In addition, the decision to administer PiRA in Year 4 allowed us to: minimise the burden on schools at the beginning of the intervention in the Autumn Term; make scores available for teachers in good time to allocate roles as first/second readers; and ensure sufficient time for randomisation, notifying schools and arranging the initial training. The Wechsler Individual Achievement Test – Third UK Edition for Teachers (WIAT-III UK-T) was used to assess the secondary outcomes of reading comprehension and oral reading fluency. In the previous trials, measurement of these secondary outcomes could not be completed because of disruption related to the pandemic. In addition, we also included a complementary measure of fluency alongside the WIAT-III UK-T. While the WIAT-III UK-T oral reading fluency subtest provides a basic measure of fluency (number of words correct per minute), the additional measure, the Multidimensional Fluency Scale (MDFS) (Rasinski, 2004) provides a qualitative measure of fluency based on judgements of: expression and volume; phrasing; smoothness; and pace. As in the previous project, the measure of self-efficacy was completed pre- and post-test online. The timing of the primary outcome post-test was also different. In this trial, the testing took place in the term following the end of intervention (Summer Term 2023), whereas in the previous trial testing was delayed by at least six months due to pandemic-related disruption. In both trials, the post-tests for the New PiRA reading assessments were administered by the teachers. In the first trial, this was because of 'no-visitor' policies in response to the pandemic. In the current trial, post-testing was administered by teachers to give them flexibility over when to conduct the test and to reduce costs. Many schools already used New PiRA reading assessments. In the previous trial, the WIAT-III UK-T reading comprehension subtest was administered remotely. In the current trial, both the oral reading fluency and reading comprehension subtests were conducted face to face by test administrators to avoid the potential logistical issues associated with administrating a fluency test remotely and given that normative data from the WIAT-III UK-T were collected face to face.

Further points of departure from the previous trial are as follows. We worked with our partner, FFT Education (part of the Fischer Family Foundation), to automate collection of enumeration data. Unlike in the previous trial, randomisation was not stratified by region. Instead, we introduced percentage free school meals (FSM) eligibility as the second stratifying variable alongside school size. The previous trial used age-standardised New PiRA scores in the primary

analysis as an outcome measure. The current trial used raw (unstandardised) New PiRA scores and controlled for the effect of age by adding it to the multilevel regression model as a pupil-level covariate. The previous trial analysed only a single class in each school, and therefore used two-level clustered designs (pupils nested in schools). We used three-level designs to account for school-, and class-level intra-cluster correlations (ICCs). The previous study used three subcategories: children who speak English as an Additional Language (EAL); children eligible for FSM; and high versus low reading achievers for their subgroup analysis. In the current trial, the subgroups of interest will be pupils that are ever-FSM, designated special educational needs and disabilities (SEND), and pupils scoring in the lowest quartile on the baseline New PiRA Summer Year 4 Test. Further exploratory analysis will examine the effects of PALS-UK for EAL pupils, but we will include only those children whose score on the baseline reading assessment falls in the lower half of the sample distribution (see the 'Impact Evaluation' section for details). The previous trial set the following compliance criteria: i) attendance at all PALS-UK initial training sessions; ii) completion of the four weeks of training to the manual; and iii) completion of PALS-UK delivery (minimum of 12 weeks delivered). The current trial defined compliance based on two criteria: i) attendance at PALS-UK training; and ii) completion of the four weeks of pupil training in line with the manual.

In relation to the IPE, in comparison to the original trial of PALS-UK, this trial collected more data. The original trial only conducted a post-intervention survey with teachers in the intervention arm of the trial. In this evaluation, all teachers @were surveyed soon after randomisation. All teachers were surveyed after the intervention was completed including those in control schools. The evaluation team conducted independent observations of PALS-UK through an onsite case study visit; observations were not undertaken by the evaluators in the first trial. The first evaluation collected interview data from headteachers and teachers at case study schools remotely. These data were collected during the onsite case study visit and on a second occasion remotely. We also collected data from pupils through conducting pupil focus groups in case study schools at the onsite visit. The previous trial did not collect data from pupils at case study schools. The collection of weekly teacher logs (recording delivery of PALS-UK sessions) was conducted for this trial but not the previous one.

Intervention

Why: Rationale, theory, or goal of the elements essential to the programme

PALS-UK is a whole-class intervention designed to improve reading comprehension and oral reading fluency. The intervention provides intensive practice of reading skills and is centred around pupils working together in pairs engaging in a set of structured reading activities three times a week for 20 weeks. For the first four weeks, pupils are trained on the PALS-UK activities and how to work well together, then in the following 16 weeks there is an emphasis on peer-directed learning in pairs, but with direction from the teacher as needed (teacher circulates to listen to readers, maintains timings, and keeps pupils on task). Each PALS-UK session is intended to take 35 minutes. Within each pair, one of the pupils is given the role of first reader. This is the pupil who the teacher considers to have a higher current attainment of reading relative to the other pupil in the pair (the second reader). The teachers' decisions around pairings are informed by data in relation to pupils' prior reading attainment, which in this trial was provided by the New PiRA Year 4 Test scores and supplemented with teacher judgement. Within the activities, pupils take turns to act as both coach and reader. Pairing is designed so that the gap in reading skills between pupils is not too wide and supports effective coaching. Together, pairs read a book that is at an appropriate level and is of interest to maintain pupil motivation. During the lessons, pupils receive differentiated instruction through feedback from peers and support from teachers as appropriate. Positive reinforcement is encouraged to motivate readers.

The lessons follow the same routine throughout the programme that is intended to keep pupils focused. Each of the PALS-UK sessions follows the structure below:

- partner reading (ten minutes): Each pupil reads for five minutes while the other pupil coaches them;
- retell (two minutes): The second reader retells what has just happened in the text that they have just read;
- paragraph shrinking (ten minutes): The first reader reads the next paragraph and then is coached to summarise the paragraph; and
- prediction relay (ten minutes): The pupils take turns to predict what will happen in the next half a page, they then read the page and check whether the prediction comes true.

Partner reading aims to develop oral reading fluency by providing pupils with regular opportunities to read aloud to each other. The first reader reads first, providing a model for the second reader, and the 'check-it' procedure supports

development of accuracy. Retell, paragraph shrinking, and prediction relay aim to support reading comprehension. The retell activity encourages pupils to read for meaning and develops their ability to accurately sequence and summarise events from the text. Paragraph shrinking also develops these skills with a particular emphasis on concisely summarising key information from the text. Prediction relay aims to develop pupils' ability to monitor their understanding as they read and to make predictions that are grounded in what they already know from their reading. All PALS-UK activities are designed to develop reading self-efficacy as they involve scaffolded tasks in a safe learning environment supported by peers. The programme is also designed to enhance pupil relationships through positive and productive interactions.

Who: Recipients of this programme

PALS-UK was delivered to Year 5 pupils (aged 9–10 years).

What: Physical or informational materials used in the programme

PALS-UK was delivered using resources developed by Nottingham Trent University and the University of Birmingham, supported by the original developers of PALS from the United States, the Fuchs Research Group at Vanderbilt University. Year 5 teachers at intervention schools were provided with a PALS-UK manual, which provided detailed instructions for the whole-class training of pupils and explained how to deliver the intervention, including guidance on pairing pupils. Teachers were also provided with videos demonstrating good practice and common mistakes, materials for pupils such as prompt cards, scripted lesson plans for pupil training, and a selection of 20 reading books.

What: Procedures, activities, and/or processes used in the programme

Participating Year 5 teachers received initial and top-up training from the developer team: Dr Emma Vardy and Luisa Tarczynski-Bowles at Nottingham Trent University and Dr Helen Breadmore at University of Birmingham. The first session was face to face, taking place before the intervention began, lasted for one day and provided the teachers with training around the skills needed for reading comprehension as well as a detailed description of the intervention. This training day also provided the teachers with the opportunity to practice the PALS-UK activities and to discuss key implementation issues such as how to pair pupils appropriately. The second training session was online and was a shorter online top-up session (90 minutes after school), which took place four weeks into the intervention. This session was designed to support teachers in selecting books appropriately, changing pupil pairings effectively (this is supposed to happen approximately every four weeks), and providing opportunities to share good practice.

The developers also provided additional 'just-in-time' support throughout the intervention in the form of discussion boards, twilight webinars, and the opportunity for teachers to contact the developers with any questions or problems at any time.

Once the teachers received their pre-intervention training, PALS-UK started with four weeks of whole-class training for the pupils. Pupils were trained to work effectively in their pairs and to work through the sequences of activities set out above. Within each training session the teacher introduced a new skill so as not to overload the pupils. Pairings were based on the teacher's assessment of the children's current reading attainment. As described above, this assessment was based on the New PiRA Year 4 Test data, which was shared back to schools, supplemented by teacher judgement (e.g. for pupils who did not complete the PiRA assessment). Teachers performed a median split based on teacher assessment of reading level. The top half of the class were assigned as first readers and the bottom half of the class were assigned as second readers—these assignments were maintained throughout the 20-week intervention. Teachers were encouraged to think carefully about whether they thought the PiRA data gave a true reflection of each pupil's reading ability, and to change the rank order if they felt it did not reflect the pupil's ability. Initially, it was recommended that teachers pair the highest attainer from the above-median group with the highest attainer of the below-median group, and so on. This was to ensure a difference between the two levels of reading attainment but one that was not too large. Pupils were guided towards books of an appropriate reading level. They were then encouraged to choose a book that was of interest, which needed to be agreed upon by the pair of readers.

Who: Programme providers/implementers

PALS was developed by the Fuchs Research Group at Vanderbilt University.

The delivery team, based at Nottingham Trent University and the University of Birmingham, adapted the PALS resources for delivery in English schools. This included extending/modernising the teacher training materials, modernising the manual, providing 'just-in-time' support for teachers, conducting observations of PALS-UK being implemented, and monitoring teacher logs to support implementation fidelity.

Year 5 teachers in English primary schools implemented PALS-UK in their classrooms and other members of staff acted as peer observers. Peer observers were either another Year 5 teacher (if a multi-form entry school) or a teacher from another year group (if a single-form entry school), nominated by the school.

How and where? Mode and location of delivery

The first training day was face to face, in regional locations so that teachers did not have to travel too far. Other training and support for teachers was provided online.

PALS-UK was delivered in Year 5 classrooms to the pupils across three English Regional School Commissioner regions: the North; the East Midlands and the Humber; and the West Midlands.

When and how much? Duration and dosage of the programme

The intervention took place from October 2022 to May/June 2023.

PALS-UK lasts for 20 weeks in total, which involved four weeks of whole-class pupil training, followed by 16 weeks of the PALS-UK sessions with pupils changing pairs (but not roles) every four weeks.

The report from the previous evaluation suggests that the intervention can be delivered with excellent fidelity (e.g.100% teachers completing the survey reported that they implemented the full 20 weeks of the intervention and 95% of teachers attended both training sessions).

Tailoring? Adaptation of the programme

Some flexibility was allowed in the pairing process so that teachers could take into account other factors (e.g. personality) that might affect the pupil's ability to work well as a pair, and teachers were advised to change pairings around, if any of the pairings were not working well.

Teachers were told that pupils with severe SEND or who were at the early stages of learning English could have a teaching assistant join their pair to act as an additional coach, if appropriate. Teachers were also told that if a pupil's needs were so severe that they were unable to engage with PALS-UK even with very simple text, then they could choose to use this time for targeted interventions. It is important to note that we would expect any such pupils to be excluded from the evaluation of PALS-UK due to not being able to provide New PiRA baseline data.

How well (planned): Strategies to maximise effective implementation

To support schools with the delivery of PALS-UK, four observations during the 20 weeks took place. These observations were meant to be supportive and help teachers. Two observations were completed by research assistants from Nottingham Trent University; the first observation took place during the initial four weeks of pupil training and the second observation took place during the 16 weeks of delivery. These observations helped to identify schools where additional support was needed. Two observations were also completed by the nominated peer observer within the school, following the same process as the research assistant; this was designed to elicit a professional opinion on how PALS-UK fits within the school environment and to create resilience, ensuring that another member of staff has familiarity with the programme and can support delivery in the event of staff absence. Both of these observations took place during the 16-week part of the programme. A structured observation sheet developed by Professor Kristen McMaster (Vanderbilt University) was used by the peer observer and the research assistants.

The logic model

The logic model (**Figure 1** below) lays out the theoretical relationships between elements of the intervention and the pupil outcomes that were measured in the impact evaluation. PALS-UK is primarily designed to develop oral reading fluency and reading comprehension. The emphasis on regular practice at reading aloud, re-reading, and receiving feedback from a peer is predicted to improve reading fluency (Rasinski, 2003). The 'check-it' procedure encourages children to read accurately as well as supporting fluency, as they have to re-read aloud after their partner says 'check-it'. The other activities of retell, paragraph shrinking, and prediction relay are each designed to encourage children to make meaning as they read, thus supporting them with opportunities to develop their reading comprehension skills.

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

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

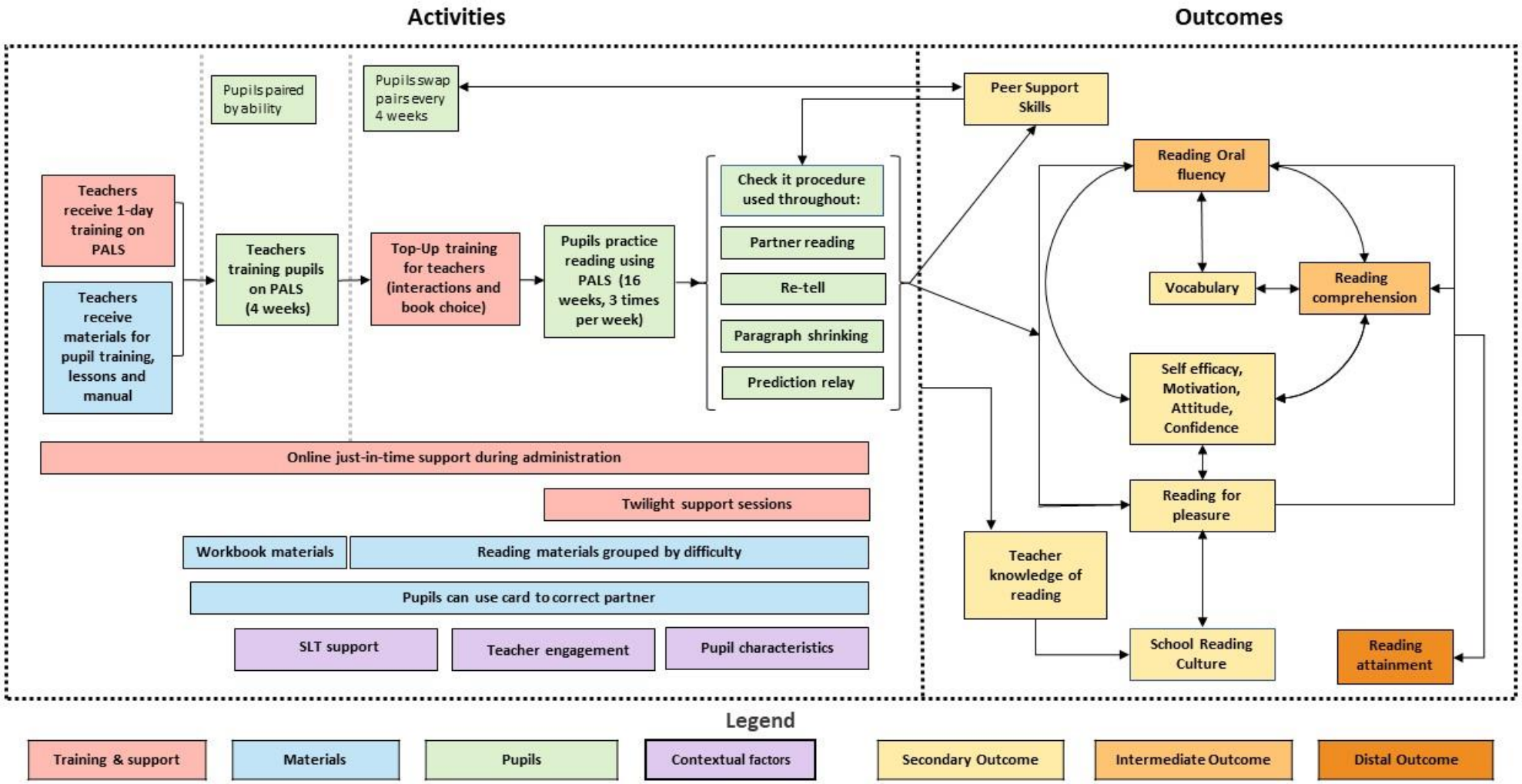
This broader school reading culture is also likely to be influenced by the teachers' knowledge of reading, which is also targeted within the intervention. The training sessions are designed to improve teachers' understanding of the skills required to be a skilled reader and provide them with pedagogical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

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

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

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

Figure 1: Logic model



Evaluation objectives

Impact evaluation: Primary research question

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

Impact evaluation: Secondary research questions

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

IPE research questions

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

The study protocol is published on the EEF website.[2] The study Statistical Analysis Plan (SAP) is also published on the EEF website.[3]

Ethics and trial registration

Ethical approval was obtained through Manchester Metropolitan University. The original submission was made on 14 January 2022 through a fast-track route and approval was granted following revisions on 1 February 2022. Then the ethics committees at Nottingham Trent University and the University of Birmingham reviewed and accepted Manchester Metropolitan University evaluation of the ethics application. The consent forms and participant information sheets were subsequently amended to ensure that participants understood that only Manchester Metropolitan University had undertaken a full ethical review. This amendment was approved on 10 February 2022. FFT Education, a sub-contractor working for Manchester Metropolitan University, requested that further information be added to the Memorandum of Understanding (MoU). This amendment was approved on 25 February 2022. A further amendment was submitted outlining the content of a video to share with pupils and a teacher guidance sheet. All participating teachers were asked to at least share the video or go through the statements on the guidance sheet with their pupils. The preference was that they did both. The video and teacher guidance provided a more accessible format for presenting the information normally included in the participant information sheet to pupils. This was approved on 9 March 2022. Finally, consent forms and participant information sheets, together with assent forms for pupils, were submitted as an amendment and approved on 21 November 2022.

[2] Available at: https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/PALSUK_Protocol_v1_20221114.pdf?v=1709048871.
[3] Available at: https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/PALS-UK-Statistical-Analysis-Plan.pdf?v=1709048871.

In sum, the process for ethical approval included providing details about the project design, information about the ethical procedures that were adopted, and copies of participant information sheets and consent/withdrawal forms. We also included the MoU and privacy notices for completeness.

The school recruitment process is detailed as follows. The delivery team identified and approached schools, which met the selection criteria in the three regions, and collected initial data including the school name, address, telephone number, unique reference number (URN), and the names and contact details of Year 5 teachers in 2022/2023. Schools were asked to sign an MoU, which provided information about the project and its aims, potential benefits for participating schools, a timetable of activities, data protection issues, and responsibilities of all parties involved. In addition, they were required to sign a separate Data Sharing Acknowledgement, which outlined how personal data would be collected and shared between the delivery team, the evaluation team, and the school. Schools then issued a withdrawal notice to all parents of pupils in Year 4. Parents were given two weeks to respond to this although they had the right to withdraw their child at any time. FFT Education then collected baseline data from each school. All recruitment documents are included in Appendix D.

This trial is registered at the Open Science Framework (OSF). The entry can be viewed at: https://osf.io/f2wgj.

Data protection

All data collected for the trial was treated with the strictest confidence and processed and stored in compliance with the General Data Protection Regulation (GDPR) and Data Protection Act 2018.

- The delivery team lead, Nottingham Trent University, was an independent data controller in respect of any personal data of pupils/and or school staff, which they processed for the purposes of delivering PALS-UK. The delivery team partner, University of Birmingham, was a data processor for the delivery team lead.
- The evaluation team, Manchester Metropolitan University was an independent data controller in respect of any personal data of pupils/and or school staff, which they processed in order to conduct the impact evaluation and the IPE. Data processors working on behalf of Manchester Metropolitan University were FFT Education who collected the baseline pupil data and AlphaPlus who conducted the 1:1 assessments of fluency.
- The EEF will become the data controller once the data is submitted to the EEF Data Archive, currently managed by FFT Education (the data processor for the archive).

The EEF project is part of a wider DfE-funded programme called the 'Accelerator Fund'. The DfE and the EEF are joint data controllers who have overarching responsibility for the programme.

Manchester Metropolitan University have ensured that all personal data collected and processed by them and the delivery team for this research project have been:

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

This evaluation was assessed for data protection and ethics as part of the research ethics approval process in place at Manchester Metropolitan University.

Manchester Metropolitan University ensured that a data sharing agreement is in place as required by the GDPR and Data Protection Act. This document clearly outlined the data sharing and protection responsibilities of Manchester Metropolitan University and Nottingham Trent University.

The delivery team ensured that all participating schools signed a Data Sharing Acknowledgement outlining what data schools would share with the delivery team and the evaluation team.

Data was processed by Manchester Metropolitan University in order to ascertain the impact of the intervention on the pupil outcomes, and to make judgements about compliance and fidelity. So that the processing of personal data relating to the pupils is fair, lawful, and transparent we used a parent information sheet, parental withdrawal form, and a privacy notice agreed with the university's data protection officer for parents and teachers. Pupils could be withdrawn from data processing at any time during the study.

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

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

Any information identifying pupils was given a unique code immediately after collection and prior to analysis in order to reduce risk. Archived data will include unique pupil numbers (UPNs) and matching to the National Pupil Database (NPD) and other administrative data may take place by the data archive manager. However, data will only be released subsequently to interested parties in an anonymised format. The information collected will be used for research purposes only and no information that can identify individuals will be used for any other purpose. Any pupil personal data collected for the impact evaluation and held by Manchester Metropolitan University and Nottingham Trent University will be destroyed in accordance with the GDPR when it is no longer required, and no later than 31 July 2024. Data from interviews (headteachers, teachers, and pupils) will be archived in an open access repository and the consent forms will be retained, if participants indicated on their consent forms that they are happy for this to happen.

Project team

The evaluation team (Manchester Metropolitan University) comprised of:

- Professor Cathy Lewin, Professor of Education, Education and Social Research Institute, Manchester Metropolitan University: Extensive experience of mixed-method evaluation of school-based interventions, including educational technology and inclusive education; joint principal investigator; and responsible for design and management of the IPE.
- Professor Stephen Morris, Professor of Evaluation, Evaluation, Policy and Evaluation Research Unit, Manchester Metropolitan University: Specialises in experimental/quasi-experimental evaluation designs; joint principal investigator; and responsible for the design and management of the impact evaluation.
- Dr Steph Ainsworth, Reader in Education, Education and Social Research Institute, Manchester Metropolitan University: Expertise in primary English teaching and the assessment of reading and early language skills; responsible for selection of the assessment measures and management of the baseline and endline testing as well as supporting the IPE.
- Sandor Gellen, Research Associate, Policy and Evaluation Research Unit, Manchester Metropolitan University: Expertise in evaluating programmes using quantitative and small-n mixed methodologies; responsible for analysis of the impact evaluation data and managing the relevant data flows.
- Dr Kate Wicker, Research Associate, Education and Social Research Unit, Manchester Metropolitan University: Expertise in mixed-methods evaluation research; responsible for coordinating the case studies and supporting the IPE.

The delivery team (Nottingham Trent University and the University of Birmingham) comprised of:

- Dr Emma Vardy, Senior Lecturer in Developmental Psychology, Psychology Department, Nottingham Trent University: Expertise includes reading development, specifically the role of affective factors and intervention evaluation; principal investigator for the delivery team; responsibilities included overall management of implementation, school recruitment, providing training and support for teachers, and managing and conducting school observations.
- Dr Helen Breadmore, Associate Professor in Psychology in Education, Department of Education and Social Justice, University of Birmingham: Expertise in literacy, assessment, and programme evaluation; co-investigator for the delivery team (acting as principal investigator during May 2022 to September 2022); responsible for the development and delivery of teacher training, quality control of implementation data, and liaising on trial design.
- Dr Luisa Tarczynski-Bowles, Research Fellow, Psychology Department, Nottingham Trent University: Expertise includes reading acquisition and interventions; particularly supporting school and pupil recruitment and retention, maintaining records of implementation, conducting school observations, and liaising between schools and the evaluation team.

Methods

Trial design

This trial is a two-armed, cluster randomised controlled trial. Entire schools were recruited to the study and randomly assigned 1:1 to either intervention or control conditions. Pupils entering Year 5 in September 2022 in schools allocated to receive the intervention had the opportunity to participate in PALS-UK. Conversely, pupils entering Year 5 in September 2022 in schools assigned to control conditions were unable to participate but instead received 'business as usual' support with their reading.

In the primary analysis, baseline scores were obtained from the New PiRA Summer Term Year 4 Test before randomisation. The primary outcome was the unstandardised score achieved by pupils in the New PiRA Summer Term Year 5 Test, which was taken in the summer of 2023. Pupil assessments were conducted face to face in schools by classroom teachers. Secondary pupil outcomes were:

- reading comprehension and oral reading fluency scores from the WIAT-III UK-T assessment;
- fluency scores from the MDFS; and
- reading self-efficacy and reading motivation scores from the Feelings About Reading (FAR) questionnaire.

The primary outcome measure (New PiRA Summer Term Year 5 Test) and the secondary outcome measures of reading self-efficacy and motivation (FAR) were collected from all Year 5 pupils in intervention and control schools at endline. The secondary outcome measures from the WIAT-III UK-T and MDFS were gathered from a subset of pupils in each school at endline (both intervention and control), sampled at random prior to random assignment of schools.

The effects of the intervention on the primary outcome were estimated for four subgroups:

- ever-FSM (using the variable EVERFSM_6);[4]
- designated SEND (binary measure);
- pupils scoring in the lowest quartile on the baseline New PiRA Summer Term Year 4 Test; and
- pupils with EAL and low prior attainment (i.e. EAL pupils obtaining a score in the lower half of all scores on the baseline New PiRA Summer Term Year 4 Test).

Table 3 describes the key features of the trial.

Table 3: Trial design

Trial design, including number of arms		• Two-armed, cluster randomised controlled trial
Unit of randomisation		• School
Stratification variable (s) (if applicable)		School size (one form per year group vs two or more forms per year group) Proportion of year group that are currently FSM (split at the median sample proportion into high or low)
Primary outcome	Variable	• Reading attainment
	Measure (instrument, scale, source)	• New PiRA Summer Term Year 5 Test, 0–45, Rising Stars
Secondary outcome(s)	Variable(s)	• Oral reading fluency (rate) • Oral reading fluency (multidimensional) • Reading comprehension • Reading self-efficacy • Motivation for reading
	Measure(s) (instrument, scale, source)	• WIAT-III UK-T: Oral reading fluency subtest, 0–121, Pearson Clinical Assessments • MDFS, 4–16

[4] The NPD was not used to gather this data. Instead, the data was collected by FFT Education in electronic form and sent to Manchester Metropolitan University to create an initial record in the trial database.

		• WIAT-III UK-T: Reading comprehension subtest, 0–44, Pearson Clinical Assessments • FAR questionnaire: Self-efficacy subscale, 20–140, developed by Carroll and Fox (2017) • FAR questionnaire: Motivation for reading subscale, 10–70, developed by Vardy *et al.* (under review)
Baseline for primary outcome	Variable	• Reading attainment
	Measure (instrument, scale, source)	• New PiRA Summer Term Year 4 Test, 0–40, Rising Stars
Baseline for secondary outcome(s)	Variable	• Reading self-efficacy • Motivation for reading
	Measure (instrument, scale, source)	• New PiRA Summer Term Year 4 Test, 0–40, Rising Stars • FAR questionnaire: Self-efficacy subscale, 20–140, developed by Carroll and Fox (2017) • FAR questionnaire: Motivation for reading subscale, 10–70, developed by Vardy *et al.* (under review)

Participant selection

Schools

Eligible schools were state-funded primary schools located in the following English Regional School Commissioner regions (prior to the DfE change to Regional Departments for Education in 2022): the North; the East Midlands and the Humber; and the West Midlands. This study was funded through the DfE's Accelerator Fund,[5] which was intended to increase access to high-quality literacy and numeracy programmes, and as a condition of funding, schools had to come from these regions. These regions had been identified as being particularly impacted by the pandemic. Schools in these regions were excluded from participation if:

- schools were involved in another Accelerator Fund efficacy trial, or another literacy project funded by the EEF targeting the cohort of pupils that would enter Year 5 in the 2022/2023 academic year;[6]
- schools that participated in the previous PALS-UK trial funded by the EEF and were allocated to the intervention group;
- schools unable to provide access to technology for online assessment of reading; and
- schools that, for any remaining reason, were using or had used PALS-UK at any point in the past.

The school recruitment process proceeded as follows. The delivery team identified and approached schools that met the selection criteria in the three regions, collecting initial data, including the school's name, address, telephone number, and URN, along with the names and contact details of Year 5 teachers in 2022/2023. Schools were requested to sign an MoU providing information about the project and its objectives, potential benefits for participating schools, a timetable of activities, data protection issues, and the responsibilities of all parties involved. Additionally, schools were required to sign a separate Data Sharing Acknowledgement outlining how personal data would be collected and shared among the delivery team, the evaluation team, and the school. Once a school signed the MoU and the Data Sharing Acknowledgement, the delivery team collected basic background information from the school. This information was sent electronically to the evaluation team and FFT Education—a partner organisation responsible for the enumeration of the pupil sample and collection of basic pupil-level information—and used to generate an initial record in the trial database.

As part of the process, schools issued a withdrawal notice to all parents of pupils in Year 4, allowing parents a two-week period to respond. However, parents retained the right to withdraw their child at any time. Following this two-week period, FFT Education gathered baseline data from each school and populated a baseline pupil dataset, subsequently sharing it with the evaluators.

[5] Available at: https://educationendowmentfoundation.org.uk/projects-and-evaluation/accelerator-fund-2022-23.
[6] This included projects such as Learning Language and Loving It, The 5Rs (Recall, Routine, Revise, Repeat, and Ready) Approach to General Certificate of Secondary Education (GCSE) Maths Resits, English Mastery, Reciprocal Reading, Children's University, Teacher Choices, and Thinking Doing Talking Science.

The research team planned to recruit approximately 120 schools—60 in each of the intervention and control arms. At the point of randomisation, 114 schools were enrolled (57 schools in each arm), and for the endline assessment, the evaluation retained 103 schools (53 in the intervention group and 50 in the control group).

Control schools did not receive any training and did not deliver the PALS-UK intervention. They followed a 'business as usual' approach to the teaching of reading. The incentive provided to control schools was £500 on the completion of the post-testing. This was in contrast to the payment of £100 intervention schools needed to make to contribute to delivery costs.

Pupils

The focal cohorts comprised pupils that entered Year 5 in September 2022 in primary schools recruited to the trial. These pupils within the cohorts were identified during Year 4, before entering Year 5.

After the schools signed the MoU and the Data Sharing Acknowledgement, they issued a withdrawal notice to all parents of pupils in Year 4. All Year 4 pupils, due to enter Year 5 in September, whose parents did not ask for them to be withdrawn were included in the study. However, pupils with SEND were excluded if, based on the school's judgement, they were unable to complete the New PiRA Summer Term Year 4 Test at baseline. These pupils, along with those that joined the school after the baseline assessment, were not included in the evaluation. Pupils joining schools in the treatment group later still participated in the intervention but were not part of the evaluation.

The final analysed sample comprised 3,628 pupils (intervention n = 1,907; control n = 1,721).

Outcome measures

Primary outcome

The primary outcome was a measure of reading attainment derived from the New PiRA Summer Term Year 5 Test. The test instrument was used in its entirety.

New PiRA termly tests[7] published by the Hodder Education are standardised assessments only available for purchase by schools and Multi-Academy Trusts (MATs). The tests are developed to assist schools in benchmarking performance and tracking progress against national averages in reading. New PiRA tests, which are UK-standardised, have been included in the EEF's Attainment Measures Database[8] and demonstrate high-test reliability (Cronbach's alpha above 0.9), face validity (aligned with national curriculum guidelines), and criterion validity, with a strong relationship with national test scores (Pearson coefficient 0.64–0.79 for Key Stage 2). Administrable to an entire class, it minimised costs and school disruption. It was designed around termly expectations, offering appropriate forms at both baseline and endline for this study (i.e. Summer Term Year 4 and Summer Term Year 5).

New PiRA is structured around the following content domains for Key Stage 2:

- vocabulary: Explain the meaning of words in context (number of marks available: 8 points[9]);
- comprehension: Retrieve and record information/identify key details (12 points);
- summary: Summarise main ideas from the text (5 points);
- inference: Make inferences from the text, explaining and justifying with evidence from the text (12 points);
- prediction: Predict what might happen (2 points);
- structure: Identify/explain how content is related and contributes to meaning (2 points);
- impact: Identify/explain how meaning is enhanced through choice of words and phrases (2 points); and
- comparison: Make comparisons within the text (2 points).

As specified in the logic model, the core activities that pupils engage in during PALS-UK are predicted to support pupils in developing their fluency, vocabulary knowledge, and reading comprehension skills. Many of the content domains

[7] For more information, see: https://www.risingstars-uk.com/series/assessment/rising-stars-pira-tests.
[8] See: https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluation/eef-outcome-measures-and-databases/attainment-measures-database
[9] The marks allocated per content domain as distributed in the New PiRA Summer Term Year 5 Test.

measured within New PiRA align with these predicted outcomes. Comprehension, summary, and prediction were expected to be particularly influenced by PALS-UK given that: retelling supports comprehension (Dunst *et al.,* 2012); paragraph shrinking entails focusing on the key main ideas, which is an essential part of summarising; and prediction relay activities are designed to support pupils in predicting upcoming events from what they have read (Fuchs *et al.,* 1997). While not the primary focus of the intervention, PALS-UK was also anticipated to contribute to improvements in other content domains measured by New PiRA. For instance, gains in fluency could potentially free up cognitive resources for comprehending inferences in texts (Klauda and Guthrie, 2008), leading to advancements in vocabulary knowledge (Yildirim *et al.,* 2014).

The baseline score for the primary outcome (reading attainment) was obtained from the digital version of New PiRA (New PiRA Summer Term Year 4 Test), which was administered prior to randomisation to pupils at the end of Summer Term 2022, when pupils were in Year 4. Schools conducted online assessments overseen by Manchester Metropolitan University, with data collection completed by the end of July 2022.

Although our original plan included conducting both the baseline and endline tests online, the administration of the New PiRA Summer Term Year 4 Test encountered challenges with the online format. To alleviate peak online testing, we distributed the administration over four weeks; however, schools faced issues like access code malfunctions, connection problems, and difficulties in saving pupils' answers. While some problems were related to browsers or internet connections and schools received guidance to address these issues, not all concerns were resolved. Technical issues prompted feedback from seven schools (four intervention schools and three control schools), highlighting the novelty of online testing for pupils, navigation difficulties, character limitations in text boxes, and time constraints affecting completion. Cumulatively, these technical difficulties raised concerns about result accuracy, with three schools (two intervention schools and one control school) noting significant pupil underperformance and doubts about their school's results. In at least three other schools, technical problems resulted in numerous 'no response' markings or missing pupil responses. While not all schools reported substantive technical issues, problems appeared to have impacted the vast majority of schools. The evaluation team concluded that performance was affected similarly across schools. Consequently, the evaluation team decided to opt for paper-based administration for follow-up testing in June and July 2023, where the endline test was administered by the school and collated by a team of subcontracted test administrators overseen by AlphaPlus (for more details, please refer to the SAP).

The possible consequences of transitioning to paper-based testing were evaluated. Nonetheless, there is limited evidence available to compare reading performance between digital and paper-based testing. One study, conducted by Støle *et al.* (2020), offered weak evidence suggesting that pupils, on average, attained lower scores on the digital test in comparison to the paper version.

Tests took between 40 and 50 minutes to complete. A pupil could score a maximum of 40 marks and a minimum of zero on the New PiRA Summer Term Year 4 Test. At endline, tests were marked by the evaluation team, who were blind from which schools were within each arm (apart from occasional exceptions, e.g. if they knew the name of a school because it was one of the six case study schools). The raw score from the baseline New PiRA reading test served as a covariate in the primary and secondary analyses discussed below. Follow-up data collection occurred towards the end of Summer Term 2023. A pupil could score a maximum of 45 marks and a minimum of zero on the New PiRA Summer Term Year 5 Test, with the raw score used as the dependent variable in the primary analysis.

Secondary outcomes

Secondary outcomes included measurements in: i) oral reading fluency (rate); ii) oral reading fluency (multidimensional); iii) reading comprehension; iv) reading self-efficacy; and v) motivation for reading.

As suggested in the logic model, PALS-UK is predicted to improve oral reading fluency, by providing pupils with frequent opportunities to read aloud with support from a partner. PALS-UK is also predicted to improve pupils' reading comprehension skills both directly, by providing structured activities that involve pupils' engaging with the meaning of the text, and indirectly, mediated by gains fluency (as discussed above in relation to New PiRA and the content domains). Given that various aspects of fluency have demonstrated independent impacts on reading comprehension (Klauda and Guthrie, 2008), there was a recognised need for a detailed measure of fluency that extends beyond mere rate and accuracy. Therefore, two fluency measures were employed. Reading fluency (rate) provided us with a basic measure of fluency by assessing the average number of words correctly read per minute, based on the oral fluency subtest of the

WIAT-III UK-T (Weschler, 2018). WIAT-III UK-T is shortlisted in the EEF attainment outcomes database. It has good construct validity and reliability and has been standardised across a UK sample.[10]

In the WIAT-III UK-T oral reading fluency substest, the pupil read two passages, assigned according to their chronological age, and a score was derived by subtracting the total errors from the combined word count of the passages. This result was then divided by the time taken to read the passages, yielding a rate per second. To convert this measure into minutes, the obtained rate per second was multiplied by 60. The total weighted raw score for this age group could range between 0 and 121. Standard administration and scoring was followed including application of 'reverse rules'. For example, when a pupil performed below a particular threshold (the time taken to read the first passage for the fluency assessment was longer than the expected amount specified within the test), pupils 'reversed' to an earlier passage (e.g. if a nine-year-old took longer than the specified time limit to read the first passage for pupils aged nine, they were then asked to read the first passage specified for children aged eight). A pupil who returns to an earlier item set, may achieve a higher standard score than a pupil who remains in the grade-appropriate set. However, it is important to note that reverting to easier sets reflects more on their range of word recognition and vocabulary than comprehension.[11]

Additionally, the MDFS (Rasinski, 2004) provided a qualitative assessment of dimensions such as: expression and volume; phrasing; smoothness; and pace. For this latter outcome measure, the MDFS offered additional insights into fluency as pupils read the WIAT-III UK-T oral reading fluency subscale texts. In other words, this scale did not involve pupils reading anything additional; rather test administrators delivered the WIAT-III UK-T subtest as normal but made an additional (qualitative) assessment of the child's reading fluency using the MDFS rubric. Test administers listened back to the recordings of the pupils after they had completed the WIAT-III UK-T test in order to make this assessment. The scale provided scores from 4 to 16.

Multiple studies have found the MDFS to be a valid and reliable instrument for assessing prosodic reading (Smith and Paige, 2019; Moser *et al.,* 2014; Paige *et al.,* 2014). For example, Moser *et al.* (2014) conducted a study to assess the inter-rater reliability of the MDFS. The authors used a generalisability study, which found that the variability attributed to readers ranged from 79.2% to 83.3% for narrative texts, and 70.7% to 81.8% for informational texts. Variability attributable to raters and rating occasions for both text genres was most often 0.0%. A decision study in the design arrived at reliability coefficients ranging from 0.94 to 0.97 for narrative texts and 0.92 to 0.98 for informational texts, indicating high consistency between raters. Similar results were found in a more recent study conducted by Smith and Paige (2019). In sum, these results suggest that the MDFS provides a reliable measurement of reading fluency.

The reading comprehension subtest of the WIAT-III UK-T tool measured comprehension—a key intermediate outcome defined by the logic model—through responses to literal and inferential comprehension questions. The total weighted raw comprehension score ranged between 0 and 44. Again, the standard WIAT-III UK-T administration and scoring procedures were followed including use of 'reverse rules', for example, if children performed below the specified threshold for their age group (e.g. if a ten-year-old pupil scored less than 2 points on the first text specified for ten-year-olds, they were asked to read texts from the nine-year-old set, and so on).

Administration times for the full assessment were approximately 25–40 minutes, with around 30 minutes for the two WIAT-III UK-T subtests per pupil, which was administered face to face in schools. Due to training requirements for valid administration of WIAT-III UK-T and MDFS, oral reading fluency and comprehension measures were collected only at endline and from a subset of randomly chosen pupils to reduce costs. For each of these measures, the baseline covariate was derived from the baseline New PiRA Summer Term Year 4 Test.

The reading comprehension and MDFS measures were scored by the trained test administrators. The test administrators were not told which schools were in which trial arm, although it was difficult to fully guarantee blinding as the test administrators may have viewed evidence of PALS-UK being delivered/spoken about inadvertently while in the school setting.

Given that the related concepts of reading self-efficacy, motivation, attitude, and reading confidence are defined secondary outcomes in the PALS-UK logic model (**Figure 1**), measuring reading self-efficacy and motivation for reading was deemed appropriate as secondary outcome measures. The final secondary outcome measures, reading self-

[10] Available at: https://educationendowmentfoundation.org.uk/measures-database/wechsler-individual-achievement-test-third-uk-edition-for-teachers.
[11] Available at: https://www.pearsonclinical.ca/content/dam/school/global/clinical/ca/assets/wiat-iii-cdn/wiat-iii-cdn-score-report-cdn-norms-can.pdf.

efficacy and motivation for reading, were assessed using an adapted version of the FAR questionnaire. Administered to the entire class, the questionnaire took approximately ten minutes to complete. The first part, gauging reading self-efficacy, comprised 20 items, while the second part, assessing motivation for reading, included ten items. Although the previous PALS-UK trial analysed only reading self-efficacy data (Culora, *et al.,* 2022), subsequent analyses indicated that both motivation for reading and reading self-efficacy contribute to unique variance in reading attainment (Vardy *et al.,* under review). Therefore, both scales were judged to be appropriate measures of the intermediate outcome measures specified in the logic model and were analysed in the current trial. Vardy *et al.* (under review) found the original scales to be valid reliable measures of the target constructs. The questionnaire used a Likert scale structure (7 points), and the potential score ranged from 30 to 210. While the motivation for reading scale initially employed a 4-point scale in a previous trial, data analysis suggested that better sensitivity could be achieved by using a 7-point scale for both parts of the questionnaire (ibid). In this study, the instrument was administered online at baseline and in paper format at endline to all pupils at the same time as the New PiRA Summer Term Year 5 Test. Similar to the New PiRA, at baseline schools conducted online assessments overseen by Manchester Metropolitan University, while the paper-based endline tests were administered by schools and collated by a team of subcontracted test administrators overseen by AlphaPlus. At baseline, the FAR scores were collected automatically by the online survey tool. At endline, the scores were transferred from the paper questionnaires to a spreadsheet by the evaluation team, who were blind to which arm each school was in (with occasional exceptions as described for the New PiRA).

The New PiRA and WIAT-III UK-T measures are not included in the report as they are commercial measures. The FAR questionnaire measure is also not included because the authors of the measure are in the process of preparing this for publication.

Sample size

The intervention delivery team had the capacity to recruit approximately 120 schools. This target number served as the basis for calculating the minimum detectable effect size (MDES) at the protocol stage. During the randomisation stage, based on the 114 schools that were successfully recruited, and accounting for a 10% school-level attrition rate, MDES calculations were based on a total sample of 102 schools.

Assumptions for Type I and II error rates, pre-/post-test correlations, and the assumption of two-sided tests for statistical significance are maintained in both sets of calculations. Assumptions made were:

- Type I and II error rates set at 5% and 20%, respectively, and aligned with standard practice in the EEF trials, representing acceptable long-run rates of error associated with hypotheses of interest.
- Randomisation to intervention and control on a 1:1 basis.
- Estimates of the correlation between pre-, and post-test New PiRA raw score for reading attainment were assumed to be 0.49 (obtained from Key Stage 1–Key Stage 2 correlation in the NPD, as reported by Allen *et al.,* 2018).
- A three-level clustered design (pupils nested within classes, classes nested within schools) was used, assuming ICC 0.10 at the school level and 0.05% at the class level. The low class-level ICC of 0.05 aligns with previous research and the widespread practice of estimating class-level ICC as half of what is found at the school level within primary education.
- An average expectation of around 25 Year 5 pupils per class and 1.55 classes per school across the entire sample was assumed (based on data released by the National Statistics containing information on school and pupil numbers in English primary schools in 2020/2021; DfE, 2021).

Average cluster sizes were converted into harmonic means to accommodate the variability in cluster sizes, following the recommendation of Dong and Maynard (2013). Using the specified parameters and assuming a continuous, normally distributed outcome, at protocol stage the MDES was calculated as 0.203, and at randomisation as 0.218. The parameters observed at the analysis stage closely resembled those assumed at the protocol stage, and the actual school-level attrition closely matched the assumption made, resulting in an MDES of 0.212.

A subgroup analysis was proposed for the primary outcome among pupils eligible for FSM at any point in the last six years (using the variable EVERFSM_6). At protocol stage we anticipated at least ten FSM-eligible pupils in each school (six pupils in each class) at analysis. Using the same assumptions, the MDES at the protocol stage for the FSM subgroup was calculated to be 0.234. At randomisation, adjusting for school recruitment figures and assuming further school-level attrition, the MDES was calculated as 0.243. Once again, the assumptions made, proved reliable as the MDES was

0.248 based on the 'as analysed' sample size. Therefore, even though being classified as an efficacy trial, the study should be powered to identify meaningful statistically significant differences between groups.

MDES was computed using the PowerUp software (Dong and Maynard, 2013). In the 'Impact evaluation' results section, **Table 9** offers a summary of statistical power during the protocol and randomisation stages. It compares assumptions and MDESs to the MDES given the sample 'as analysed'.

Randomisation

By the end of Summer Term 2022, evaluators were informed of the identities of each participating school and the pupils within them. Baseline reading assessments were collected, reading self-efficacy, and motivation for reading measures for each pupil. Before randomisation at the beginning of Autumn Term 2022, we randomly sampled one class per school. Within the selected class, ten pupils were chosen at random for additional data collection. These pupils were selected to complete the WIAT-III UK-T reading comprehension and oral reading fluency subscales in Summer Term 2023 at endline. To maximise response rates, two further selection criteria were applied:

- only pupils with a valid baseline New PiRA score have been included in the pre-selection process; and
- five further pupils have also been selected at random in each school within the same process. These pupils were only approached to complete testing if pre-selected pupils were not present in school on the day tests are administered.

Randomisation occurred only after these tasks were completed.

The random assignment of schools to treatment and control conditions was executed by Manchester Metropolitan University researchers on 2 September 2022. This process involved 114 schools that had signed an MoU. Schools were informed of their allocations on the same day, by the delivery team. As outlined in the protocol, the unit of randomisation in this trial was the school. Randomisation was stratified to achieve balance on key school-level covariates, including school size (single-form entry vs multi-form entry) and the proportion of the year group currently eligible for FSM.

Using statistical software STATA 18 (StataCorp LLC, College Station, TX, USA), randomisation was conducted by a research associate from the Policy Evaluation and Research Unit who was blinded to the identity of the school at randomisation. To address any uneven treatment fractions, specific codes were incorporated into the randomisation syntax (see the full randomisation code in the Appendix G).

Statistical analysis

The analysis proceeded based on the principle of intention-to-treat (ITT). In this approach, pupils were identified in the analysis as members of the intervention or control group based on their school's allocation to intervention and control conditions at randomisation, regardless of whether the school or pupil subsequently participated in the intervention or not. If schools withdrew from the study after randomisation and requested data deletion, records for the relevant pupils were removed from the sample file. Approaches to assess the consequences of sample loss and strategies for handling missing data are discussed below.

Primary analysis

The primary analysis aimed to estimate the average effect of ITT (AITT) of the intervention on scores obtained from the New PiRA Summer Term Year 5 Test for Year 5 pupils. The domains for this measure included vocabulary, comprehension, summary, inference, prediction, and structure. A sample estimate of AITT was obtained from a multilevel linear regression model with the following three-level form:

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

Here, Y_{ijk} is the unstandardised score obtained by pupil 'i' in class 'j' and school 'k' from the New PiRA Summer Term Year 5 Test. The variable T_k took the value one if the pupil is in a school randomised to the intervention, zero otherwise. The sample estimate of the parameter of β_1 is the estimate of AITT. X_{ijk} represents the raw score obtained by pupil 'i' in the New PiRA Summer Term Year 4 Test. Z_{ijk} is a pupil's month of birth obtained from the baseline demographic data for pupil 'i' in class 'j' and school 'k', and S_k is a collection of school-level stratum variables. w_k is a school-level random effect, u_{jk} is a class-level random effect and ε_{ijk} a pupil-level residual term.

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

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

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

Uncertainty for the treatment effects was reported in the form of continuous p-values and frequentist 95% confidence intervals (CIs). Regression estimates for treatment effects were converted to effect sizes consistent with Hedges' g, as discussed in the effect size calculation section.

Parameter estimates were obtained using statistical software STATA 18.

Secondary analysis

A range of secondary analyses were conducted. The effects on reading self-efficacy and motivation for reading outcomes for the full sample were estimated. Additionally, secondary analyses involved estimating effects on the MDFS outcome and the WIAT-III UK-T outcomes for the subset of ten pupils per school selected at random, prior to randomisation. For ease of administration the ten pupils were selected from the same class. Consequently, in multi-form entry schools, a single class was randomly chosen before sampling the ten pupils for WIAT-III UK-T and MDFS administration. Thus, the effect estimates for WIAT-III UK-T and MDFS are obtained from a two-level model instead of a three-level model, given the samples are drawn from a single class per school.

Sample estimates of AITT for reading self-efficacy, motivation for reading, oral fluency, and comprehension were obtained by fitting mixed effects linear regression models to the relevant data, in a manner similar to that of the primary analysis model. As WIAT-III UK-T and the MDFS were not administered at baseline, the main covariate included in the analysis was derived from the baseline New PiRA Summer Term Year 4 Test.

For all primary and secondary analyses, unadjusted means, adjusted means, CIs, and effect sizes are reported.

The models are described and summarised in **Table 4** below.

Table 4: Secondary analysis model description

Dependent variable	Model	Outcome measure	Baseline covariate	Sample
Oral reading fluency (rate)	Hierarchical linear model	WIAT-III UK-T	Reading attainment (from the New PiRA Summer Term Year 4 Test)	Ten pupils / school
Oral reading fluency (multidimensional)	Hierarchical linear model	MDFS	Reading attainment (from the New PiRA Summer Term Year 4 Test)	Ten pupils / school
Reading comprehension	Hierarchical linear model	WIAT-III UK-T	Reading attainment (from the New PiRA Summer Term Year 4 Test)	Ten pupils / school
Reading self-efficacy	Hierarchical linear model	FAR questionnaire	Reading self-efficacy (from the FAR questionnaire)	All pupils
Motivation for reading	Hierarchical linear model	FAR questionnaire	Motivation for reading (from the FAR questionnaire)	All pupils

Analysis in the presence of non-compliance

PALS-UK is conceived as a whole-class/school intervention, with compliance defined at the school level. If a school satisfied all compliance criteria, all pupils within that school were considered compliers and exposed either directly or indirectly to the intervention. For the purposes of compliance average causal effect (CACE) analysis, a complying school was defined as one where:

- at least one teacher from an intervention school attended the initial training event;[12] and
- there was evidence from that teacher's school that one or more pupils completed all four weeks of pupil training.

Attendance at the training events was recorded by the delivery team. Pupil training compliance data could be derived from class teacher weekly logs, research assistant observations, or surveys completed during top-up training, confirming completion of training by that point. Schools were deemed compliant if they fulfilled both compliance criteria.[13] Conversely, schools that only fulfilled one or none of the criteria were deemed non-compliant.

It was proposed that in the event of significant non-compliance (around 10% or more[14]), CACE analysis would be conducted using instrumental variables (IVs) based on two stage least squares (2SLS) and cluster robust standard errors. The purpose of this analysis was to estimate the impact of PALS-UK for pupils who complied by attending a compliant school, as distinct from the main analysis, which follows the principle of ITT. As the overall proportion of non-compliance was considered negligible, we provide a descriptive report on non-compliance at the school level and refrain from further analysis.

Missing data analysis

As missingness occurring before randomisation was unlikely to introduce bias in estimated treatment effects and turned out to be limited, sensitivity tests focused on whether missing data at follow-up resulted in bias or had implications for precision.

In the initial screening stage, the type of missingness was examined—whether it was missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR). This involved calculating and comparing the rate of missing data in the trial arms. If the level of missingness was deemed problematic (exceeding 5% in both control and treatment groups), an assessment of whether available baseline covariates explained missingness was made. This involved fitting a multilevel logistic regression model, where the dependent variable recorded whether a pupil in the full as randomised sample provided an observation on the primary outcome at endline. The model incorporated baseline measurements of gender, FSM status, SEND status, EAL status, and school size as explanatory variables. Additionally, further bivariate tests were conducted to assess the correlation between the presence of an endline New PiRA Summer Year 5 Test score and other explanatory variables: gender; FSM status; SEND status; EAL status; school size; baseline self-efficacy scores; baseline motivation scores; endline self-efficacy scores; and endline motivation scores. Covariates associated with the New PiRA Summer Term Year 5 Test scores or significantly related to missingness (with a 95% CI) were used as auxiliary variables and incorporated into a multiple imputation (MI) model.

As missing data on the New PiRA Summer Term Year 5 Test at endline exceeded 5%, and evidence from the drop-out model indicated missingness associated with included covariates, further sensitivity tests examined the consequences of missing data for the primary analysis. Multiple imputation using chained equation (MICE) was employed to impute missing values for each affected variable in the analysis. The MI process involved selecting cases with a baseline record on the New PiRA Summer Term Year 4 Test, specifying an imputation model for each variable with missing data, determining the creation of ten imputed data sets, setting a random number seed, and running mixed effects linear regression models on the imputed data sets. Rubin's rule was applied to combine the results. To perform MI and subsequent analyses, the 'mi impute chained' and 'mi estimate' commands in STATA 18 were used. The results of the combined analysis based on imputation were compared to those obtained in the primary analysis to assess the sensitivity of the results to missing data under the assumption of MAR.

The statistical software used in the analyses presented in this report does not allow for a multilevel data structure when performing imputation. For this reason, and to maintain consistency, we re-analysed the primary analysis, containing exactly the same covariates, but obtaining parameter estimates using linear ordinary least squares (OLS) regression, with cluster robust standard errors (cluster heteroskedasticity-consistent 2 [HC2] standard errors).[15] We then re-ran the

[12] N.B. It is not possible for teachers to attend the top-up training event unless they have first attended the initial training. It is also not possible for pupils to receive training in PALS-UK unless their teacher or at least one teacher in the school has attended initial training.
[13] Compliance criteria and definitions were determined following extensive discussions with both the delivery team and the EEF and are on balance, felt most appropriate.
[14] This cut-off was based on judgement following discussions with the delivery team about the level of compliance required for non-compliance to have substantive consequences.
[15] N.B. This is equivalent to a population average treatment effect as opposed to a cluster-specific effect obtained from the multilevel model with random effects at the class and school levels.

MI as before, but this time analysed the imputed data sets also using linear OLS regression with cluster robust standard errors. We then compared results to examine whether maintaining a single-level approach throughout (though correcting standard errors for clustering) led to any substantive differences in our results and hence our interpretation.

Subgroup analyses

Subgroup analysis examined AITT of PALS-UK on reading attainment scores for specific pupil categories: FSM (using the variable EVERFSM_6), designated SEND, and pupils scoring in the lowest quartile on the baseline New PiRA Summer Term Year 4 Test. It is important to note that, before randomisation, a small proportion of pupils with SEND in the sample were withdrawn (n = 57) because their teachers deemed New PiRA administration inappropriate for these pupils. Therefore, the SEND subgroup analysis did not include all initially recruited pupils with SEND.

In addition, exploratory analysis investigated the effects of PALS-UK for EAL pupils. This involved creating an indicator combining the NPD-type binary measure of EAL with a pupil's raw score on the baseline New PiRA Summer Term Year 4 Test; specifically grouping EAL pupils by whether their scores fell above or below the median for the sample. Essentially, a binary indicator was created at the pupil level, taking the value '1' if a pupil was EAL and their baseline reading test score fell below the median, '0' otherwise.

Separate analyses were conducted for all subgroups. This involved including an interaction term in the primary analysis model with the relevant subgroup indicator interacted with the treatment dummy indicator. Based on the primary analysis model, the regression model took the form:

$$Y_{kji} = \beta_0 + \beta_1 T_k + \beta_2 X_{kji} + \beta_3 Z_{kji} + \beta_4 S_k + \beta_5 subgroup_{kji} + \beta_6 T_k * subgroup_{kji} + w_k + u_{kj} + e_{kji}$$

When notable interaction effects were observed, causal impact estimates of PALS-UK on reading attainment for specified subgroups were expressed as effect sizes, consistent with Hedges' g. Effect sizes were calculated by taking the sample estimate of β_1 from a regression model that exclusively considered the subgroup, and then dividing it by the pooled unconditional standard deviation (SD) of the outcome for that group.

Additional analyses and robustness checks

For the primary outcomes, two additional analyses were conducted. The first form of sensitivity analysis involved a reduced primary analysis model, which excluded the pupil baseline measure of reading attainment:

$$Y_{kji} = \beta_0 + \beta_1 T_k + \beta_2 Z_{kji} + \beta_3 S_k + w_k + u_{kj} + e_{kji}$$

This specification allowed an assessment of how the inclusion of the baseline New PiRA Summer Term Year 4 Test score as a covariate influenced the precision of the sample estimates. The second form of sensitivity analysis mirrored the regression model used for the primary analysis, but with the age-standardised New PiRA Summer Term Year 5 Test score obtained at follow-up as the dependent variable instead of the raw score. This specification omitted the month of birth covariate previously included. This second specification enabled an assessment of how age-standardisation might influence results, given that the previous trial used age-standardised New PiRA scores in the primary analysis as an outcome measure. Floor and ceiling effects for the age-standardised New PiRA Summer Term Year 5 Test score were assessed using histograms.

Estimation of effect sizes

The estimate of the causal impact of PALS-UK on the primary outcome was expressed as an effect size, consistent with Hedges' g. The equation for Hedges' g is as follows:

$$g = \frac{\hat{\beta}}{\sigma} \times \left(\frac{N-3}{N-2.25}\right) \times \sqrt{\frac{N-2}{N}}$$

In our application $\hat{\beta}$ is the sample estimate of β_1 from the primary analysis model, and σ is the unconditional pooled SD for the dependent variable in the primary analysis model calculated across different levels in the data. The two factors to the right in the equation above adjust for bias in small samples. Due to the size of the available sample, these factors were considered trivial and were therefore, ignored. It is important to note that population SDs were not used in calculating the effect size. This decision is justified in that our sample was not a random selection of schools from the population, and inferences were drawn specifically from the sample rather than the entire population of schools. This approach was deemed acceptable as the trial functioned as an efficacy study, aiming to test the intervention under ideal circumstances.

For both primary and secondary analyses, a 95% CI for the effect size was obtained using bootstrap procedures conducted over 5,000 cycles. This approach facilitated the creation of a CI relying on an empirical distribution derived from observed data through resampling, eliminating the need for parametric assumptions. In contrast, for all other analyses, including subgroup analysis and sensitivity tests, the 95% CI for the effect size was computed by dividing the 95% CI limits for the adjusted mean difference obtained directly from the regression results by the same denominator used in the calculation of the effect size itself.

Estimation of ICC

The ICCs used in the previously mentioned power calculations considered ICCs at both the school and class levels. This three-level model (i.e. pupils nested in classes nested in schools) aligns with prior trials by the EEF (Boylan *et al.,* 2018; Jay *et al.,* 2017), where the assumed ICCs were conservative estimates of between-school and between-class variances. Unlike some other trials (e.g. Gorard *et al.,* 2017; Humphrey *et al.,* 2020; O'Hare *et al.,* 2019; Rudd *et al.,* 2017) funded by the EEF that typically neglect clustering at the class level, our approach considers it. This decision was influenced by the EEF research, which argues that disregarding class-level clustering could adversely affect trial sensitivity and statistical power (Demack, 2019). We report both the conditional and unconditional ICCs at school level and class level at the analysis stage based on the primary outcome measure.

IPE[16]

The IPE was designed to gather additional information in relation to the programme delivery and also to gather evidence on the 'business as usual' offered by schools allocated to the control arm of the study. As PALS-UK is a manualised intervention, fidelity is a key implementation dimension for this evaluation. The IPE focused on the training and support offered by the delivery team and the documentation provided as part of the PALS-UK programme. Gathering perceptions from all stakeholders (headteachers, teachers, and pupils) and observing PALS-UK lessons were also essential in order to explore fidelity. Stakeholder perceptions enabled more in-depth analyses of quality, dosage, responsiveness, and identification of teacher adaptations. The IPE complemented the impact evaluation in order to test the logic model, particularly in relation to contextual factors (senior leadership support, teacher engagement, and pupil characteristics), teacher knowledge of reading, pupils' reading for pleasure, pupils' attitudes and confidence, and school reading culture. This enabled the logic model to be fully explored through linking together the findings from the impact evaluation and IPE in relation to: training, support, and provision of materials; teacher and pupil engagement; contextual factors; and outcomes and impact. Thus, the IPE plays a key role in interpreting and explaining how PALS-UK impacts on pupils, teachers, and schools.

Given the programme was delivered in one academic year, a mixed-methods approach was adopted following a convergent parallel design. That is quantitative and qualitative IPE data were collected in parallel (Humprey *et al.,* 2016). Case studies were also incorporated in order to generate richer insights into the implementation of PALS-UK, drawing heavily on stakeholder perceptions.

Research methods

The primary forms of data collection were surveys, interviews, observations, and gathering relevant documents and resources. Surveys included a mixture of closed and open questions.

As stated above, the research questions for the IPE were:

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

[16] See **IPE guidance** for further details.

Research question 1 was addressed through observations of the training offered by the delivery team, surveys of headteachers and teachers in the intervention arm of the trial, and case studies of six schools participating in the intervention. This facilitated data collection from all participants (via surveys) alongside more in-depth probing of participants' experiences (via interviews). Observations were important, supporting triangulation. Research question 2 was addressed through surveys of headteachers and teachers in the control arm of the trial, enabling information to be gathered from all participating schools on their existing practices at the start of the intervention and any changes to practice during the academic year. **Table 5** provides an overview of the planned data collection, analyses, and relationship to the research questions, implementation dimensions, and logic model.

Table 5: IPE methods overview

Research methods	Data collection methods	Participants/ data sources (planned sample)	Data analysis methods	Research questions addressed	Implementation / logic model relevance
Mixed methods	Survey: pre-intervention	Headteachers (all, n = 120)	Descriptive cross-tabulations, mixed coding, thematic analysis	1.4 2	Usual practice
Mixed methods	Survey: pre-intervention	Teachers (all, n = >120)	Descriptive cross-tabulations, mixed coding, thematic analysis	1.4 2	Usual practice
Qualitative	Observation field notes of: initial training (two sessions); Top-up training (two sessions); Twilight training (one session)	Delivery team (two), Teachers: initial training (all attending session); top-up training (all attending session); twilight training (all attending session)	Mixed coding, thematic analysis	1.1 1.3	Fidelity Compliance Quality Responsiveness Logic model
Qualitative	Delivery documentation	Documents provided to schools relating to PALS-UK	Mixed coding, thematic analysis	1.2	Quality Logic model
Qualitative	Observation field notes of PALS-UK session	Case study class / field notes (at all six case studies on two occasions, n = 12)	Mixed coding, thematic analysis	1.1 1.2	Fidelity Quality Responsiveness Logic model
Quantitative	Observation checklists of four x PALS-UK sessions (two conducted by a researcher, two conducted by a peer review teacher)	Teachers (intervention only, all, n = >60)	Descriptive cross-tabulations	1.1	Fidelity Dosage Quality Compliance Adaptations
Quantitative	Teacher logs of PALS-UK delivery	Schools (intervention only, all, n = >60)	Descriptive cross-tabulations	1.1	Fidelity Dosage
Qualitative	Interviews with peer observers in case study schools (second visit only)	Teachers (intervention only, n = <6)	Mixed coding, thematic analysis	1.1	Fidelity Quality Dosage Responsiveness Adaptations
Qualitative	Interviews with teachers and headteachers (two visits)	Headteachers/senior leaders (n = 6 x 2) Teachers (n = 6 x 2)	Mixed coding, thematic analysis	1.1 1.2 1.3 1.4 1.5	Fidelity Quality Dosage Responsiveness Adaptations Logic model
Qualitative	Interviews with a group of pupils (two visits)	Case study school pupil sample (from six schools, six to eight pupils on two occasions, n = >36)	Mixed coding Thematic analysis	1.1 1.2 1.3	Fidelity Logic model

				1.4 1.5	
Mixed methods	Survey: post-intervention	Headteachers (intervention only, n = 60)	Descriptive cross-tabulations, mixed coding, thematic analysis	1.2 1.4 1.5	Cost Responsiveness Logic model
Mixed methods	Survey: post-intervention	Headteachers (control only, n = 60)	Descriptive cross-tabulations, mixed coding, thematic analysis	2	Usual practice
Mixed methods	Survey: post-intervention	Teachers (intervention only, n = >60)	Descriptive cross-tabulations, mixed coding, thematic analysis	1.1 1.2 1.3 1.4 1.5 2	Fidelity Dosage Quality Responsiveness Adaptations Logic model
Mixed methods	Survey: post-intervention	Teachers (control only, n = >60)	Descriptive cross-tabulations, mixed coding, thematic analysis	2	Usual practice

Survey and interview questions took account of questions asked in the first trial to enable comparisons to be made if deemed appropriate. Questions were developed in consultation with the delivery team, and also consulted with in-house literacy specialists at Manchester Metropolitan University.

Prior to the start of the intervention, all headteachers were asked to complete an online survey. It was administered in September 2022. As the survey was administered soon after randomisation, questions were phrased sensitively given that 50% of the schools had been allocated to the control arm of the trial.

All teachers identified as teaching Year 5 classes (or classes including Year 5 pupils) in September 2022 were asked to complete a survey in September 2022, soon after randomisation. Teachers acting as a peer observer from single-form entry schools were also asked to complete the survey. After schools had withdrawn from the trial they were not sent further survey links; this includes those that remained in the evaluation, in order to encourage their completion of the primary and secondary outcome measures.

Training and support are key to ensuring the fidelity of the implementation. We observed two of the initial training events (October 2022) and two of the top-up training events (November 2022–December 2022). Participants elected to attend whichever session suited them best so the numbers attending these events varied. The key source of data was structured field notes. The documentation (e.g. manual) and support materials (e.g. videos) provided to teachers were reviewed. Attendance data from the training, top-up training, and other support activities, collected by the delivery team, were also considered.

Fidelity was evaluated in two ways. First, as part of the delivery of PALS-UK, four structured observations were undertaken of all teachers in intervention schools. This data collection was overseen by the delivery team. A research assistant undertook two separate observations in all schools delivering the intervention, the first during the first four weeks (one to four, pupil training) and the second during Weeks 10-14. Each teacher was also trained as a peer observer. In single-form entry schools, another year group teacher was invited to the initial training session to act as a peer observer for the Year 5 teacher. Peer observers were expected to undertake two further structured observations between Weeks 5-9 and Weeks 15-20.

Six intervention schools, including at least one from each region, were recruited as case study schools. To keep the selection process as simple as possible and given the tight timeline, all intervention schools were invited to express their interest in participating in this additional layer of data collection. It was made clear to all intervention schools what was involved and the potential advantages of participating. Six schools expressed an interest, and all were included on the basis that they provided a diverse sample of settings in terms of school size, urban/rural location, and headteacher commitment (from the baseline survey). Recruiting more engaged schools was intended to generate the richest data and provide a deeper understanding of the participants' experiences and impact of the implementation.

One half-day visit to each case study school was undertaken, midway through the intervention (January 2023–February 2023). During each visit we planned to undertake the following activities:

- observation of PALS-UK (35 minutes);
- 20-minute teacher interview;
- 20-minute headteacher or senior leader interview;
- 20-minute pupil focus group (six to eight pupils, a different group each time); and
- 20-minute peer observer interview (single-form entry schools [n = 4], second visit only).

For schools with more than one class using PALS-UK, the teacher was randomly selected, and the group of pupils were from the selected teacher's class. The teacher was asked to identify a diverse group of pupils in relation to gender and ability although this relied on parents' consent as well as pupils' assent to participate in the focus group.

A second half-day visit was originally planned but was replaced with online interviews with senior leaders, teachers, and peer observers (in single-form entry schools). In addition, teachers were asked to obtain feedback from their pupils prior to the interview and were provided with some questions as prompts. The online interviews took place towards the end of the delivery of PALS-UK between May 2023 and July 2023.

Post-intervention surveys were administered to all headteachers and teachers in June 2023 and July 2023.

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

'Business as usual' was verified through the headteacher and teacher surveys administered at baseline to all schools and administered at post-test to control schools. Post-test surveys asked about current practices, which could then be compared with baseline responses, and also about any specific changes that took place during the academic year.

Fidelity was assessed through data collected about training and support (e.g. attendance logs, and uptake of optional support) and from the PALS-UK observation checklist. Intervention teachers were asked about any adaptations made to the programme in the post-test survey. Case study interviews also elicited data in relation to fidelity. Dosage is a key element relating to fidelity. The delivery team collected data from teachers on the number of sessions delivered via a logbook completed weekly on Microsoft Teams.

The support of senior leaders was considered to be an important success factor in the previous efficacy trial. Therefore, responsiveness was another key implementation dimension, and this was evaluated through surveys and the case studies.

Issues of bias were addressed through the following means. Data were collected from multiple sources (e.g. surveys, interviews, and observations) and from different stakeholders. Structured observations, structured interview questions, and survey questions ensured that data collection was rigorous. Surveys were administered online and reminders were sent at least twice to minimise potential bias from non-responders. Qualitative data were coded and analysed thematically, ensuring that a consistent approach is adopted by all those involved. Pupils invited to participate in focus groups were selected by their teacher who was advised to identify pupils who were representative in terms of criteria such as gender and ability.

Summary of IPE data collected by method

Interviews

As detailed above, we planned to conduct interviews with staff and pupils at six case study schools across two points in time. All six case study schools were retained throughout the trial.

In total, 15 members of staff participated in an interview. This included five headteachers/senior leaders, six teachers, and four other staff. A total of 45 pupils took part in focus groups. A summary of the interviews completed in each case study school is included in **Table 6** below. It shows that in the school named 'CS5 (case study 5)', we were unable to talk to the headteacher or senior leader on either visit, at endline we were unable to talk to the headteacher or senior leader at two further schools (CS3 and CS6), in two of the single-form entry schools we were unable to talk to the peer observer (CS5 and CS6) and in one school we did not receive pupil feedback on the second visit (CS2). In all cases these omissions were due to staff availability and capacity.

Table 6: Interviews completed in case study schools

Case study (CS)	Visit 1			Visit 2			
	Headteacher / senior leader interview	Teacher interview	Pupil focus group	Headteacher / senior leader interview	Teacher interview	Peer observer interview	Pupil feedback received
CS1	Yes	Yes	Yes (n = 7)	Yes	Yes	Yes	Yes
CS2[17]	Yes	Yes	Yes (n = 6)	Yes	Yes	Yes	No
CS3[18]	Yes	Yes	Yes (n = 7)	No	Yes	N/A	Yes
CS4	Yes[19]	Yes	Yes (n = 14)	Yes	Yes	N/A	Yes
CS5	No	Yes	Yes (n = 4)	No	Yes	No	Yes
CS6	Yes	Yes	Yes (n = 7)	No	Yes	No	Yes

N/A, not applicable.

Baseline and endline surveys

At baseline, there were separate surveys for teachers and headteachers. We received 86 responses to the headteacher survey, including 41 from control schools and 45 from intervention schools. All responses were valid; the respondent had answered the questions. There were no duplicates. Two responses were received from schools that subsequently withdrew from the trial (one from a control school and one from an intervention school), and two from intervention schools that withdrew from the intervention but remained in the evaluation. The former was deleted from the analysed data set and the latter retained.

For the baseline teacher survey, we received 163 responses: 78 responses from 48 control schools; and 85 responses from 47 intervention schools. All responses were valid. There were two duplicate responses (one from a control school and one from an intervention school), where two teachers had each filled in the questionnaire twice. The first submission was included for analysis, with the addition of any extra free-text details that had been included in the second submission. Ten responses were received from control schools that withdrew completely from the trial and seven from intervention schools that withdrew from the intervention but remained in the evaluation. The former were deleted for the purposes of analysis and the latter were retained.

At endline, there were separate surveys for four groups of respondents: headteachers and teachers in control schools; and headteachers and teachers in intervention schools. We received 31 responses from control school headteachers and 26 from intervention school headteachers. There were no duplicate or invalid responses, and no responses were received from withdrawn schools.

In response to the endline teacher survey we received 50 responses from 38 control schools and 72 responses from 41 intervention schools. All responses were valid and there were no duplicates. Three responses were received from control schools that withdrew during endline testing and these were removed from the dataset for analysis.

The data in **Table 7** shows the number of survey responses obtained compared to intended, and the percentage response rate. Schools that withdrew from the trial are not included in the expected figures.

Table 7: Composition of the dataset from the staff surveys

Intended respondent	Control		Intervention	
	Baseline	Endline	Baseline	Endline
Headteacher	40/52 (79%)	31/52 (60%)	44/53 (83%)	26/50 (52%)
Teacher	67/71 (94%)	47/71 (66%)	84/86 (98%)	72/79 (91%)

Observations

We observed a sample of the training events. As planned, this included two initial training events (36 teachers attended one and five attended the other) and two top-up training sessions (with 13 participants at one and 20 at the other). We also observed one twilight session (with seven participants).

[17] A higher level teaching assistant involved in the delivery of PALS-UK was also interviewed on this visit.
[18] The school's reading lead was also interviewed on this visit.
[19] Both the headteacher and deputy headteacher were interviewed on this visit.

At these events we made field notes on the structure and content of the sessions and recorded our reflections on delivery and the engagement of teachers.

Further data collection

The delivery team provided data relevant to fidelity generated as part of their own programme monitoring. This included training attendance numbers and the results of in-school PALS-UK observations conducted by peer observers and research assistants. The evaluation team conducted further observations of PALS-UK lessons (n = 8) at the first case study school visit, with at least one observation completed in each case study school.

We approached staff in two intervention schools that had withdrawn from the trial partway through to invite them to a short interview about their experiences with the programme and reasons for leaving. One member of staff from one school gave us written comments.

Analysis

Quantitative data from surveys have been represented as graphs where appropriate. Qualitative data were analysed using software NVivo and thematic analysis (Braun and Clarke, 2006) with a mixed coding method. A coding framework derived from the logic model was applied deductively and additional themes derived inductively, allowing for unexpected mediators and outcomes to be identified. This was applied to open questions from the surveys and interview transcripts. These data will ensure that the research questions can be addressed drawing on evidence of the experiences of intervention and control schools (business as usual). Findings from case studies have enabled interpretation of the impact analyses.

Using structured approaches has ensured that the analyses are rigorous and that the findings, as far possible, are representative and unbiased.

Costs

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

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

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

- programme fees: Reflecting school access to training and materials based on market value;
- pre-requisite costs;
- staff time for teacher training, preparation, and delivery of PALS-UK, with a separate identification of the cost of new hires and supply staff; and
- any additional (unpaid) staff time supporting the delivery of PALS-UK, as reported by headteachers.

Costs were divided into pre-requisites, start-up costs, and recurring costs in accordance with the EEF's cost evaluation guidance (EEF, 2019). Programme fees calculation relied on information provided by the delivery team. Additionally, cost data was collected through post-intervention headteacher and teacher surveys participating in the IPE.

Timeline

Table 8: Timeline

Dates	Activity	Staff responsible / leading
December 2021–January 2022	Start-up meetings/review theory of change	Nottingham Trent University and the University of Birmingham / Manchester Metropolitan University
December 2021–February 2022	Data governance/MoU, parental withdrawal, and data processing notices drafted and agreed	Manchester Metropolitan University
February 2022	Ethical approval	Manchester Metropolitan University, Nottingham Trent University, and the University of Birmingham
February 2022	Finalise trial/IPE design	Manchester Metropolitan University
February 2022–June 2022	Recruitment of schools	Nottingham Trent University and the University of Birmingham
June 2022–July 2022	Baseline data collection (including enumeration of the sample and administration of New PiRA and FAR questionnaire online)	FFT Education / Manchester Metropolitan University
May 2022–November 2022	Protocol development	Manchester Metropolitan University
August 2022	Sampling for endline New PiRA administration	Manchester Metropolitan University
September 2022	Randomisation	Manchester Metropolitan University
September 2022	Inform schools of outcome	Nottingham Trent University and the University of Birmingham
September 2022–October 2022	Administration of baseline surveys	Manchester Metropolitan University
October 2022	Delivery of initial training for intervention, with sample observed	Nottingham Trent University with the University of Birmingham (delivery) and Manchester Metropolitan University (observations)
November 2022	SAP submitted	Manchester Metropolitan University
November 2022–December 2022	Delivery of top-up training for intervention, with sample observed	Nottingham Trent University with the University of Birmingham (delivery) and Manchester Metropolitan University (observations)
November 2022–May 2023	PALS-UK delivery	Nottingham Trent University and the University of Birmingham
January 2023–February 2023	First case study school visits	Manchester Metropolitan University
May 2023–July 2023	Administration of endline New PiRA and FAR questionnaire	Manchester Metropolitan University
May 2023–July 2023	Administration of WIAT-III UK-T tests and MDFS assessments	AlphaPlus
May 2023–July 2023	Second case study school visits (online)	Manchester Metropolitan University
June 2023–July 2023	Administration of endline surveys	Manchester Metropolitan University
December 2023–February 2024	Analysis	Manchester Metropolitan University
29 February 2024	Draft report submitted	Manchester Metropolitan University

Impact evaluation

Participant flow including losses and exclusions

School recruitment and attrition

The delivery team approached 5,325 schools and asked them to participate in the study. Of these, 124 schools signed the MoU, but ten schools withdrew after signing the MoU, but before the baseline assessments took place primarily citing staffing issues as the reason for their withdrawal but also logistical problems arranging for the baseline testing to take place.

A total of 114 schools signed an MoU, completed baseline data collection and were randomised: 57 to the intervention group and 57 to the control group. Following randomisation but before endline testing, eight schools withdrew from the evaluation—five from the control arm and three from the intervention arm—primarily citing staffing issues as the reason for their withdrawal. Additionally, one control school did not complete the endline testing for the primary outcome, and two schools—one from the control arm and one from the intervention arm—had missing test results (lost in transit). At the time of endline testing, 103 schools remained in the evaluation (intervention, n = 53; control, n = 50). Two of these intervention schools withdrew from the delivery of PALS-UK citing staffing and capacity issues but were retained in the evaluation. These schools did not meet any of the compliance criteria (i.e. they did not attend the initial training event) so it is assumed that they did not even partially receive the intervention.

Pupil recruitment and attrition

For the schools that had signed the MoU, the number of pupils eligible for participation in the intervention—those starting Year 5 in September 2022—was 4,927. Before the randomisation process, 87 pupils (1.8%) were excluded from the sample for the following reasons:

- 51 pupils were excluded for not meeting the pupil eligibility criteria or meeting the exclusion criteria;[20]
- 13 pupils were excluded due to parental withdrawal; and
- 23 pupils left the school before randomisation.

In total, 4,840 pupils were included in the sample at the time of randomisation.

Baseline New PiRA assessments

At baseline, assessments were conducted in all 114 schools, involving a total of 4,481 pupils (intervention, n = 2,290 ,51.1%; control, n = 2,191, 48.9%). Baseline New PiRA Summer Term Year 4 Test results were not obtained from 359 pupils (7.4%) for the following reasons:

- technical issues occurred during the online test for 29 pupils, leading to unrecorded responses; and
- another 330 pupils were absent on the day of assessment.

Schools were not informed of their allocation until after completing the baseline testing.

Endline New PiRA assessments

In total, the evaluation team received 3,845 endline New PiRA Summer Term Year 5 Tests from 103 of the 106 schools that remained in the evaluation. Endline testing was originally planned for all pupils who stayed in the evaluation, but the primary analysis was restricted to those with both baseline and endline New PiRA assessments. The final analysed sample comprised 3,628 pupils (intervention n = 1,907; control n = 1,721), with most missing data due to school-level attrition. Completed tests from 128 pupils were lost in transit before they could be marked.

Compared to the sample ‘as randomised’, the ‘as analysed’ sample was affected by the following data loss:

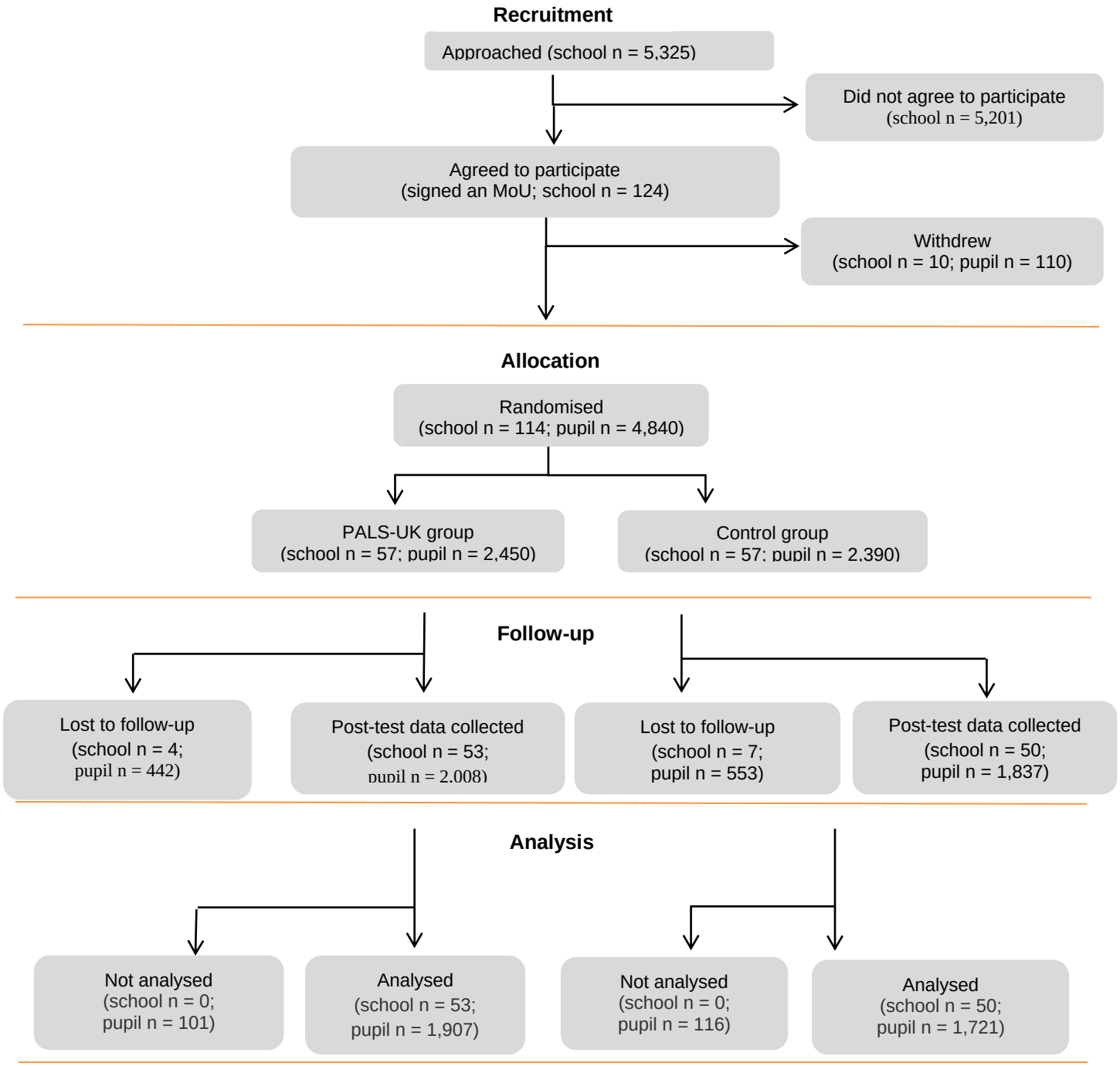
- pupil missingness due to their school leaving the evaluation: n = 412 (intervention n = 134; control n = 278);

[20] Pupils were ineligible if, as determined by the school, unable to complete the New PiRA Summer Year 4 Test at baseline.

- pupil left school: n = 88 (intervention n = 49; control n = 39);
- withdrawn by parent: n = 11 (intervention n = 5; control n = 6);
- withdrawn by school: n = 6 (intervention n = 1; control n = 5);
- refused: n = 2 (intervention n = 0; control, n = 2);
- test has been lost in transit: n = 128 (intervention n = 41; control n = 87);
- absent on assessment days: n = 295 (intervention n = 184; control n = 111); and
- missing baseline test score: n = 270 (intervention n = 129; control n = 141).

Figure 2 shows the progression of schools and pupils through the trial. The primary analysis is based on 103 schools (and 3,628 pupils) from the 114 schools (and 4,840 pupils) that initially enrolled. This represents a school response rate of 90% and a pupil response rate of 75%.

Figure 2: Participant flow diagram (two arms)



As per protocol, our goal was to achieve 80% power for detecting an effect size of 0.20 of an SD across 120 schools. This calculation assumed a baseline/endline correlation of 0.70, an ICC of 0.10 at the school level, and 0.05 at the class level. At the randomisation stage, we factored in a 10% expected attrition after recruiting the 114 schools, leading to an expectation of achieving 102 schools at analysis. Under otherwise identical assumptions, the MDES was calculated to be 0.22. In the primary analysis model, the actual observed ICC at the school level was 0.11, and at the class level was 0.04. The overall correlation between pre- and post-test scores for participants in the primary analysis was 0.73. Considering the number of schools included in the primary analysis model (n = 103) and the observed ICC and baseline/endline correlation, the estimated MDES for the primary outcome based on the sample 'as analysed' was determined to be 0.22 (see **Table 9**).

Table 9: MDES at different stages

		Protocol		Randomisation		Analysis	
		Overall	FSM	Overall	FSM	Overall	FSM
MDES[21]		0.203	0.234	0.218	0.243	0.212	0.248
Pre-/post-test correlations	Level 1 (pupil)	0.7	0.7	0.7	0.7	0.73	0.72
	Level 2 (class)	N/A	N/A	N/A	N/A	N/A	N/A
	Level 3 (school)	N/A	N/A	N/A	N/A	N/A	N/A
ICCs	Level 2 (class)	0.05	0.05	0.05	0.05	0.04	0.04
	Level 3 (school)	0.10	0.10	0.10	0.10	0.10	0.10
Alpha		0.05	0.05	0.05	0.05	0.05	0.05
Power		0.8	0.8	0.8	0.8	0.8	0.8
One-sided or two-sided?		Two	Two	Two	Two	Two	Two
Average cluster size for level 1 (per level 2 unit)[22]		20	5	19.36	5.22	20.00	5.48
Average cluster size for level 2 (per level 3 unit)		1.30	1.30	1.42	1.50	1.35	1.41
Number of schools	Intervention	60	60	51	51	53	50
	Control	60	60	51	51	50	45
	Total:	120	120	102	102	103	95
Number of pupils	Intervention	2,280	570	2,205	721	1,907	606
	Control	2,280	570	2,151	796	1,721	597
	Total:	4,560	1,140	4,356	1,517	3,628	1,203

N/A, not applicable.

Attrition

Out of the total randomised participants, 3,628 pupils, constituting 75% of the randomised cohort, possessed valid baseline and endline test scores and were consequently incorporated into the primary outcome analysis. The ratios of 'as analysed' to 'as randomised' samples for the two trial arms were 1,907:2,450 for the intervention arm and 1,721:2,390 for the control arm (see **Table 10**).

[21] Estimates at randomisation stage assume 10% school-level attrition at post-test.
[22] The harmonic mean here is reported to account for varying cluster size.

Table 10: Pupil- and school-level attrition from the trial (primary outcome)

		Intervention	Control	Total
Number of pupils	Randomised	2,450	2,390	4,840
	Analysed	1,907	1,721	3,628
Pupil attrition (from randomisation to analysis)	Number	543	669	1212
	Percentage	22.16	28.00	25.04
Number of schools	Randomised	57	57	114
	Analysed	53	50	103
School attrition (from randomisation to analysis)	Number	4	7	11
	Percentage	7.02	12.28	9.65

Pupil and school characteristics

Characteristics of the 114 randomised schools and 4,840 participating pupils are presented in **Table 11**. The data reveal a reassuring degree of similarity between the two trial arms. Schools in the intervention and control groups exhibited very similar Office for Standards in Education, Children's Services and Skills (Ofsted) rating proportions. In the intervention group, 27.3% of the pupils on the school rolls were eligible for FSM, as opposed to 28.1% in the control arm. The proportions of pupils with EAL on the school rolls were also comparable between the intervention and control arms (15.2% and 17.6%, respectively). The trial arms exhibited nearly identical prior Key Stage 2 reading average scaled scores for the school (104.3 and 104.2), and closely aligned with the national average mean (104.4).

Looking at the overall percentages of the ‘as randomised’ sample, the trial arms were also well balanced in terms of pupils with SEND (17.1% in the treatment group and 17.2% in the control group) and lowest quartile attained New PiRA Summer Term Year 4 Test scores at baseline (24.9% and 25.3%, respectively). There was a balance between males and females both in the intervention group (male n = 1,247, 50.9%; female n = 1,203, 49.1%) and in the control group (male n = 1,219, 51.0%; female n = 1,171, 49.0%).

Some differences were observed for EAL between groups in the ‘as randomised’ sample. The control group had a larger proportion of EAL pupils (n = 579, 24.2%) compared to the intervention group (n = 477, 19.5%). Additionally, a higher proportion of pupils eligible for FSM was noted in the control group, with 885 (37.0%), compared to 801 (32.7%) in the intervention group. Finally, differences were observed between trial arms in terms of EAL with a score on the baseline reading assessment falling in the lower half of the sample distribution: 246 (10.7%) in the treatment arm and 300 (13.7%) in the control arm.

In terms of continuous pupil-level characteristics, no differences were observed in terms of the average age between trial arms (127.5 months on average in both arms). Furthermore, baseline covariates used for the primary and secondary analyses, namely New PiRA Summer Term Year 4 Test raw score, FAR self-efficacy score, and FAR motivation score, were compared and converted into Hedges’ g effect size, with no notable differences between the arms observed.

Table 11: Baseline characteristics of groups as randomised

School level (categorical)		Intervention group		Control group	
		n/N (missing)	Count (%)	n/N (missing)	Count (%)
School area: Rural Urban		57/57 (0)	 6 (10.5) 51 (89.5)	57/57 (0)	 9 (15.8) 48 (84.2)
Ofsted ratings:					

Outstanding Good Requires improvement		50/57 (7)	5 (10.0) 39 (78.0) 6 (12.0)	52/57 (5)	<5 (-) 45 (86.5) <5 (-)	
Ofsted ratings: Outstanding Good Requires improvement		50/57 (7)	5 (10.0) 39 (78.0) 6 (12.0)	52/57 (5)	<5 (-) 45 (86.5) <5 (-)	
School type: Academy Maintained		57/57 (0)	24 (42.1) 33 (57.9)	57/57 (0)	29 (50.9) 28 (49.1)	
School type subgroups: Academy converter Academy sponsor-led Community Foundation Voluntary aided Voluntary controlled		57/57 (0)	20 (35.1) <5 (-) 22 (38.6) <5 (-) <5 (-) 6 (10.5)	57/57 (0)	23 (40.4) 6 (10.5) 17 (29.8) <5 (-) 6 (10.5) <5 (-)	
School level (continuous)	National level mean	n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	
Proportion of FSM pupils	24.5	57/57 (0)	27.3 (14.2)	57/57 (0)	28.1 (16.8)	
Proportion of EAL pupils	17.1	57/57 (0)	15.2 (18.4)	57/57 (0)	17.6 (25.8)	
Key Stage 2 reading average scaled score	104.4	54/57 (3)	104.3 (2.3)	56/57 (1)	104.2 (2.4)	
Pupil level (categorical)		n/N (missing)	Count (%)	n/N (missing)	Count (%)	
Gender: Male Female		2,450/2,450 (0)	1,247 (50.9) 1,203 (49.1)	2,390/2,390 (0)	1,219 (51.0) 1,171 (49.0)	
EAL pupil: Yes No		2,450/2,450 (0)	477 (19.5) 1,973 (80.8)	2,390/2,390 (0)	579 (24.2) 1,811 (85.8)	
FSM pupil: Yes No		2,450/2,450 (0)	801 (32.7) 1,649 (67.3)	2,390/2,390 (0)	885 (37.0) 1,505 (63.0)	
Designated SEND: Yes No		2,450/2,450 (0)	419 (17.1) 2,031 (82.9)	2,390/2,390 (0)	411 (17.2) 1,979 (82.8)	
PiRA lowest quartile: Yes No		2,290/2,450 (160)	571 (24.9) 1,719 (75.1)	2,191/2,390 (199)	555 (25.3) 1,636 (74.7)	
EAL and PiRA lower half: Yes No		2,290/2,450 (160)	246(10.7) 2,044 (89.3)	2,191/2,390 (199)	300 (13.7) 1,891 (86.3)	
Pupil level (continuous)		n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	Hedges' g effect size
Age (in months)		2,449/2,450 (1)	127.5 (3.6)	2,390/2,390 (0)	127.5 (3.6)	N/A

PiRA raw score		2,290/2,450 (160)	16.1 (8.2)	2,191/2,390 (199)	16.3 (8.2)	-0.01
FAR self-efficacy score		1,977/2,450 (473)	105.5 (23.8)	1,864/2,390 (526)	106.5 (23.2)	-0.05
FAR motivation score		1,977/2,450 (473)	53.4 (12.1)	1,864/2,390 (526)	53.3 (12.7)	0.01

N/A, not applicable.

Table 12 presents baseline pupil characteristics for the 'as analysed' sample. The total number of schools providing valid endline data for the primary analysis was 103 (53 in the control and 50 in the treatment arm). School-level attrition did not lead to appreciable imbalance in baseline school characteristics in the 'as analysed' sample. The proportion of schools in terms of the rural/urban split, Ofsted ratings, and school types was very similar between trial arms.

In intervention schools in the 'as analysed' sample, 27.4% of the school roll was eligible for FSM, compared to 28.3% in the control arm. The percentage of the school roll with EAL was comparable in treatment and control arms (15.9% and 17.8%, respectively). The difference in mean Key Stage 2 reading average school scaled scores was almost identical in the treatment and control arms (104.4 and 104.3, respectively), furthermore, nearly identical to the national average mean for Key Stage 2 reading scores (104.4). There were also similar proportions of pupils designated as SEND on the school rolls for both trial arms (14.6% in the treatment group and 15.8% in the control group), as well as proportions in the lowest quartile of the New PiRA Summer Term Year 4 Test scores (23.8% and 24.3%, respectively).

In terms of the 'as analysed' sample, approximately half of the pupils were male: 941 (49.3%) in the intervention group and 873 (50.7%) in the control group. The remaining pupils were female, with 966 (50.7%) in the intervention group and 848 (49.3%) in the control group. Some differences between the trial arms did emerge for proportions of EAL. The control group had a larger proportion of EAL (n = 385, 22.4%) compared to the intervention group (n = 373, 19.6%). Additionally, a higher proportion of pupils eligible for FSM was noted in the control group (n = 597, 34.7%), compared to in the intervention group (n=606, 31.8%). Finally, differences were observed between trial arms in terms of EAL pupils with a score on the baseline reading assessment falling in the lower half of the sample distribution: n = 207 (10.9%) in the treatment arm and n = 235 (13.7%) in the control arm.

In terms of continuous pupil-level characteristics, again, no notable differences were observed in terms of the average age between trial arms (127.6 in the treatment arm and 127.5 in the control arm) in the 'as analysed' sample. Furthermore, baseline covariates used for the primary and secondary analyses, namely, the New PiRA Summer Term Year 4 Test raw score, FAR self-efficacy score, and FAR motivation score, were compared and converted into Hedges' g effect size. No appreciable differences were observed between trial arms.

Table 12: Comparison of baseline pupil characteristics, as included in the primary analysis

School level (categorical)		Intervention group		Control group		
		n/N (missing)	Count (%)	n/N (missing)	Count (%)	
School area: Rural Urban		53/53 (0)	6 (11.3) 47 (88.7)	50/50 (0)	8 (16.0) 42 (84.0)	
Ofsted ratings: Outstanding Good Requires improvement		46/53 (7)	5 (10.9) 35 (76.1) 6 (13.0)	45/50 (5)	<5 (-) 40 (88.9) <5 (-)	
School type: Academy Maintained		53/53 (0)	22 (41.5) 27 (58.5)	50/50 (0)	23 (46.0) 27 (54.0)	
School type subgroups: Academy converter Academy sponsor-led		53/53 (0)	19 (35.9) <5 (-)	50/50 (0)	18 (36.0) 5 (10.0)	

Community Foundation Voluntary aided Voluntary controlled			21 (39.6) <5 (-) <5 (-) 5 (9.4)		16 (32.0) <5 (-) 6 (12.0) <5 (-)	
School level (continuous)	National level mean	n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	
Proportion of FSM pupils	24.5	53/53 (0)	27.4 (14.3)	50/50 (0)	28.3 (17.6)	
Proportion of EAL pupils	17.1	53/53 (0)	15.9 (19.1)	50/50 (0)	17.8 (25.7)	
Key Stage 2 reading average scaled score	104.4	51/53 (2)	104.4 (2.3)	49/50 (1)	104.3 (2.5)	
Pupil level (categorical)		n/N (missing)	Count (%)	n/N (missing)	Count (%)	
Gender: Male Female		1,907/1,907 (0)	941 (49.3) 966 (50.7)	1,721/1,721 (0)	873 (50.7) 848 (49.3)	
EAL pupil: Yes No		1,907/1,907 (0)	373 (19.6) 1534 (80.4)	1,721/1,721 (0)	385 (22.4) 1,336 (77.6)	
FSM pupil: Yes No		1,907/1,907 (0)	606 (31.8) 1301 (68.2)	1,721/1721 (0)	597 (34.7) 1,124 (65.3)	
Designated SEND: Yes No		1,907/1,907 (0)	279 (14.6) 1,628 (85.4)	1,721/1,721 (0)	273 (15.8) 1,448 (84.2)	
PiRA lowest quartile: Yes No		1,907/1,907 (0)	454 (23.8) 1,453 (76.2)	1,721/1,721 (0)	418 (24.3) 1,303 (75.7)	
EAL and PiRA lower half: Yes No		1,907/1,907 (0)	207 (10.9) 1,700 (89.1)	1,721/1,721 (0)	235 (13.7) 1,486 (86.7)	
Pupil level (continuous)		n/N (missing)	Mean (SD)	n/N (missing)	Mean (SD)	Hedges' g effect size (95% CI)
Age (in months)		1,907/1,907 (0)	127.6 (3.5)	1,721/1,721 (0)	127.5 (3.6)	N/A
PiRA raw score		1,907/1,907 (0)	16.3 (8.0)	1,721/1,721 (0)	16.5 (8.1)	-0.02 (-0.18, 0.14)
FAR self-efficacy score		1,661/1,907 (246)	106.4 (23.5)	1,461/1,721 (460)	107.0 (23.1)	-0.02 (-0.11, 0.07)
FAR motivation score		1,661/1,907 (246)	53.5 (12.1)	1,461/1,721 (460)	53.5 (12.6)	-0.01 (-0.11, 0.09)

N/A, not applicable.

Taken together, these findings suggest that, from randomisation to analysis, the characteristics of the sample remained relatively stable, with a slight decrease in both FSM and pupils with SEND in both arms in the 'as analysed' sample. However, since the primary attrition was due to school withdrawals, these minor changes are likely to be associated with schools that withdrew having higher proportions of FSM and pupils with SEND than average. Again, this change did not impact the overall balance between arms.

While certain subgroups are slightly overrepresented in the control arm (i.e. proportionately, there is around 3% more FSM and pupils with SEND), the equality between the two arms of the trial in the 'as analysed' sample in terms of their

average baseline New PiRA Summer Term Year 4 Test scores is reassuring. Generally, we conclude that the 'as analysed' sample upon which our primary analysis is based is well balanced by trial arm given this assessment drawing on the variables available to us.

Outcomes and analysis

Primary analysis

The primary outcome and thus the focus of attention in our primary analysis is a measure of reading attainment derived from the New PiRA Summer Term Year 5 Test. Among the 114 schools randomised, 103 provided a valid New PiRA reading attainment score at both baseline and endline for at least one pupil. Both baseline and endline New PiRA scores were available for 3,628 pupils (75.0%), with a correlation between the two measures of 0.73. Histograms examining the distributions of baseline and endline New PiRA scores by intervention and control groups indicate approximately normal distributions (see **Figure 3**).

The mean New PiRA Summer Term Year 5 Test scores at endline were 22.2 (95% CI: 21.5 to 23.0) in the intervention arm and 21.3 (95% CI: 20.2 to 22.4) in the control arm. The unadjusted mean difference was 0.97 (95% CI: -0.42 to 2.35) (see **Appendix C**, **Table D1**). The adjusted mean difference in endline scores is obtained from a multilevel linear model as discussed in the methods section above. In checking model assumptions, a QQ (Quantile-Quantile) plot confirmed the normality of residuals (refer to **Figure 4**).

Figure 3: Histograms of pre- and post-intervention unstandardised New PiRA scores, for the randomised sample (values at the extremes of the range have been suppressed to avoid statistical disclosure where there were values with fewer than ten pupils)

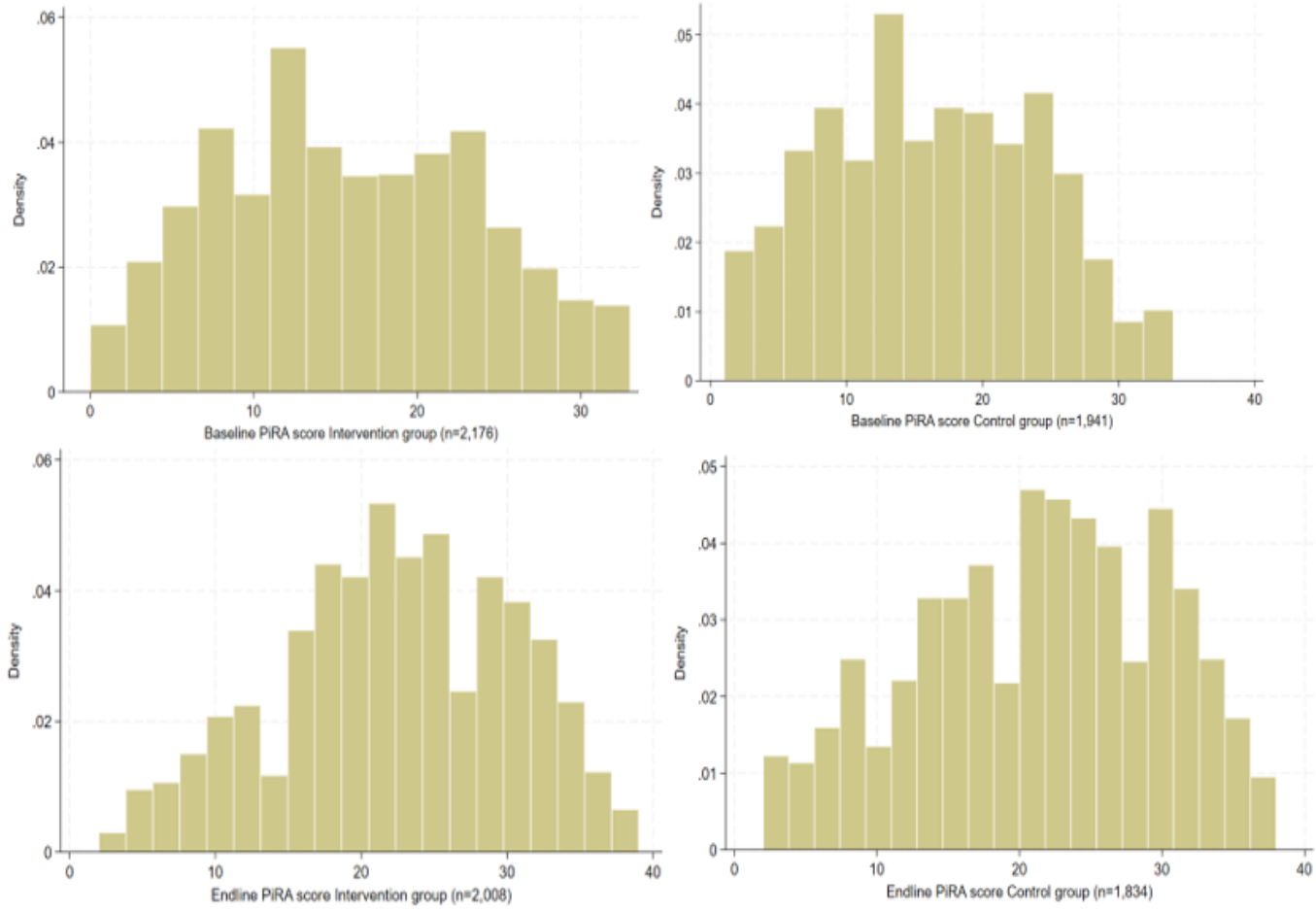
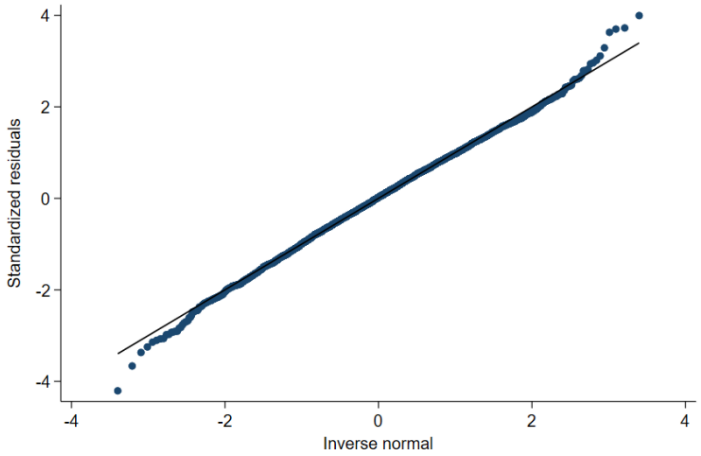


Figure 4: QQ plot of the standardised residuals, to assess normality, from the primary analysis model



Model results reveal that the adjusted mean difference in the New PiRA Summer Term Year 5 Test scores at endline between the intervention and control groups was 1.02 (95% CI: 0.12 to 1.92; see **Appendix C**, **Table D1**). In other words, intervention group pupils scored on average just over 1 point higher in their endline New PiRA assessment than the control group. This is equivalent to an effect size (Hedges' g) of 0.12 (95% CI: 0.01 to 0.23), which equates to approximately two months' additional progress in the intervention group relative to the control. We note, however, that although the CI associated with this result does not contain zero, it is quite wide and suggests that quite small values (close to zero) for the true difference between the groups cannot be rejected at the 95% confidence level. Having said this, the p-value associated with this test reveals that the probability of observing an effect size of this magnitude or larger under the null hypothesis is quite low (p = 0.03). The total variance used for effect size calculation was 73.56, comprising variances of 62.17 at the pupil level, 2.88 at the class level, and 8.51 at the school level. The ICC associated at the school level from the adjusted model was 0.10 (95% CI: 0.06 to 0.15) and at the class level 0.04 (95% CI: 0.00 to 0.04), while the ICC (95% CI) for the empty model (without covariates) was 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.

Table 13: Primary analysis

	Unadjusted means (post-intervention)				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges' g (95% CI)	P-value
Endline New PiRA score	1,907 (403)	22.2 (21.5, 23.0)	1,721 (378)	21.3 (20.2, 22.4)	3,628 (1,907; 1,721)	0.12 (0.01, 0.23)	0.03

Secondary analysis

Secondary analysis consists of the estimation of the effects of PALS-UK on five secondary outcomes. These five outcomes include two measures of reading fluency, one derived from the WIAT-III UK-T instrument and one from the MDFS. In addition, a reading comprehension outcome from WIAT-III UK-T is also examined along with a reading self-efficacy measure and reading motivation score both obtained from the FAR questionnaire. We look at the estimates of effectiveness on each of these outcomes in turn.

Histograms were used to examine the distributions of the various post-test scores, and QQ plots to assess the distribution of the standardised residuals resulting from the relevant regressions models from which sample estimates of the effects were obtained (see Appendix H).

Oral reading fluency (rate) score

Oral reading fluency (rate) scores, collected exclusively at follow-up, were acquired using the WIAT-III UK-T reading fluency subscale. The target sample comprised ten randomly selected pupils from a single class in each school. The

final analysed sample included pupils with both the New PiRA Summer Term Year 4 Test baseline scores and endline fluency scores and consisted of 999 pupils across 106 schools (intervention n = 506; control n = 493). The mean unadjusted oral reading fluency (rate) score for the intervention group was 48.65 (95% CI: 47.8 to 49.5), while the control group score was 47.2 (95% CI: 46.3 to 48.1). The unadjusted mean difference between the two groups was 1.44 (95% CI: 0.19 to 2.69) (refer to **Appendix C**, **Table D1**).

The adjusted mean difference in oral reading fluency (rate) comes from a regression model in which a pupil's baseline New PiRA Summer Term Year 4 Test score is entered into the model as a covariate. The correlation between baseline New PiRA Summer Term Year 4 Test score and endline reading fluency (rate) score was 0.57. As the pupil sample was selected at random from a single class in each school, and the class was also selected at random, a two-level model was fitted to the data. After adjusting for covariates, the mean difference in post-test scores between the intervention and control groups was 1.16 (95% CI: 0.80 to 2.25, p = 0.04; see **Appendix C**, **Table D1**). The estimated Hedges' g effect size was 0.14 (95% CI: -0.01 to 0.27). This corresponds to approximately two months' of additional progress in the intervention over the control group. We note that CI obtained from the regression model does not include zero, while the interval associated with the effect size (Hedges' g) does. Both intervals are quite wide and the lower bound of the interval obtained directly from the regression is close to zero, implying that given our result we cannot reject quite small values for the true effect on the reading fluency rate at the 95% confidence level. The CI for the effect size was, consistent with the SAP for the study, derived using bootstrap procedures. This is a sampling-based approach and can therefore lead to coverage that will vary. As an alternative (as we have done elsewhere), we could derive approximate limits for the effect size CI by dividing through the limit obtained direct from the regression results by the effect size denominator (see following paragraph). Were we to do so, we would obtain upper and lower limits of ES=0.26 (2.25 / sqrt(72.52)) and ES=0.01 (0.08/sqrt(72.52)). In other words, a CI where in strict frequentist terms we would reject the null hypothesis of equality in means. The variation in results we obtain, reflects the natural variability of sampling-based methods and reveals that our adjusted result sits very close to the typical 5% significance threshold.

The total variance used to calculate the effect size, obtained from a model without covariates, was 72.52. This comprised variance components of 68.65 (pupil level) and 3.87 (school level). The ICC (from the empty model) associated with schools was 0.05 (95% CI: 0.02 to 0.11).

Oral reading fluency (multidimensional) score

As mentioned, scores for oral reading fluency (multidimensional), were obtained only at follow-up using the MDFS. The targeted sample included ten randomly selected pupils from a single class in each school. The final analysed sample, encompassing participants with both the New PiRA Summer Term Year 4 Test baseline scores and endline multidimensional fluency scores, totalled 1,009 pupils across 106 schools (intervention n = 512; control n = 497).

The mean unadjusted oral reading fluency (multidimensional) score for the intervention group was 12.0 (95% CI: 11.6 to 12.4), while the control group achieved a score of 11.8 (95% CI: 11.4 to 12.2). The unadjusted mean difference between the two groups was 0.19 (95% CI: -0.38 to 0.75) (see **Appendix C**, **Table D1**).

The adjusted difference in fluency score (MDFS) between intervention and control groups again comes from a two-level regression model in which a pupil's baseline New PiRA Summer Term Year 4 Test score was entered as a covariate. The correlation between baseline New PiRA Summer Term Year 4 Test score and endline reading fluency (multidimensional) scores was 0.54. Upon adjusting for covariates, the mean difference in post-test scores between the intervention and control groups was 0.10 (95% CI: -0.38 to 0.58, p = 0.68; see **Appendix C**, **Table D1**). The estimated Hedges' g effect size was 0.04 (95% CI: -0.13 to 0.19). The CI includes zero and the p-value obtained from the regression model is quite large. This suggests our results are relatively compatible with equality between the two groups. The total variance used to calculate the effect size, obtained from a model without covariates, was 8.29. This comprised variance components of 6.87 (pupil level) and 1.42 (school level). The ICC (from the empty model) associated with schools was 0.17 (95% CI: 0.12 to 0.24).

Oral reading comprehension score

Comprehension scores for oral reading were obtained at follow-up from the WIAT-III UK-T instrument. The selected sample consisted of ten randomly chosen pupils from a single class in each school. Upon receipt of endline scores, 75 pupils (from nine schools) had to be excluded due to the incorrect implementation of the WIAT-III UK-T instrument (despite rigorous quality assurance processes). Specifically, one test administrator prevented pupils from keeping the text in front of them while answering, while the test guidance states that pupils are supposed to be able to reference the text that they have read. Additionally, this administrator frequently failed to correctly apply the reverse rule (i.e. if a pupil

scores two or less on the reversal items for a given grade start point, one grade start point should be reversed, not just one item set). These systematic errors led to pupils from schools assigned to this particular test administrator consistently underscoring on the WIAT-III UK-T comprehension test. The final analysed sample, including participants with both the New PiRA Summer Year 4 Test baseline scores and endline comprehension scores, consisted of 922 pupils across 97 schools (treatment n = 467; control n = 455).

The mean unadjusted comprehension score for the intervention group was 57.5 (95% CI: 56.2 to 58.8), whereas for the control group it was 55.7 (95% CI: 54.4 to 57.0). The unadjusted mean difference between the two groups was 1.81 (95% CI: -0.03 to 3.64) (see **Appendix C**, **Table D1**).

Adjusted estimates for the mean difference in comprehension scores between intervention and control groups are, as elsewhere, obtained from a regression model in which a pupil's baseline New PiRA Summer Term Year 4 Test score is entered as a covariate. As the outcome is obtained from a subsample of pupils the model is a two-level hierarchical linear model. The correlation between baseline New PiRA Summer Term Year 4 Test score and endline reading comprehension scores was 0.60. After adjusting for covariates, the mean difference in post-test scores between the intervention and control groups was 1.47 (95% CI: 0.16 to 2.77, p = 0.03; see **Appendix C**, **Table D1**). The estimated Hedges' g effect size was 0.16 (95% CI: 0.01 to 0.31), equivalent to approximately two months' additional progress in the intervention relative to the control group. While acknowledging that the CI for this result excludes zero, it is important to note that the interval is again relatively wide. This implies that relatively small values, approaching zero, for the true difference between the groups cannot be rejected at the 95% confidence level. However, it is important to highlight that the associated p-value is small (p = 0.03), suggesting that our findings are relatively inconsistent with equality in average comprehension scores between the groups. The total variance used to calculate the effect size, derived from a model without covariates, was 81.31. This encompassed variance components of 66.63 (pupil level) and 14.69 (school level). The ICC (from the empty model) associated with schools was 0.18 (95% CI: 0.12 to 0.25).

Reading self-efficacy score

Baseline and follow-up data on reading self-efficacy scores were collected from 3,025 pupils across 102 schools (n = 1,563; control, n = 1,462). The mean unadjusted reading self-efficacy score at endline was 105.1 (95% CI: 103.6 to 106.7) for the intervention group and 104.9 (95% CI: 103.3 to 106.5) for the control group. The unadjusted mean difference between the two groups was 0.23 (95% CI: -2.00 to 2.46) (see **Appendix C**, **Table D1**).

An estimate of the adjusted difference in means in reading self-efficacy scores between intervention and control groups due to PALS-UK was obtained from a multiple regression model in which a baseline measure of pupil's reading self-efficacy was included as a covariate. The correlation between baseline and endline scores was 0.48. After adjusting for covariates, the mean difference in post-test scores between the intervention and control groups was 0.35 (95% CI: -1.67 to 2.39, p = 0.73; see **Appendix C**, **Table D1**). This translates to a Hedges' g effect size of 0.02 (95% CI: -0.08 to 0.12), indicating that the observed effect is negligible, and the possibility of a modest positive or negative impact cannot be dismissed. The CI includes zero, which means that we cannot reject zero (among other values) as the true value for the parameter at the 95% confidence level. The total variance utilised to calculate the effect size, derived from a model without covariates, was 392.69. This total variance comprised 370.97 (pupil level), 13.50 (class level), and 9.26 (school level). The ICC for the empty model (i.e. without covariates) was 0.02 (95% CI: 0.01 to 0.07) at school level and 0.03 (95% CI: 0.02 to 0.05) at class level.

Motivation for reading score

Finally, motivation for reading was assessed at baseline and endline through the FAR questionnaire. The analysed sample involved 3,019 pupils from 102 schools (intervention n = 1,561; control n = 1,458). At follow-up, the unadjusted mean motivation for reading score was 52.3 (95% CI: 51.3 to 53.3) for the intervention group and 51.4 (95% CI: 50.4 to 52.4) for the control group. The unadjusted mean difference between the two groups was 0.84 (95% CI: -0.56 to 2.25) (see **Appendix C**, **Table D1**).

As with other estimates, the adjusted difference in mean reading motivation between intervention and control groups in response to PALS-UK was obtained from a regression model in which baseline reading motivation was included as a covariate. The correlation between baseline and endline scores was 0.50. Upon adjusting for covariates, the mean difference in post-test scores between the intervention and control groups was 0.89 (95% CI: -0.33 to 2.11, p = 0.15; see **Appendix C**, **Table D1**). The estimated Hedges' g effect size was 0.07 (95% CI: -0.02 to 0.16), indicating a small effect, and additional progress of approximately one month in the intervention group relative to the control group. The CI includes zero, so that we cannot reject zero (among other values) as the true value for the parameter at the 95%

confidence level. The total variance used to calculate the effect size, obtained from a model without covariates, was 167.14, with variance components of 157.07 (pupil level), 7.87 (class level), and 2.20 (school level). The ICC (from the empty model) associated with schools and classes was 0.01 (95% CI: 0.00 to 0.09) and 0.05 (95% CI: 0.03 to 0.07), respectively.

Table 14: Secondary analysis

	Raw means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges' g (95% CI)	P-value[23]
Oral reading fluency rate (WIAT-III UK-T)	506 (1,804)	48.7 (47.7, 49.5)	493 (1,606)	47.2 (46.3, 48.1)	999 (506; 493)	0.14 (-0.01, 0.27)	0.04
Oral reading fluency (MDFS)	512 (1,798)	12.0 (11.6, 12,4)	497 (1,602)	11.8 (11.4, 12.2)	1,009 (512; 497)	0.04 (-0.13, 0.19)	0.68
Reading comprehension	467 (1,843)	57.5 (56.2, 58.8)	455 (1,644)	55.7 (54.4, 57.0)	922 (467; 455)	0.16 (0.01, 0.31)	0.03
Reading self-efficacy	1,563 (747)	105.1 (103.6, 106.7)	1,462 (637)	104.9 (103.3, 106.5)	3,025 (1,563; 1,462)	0.02 (-0.08, 0.12)	0.73
Motivation for reading	1,561 (749)	52.3 (51.3, 53.3)	1,458 (641)	51.4 (50.4, 52.4)	3,019 (1,561; 1,458)	0.07 (-0.02, 0.16)	0.15

Analysis in the presence of non-compliance

For the purposes of CACE analysis, compliance was defined at the school level, such that if a school complied, all pupils within the school were also deemed compliers. This is justified on the basis that if a school complies and engages in PALS-UK, we cannot rule out consequences for pupils from the relevant year group cohort in that setting. A complying school is one where:

- at least one teacher from an intervention school attends the initial training event;[24] and
- there is evidence from that teacher's school that one or more pupils completed all four weeks of training.

The attendance at the training events was documented by the delivery team. Data on pupil training compliance was recorded in class teacher weekly logs, research assistant observations,[25] or the survey completed during top-up training to confirm training completion. Schools were regarded as compliant if they met both compliance criteria. Conversely, schools that fulfilled only one of the two criteria or none of the criteria were considered non-compliant. CACE analysis conducted using IVs based on 2SLS was planned to obtain estimates of causal effects on compliers with cluster robust standard errors.

Initially, we assessed the extent of non-compliance by creating and evaluating a binary compliance variable (1 = complied, 0 = not complied). The aim of this analysis was to determine if appreciable non-compliance occurred (around 10% or more[26]). In such a case, further CACE analysis would be undertaken. **Table 15** presents non-compliance levels across intervention schools. Since the overall non-compliance proportion was deemed quite small—only 2 out of 53 intervention schools with valid baseline and endline reading attainment scores were non-compliant—no additional

[23] P-values obtained from the regression model.
[24] N.B. It is not possible for teachers to attend the top-up training event unless they have first attended the initial training. It is also not possible for pupils to receive training in PALS-UK unless their teacher or at least one teacher in the school has attended initial training.
[25] To support schools with the delivery of PALS-UK, two observations were completed by research assistants from Nottingham Trent University: the first observation took place during the initial four weeks of pupil training; and the second observation took place during Weeks 10-14.
[26] This cut-off was arrived based on judgement following discussions with the delivery team about the level of compliance required for non-compliance to have substantive consequences.

analysis was conducted. Adjusting the ITT results from the primary analysis by a level of non-compliance of this magnitude would have no substantive implications.

Table 15: Summary of compliance evidence as reported by the delivery team

Core components	Evidence	Number (%) of intervention schools
At least one teacher from an intervention school attends the initial training event	Yes	51 (96.2)
	No	2 (3.8)
There is evidence from teacher's school that one or more pupils completed all four weeks of training (from information held by the delivery team, which could be teacher logs, survey data, or observations)	Yes	51 (96.2)
	No	2 (3.8)
There is evidence from teacher's school that one or more pupils completed all four weeks of training (from Qualtrics[27])	Yes	30 (56.6)
	No	23 (43.4)
Evidence that schools met both compliance criteria	**Yes**	**51 (96.2)**
	No	**2 (3.8)**

Missing data analysis

The protocol and SAP specified a range of analyses with the aim of examining the consequences of missing data for the primary analysis. These analyses are explained in the 'Methods' section above. Overall, 11.9% (489 out of 4,117) of pupils were excluded from the primary analysis due to lacking an endline New PiRA Summer Term Year 5 Test score (intervention group: n = 269 out of 2,176, 12.4%; control group: n = 220 out of 1,941, 11.3%). The main reasons for exclusions were lost test results, pupils leaving the school, or being absent on the testing day, as displayed in the 'Participant flow' section (refer to the subsection 'Endline New PiRA assessments').

In the adjusted mixed-effect logistic regression with the presence or absence of endline New PiRA Summer Term Year 5 Test scores as the outcome variable, the so-called drop-out model, two covariates stood out as statistically significant predictors: ever-FSM and SEND (pupils eligible for FSM and those designated SEND were more likely to have missing data). Additional variables, such as secondary outcome scores measuring self-efficacy and motivation at baseline and follow-up, gender, and EAL, were considered as auxiliary variables in the MI process.

Missing post-test New PiRA data underwent imputation using MI by chained equations. The primary analysis model was then rerun on a series of imputed datasets, and Rubin's rule was applied to combine the treatment estimates from across these analyses.

Following MI, the adjusted mean difference in New PiRA post-test scores between the two groups was 0.95 (95% CI: 0.14 to 1.77, p = 0.02), which was comparable to the adjusted difference obtained from the primary mixed-effect model (1.02; 95% CI: 0.12 to 1.92, p = 0.03).

Additional analysis on the same imputed dataset was conducted to examine missingness using linear regression with cluster robust standard errors, maintaining consistency with the analytical model applied during the imputation phase. Estimates from the MI model were then compared to those from primary analysis results derived also from OLS regression. Once again, the adjusted mean difference in New PiRA post-test scores between the two groups (0.94; 95% CI: 0.07 to 1.80, p = 0.03) closely resembled the adjusted difference obtained from the primary OLS model (0.97; 95% CI: 0.01 to 1.93, p = 0.048).

Substantively therefore, imputing missing observations on the endline New PiRA Summer Term Year 5 Test score as well as relevant covariates and re-running our analysis on data sets with missing values filled-in using MI led to no appreciable differences with the main primary analysis such that it might alter our conclusions. Given that MI proceeds under the assumption that missing data are MAR, we can at least discount the possibility that in our case MAR processes may have given rise to the results obtained in the main primary analysis (see Appendix H for regression outputs).

Subgroup analyses

In the SAP, we specified our intention to examine subgroup effects within three distinct categories: pupils eligible for FSM at any point in the last six years (using the variable EVERFSM_6); those designated SEND; and pupils scoring in the lowest quartile on the baseline New PiRA Summer Term Year 4 Test. Additionally, we conducted exploratory analyses to investigate the impact of PALS-UK for lower attaining EAL pupils. To achieve this, we created an indicator

[27] Completed at top-up training/sent by email to non-attenders.

that combined the binary measure of EAL with a pupil's raw score on the baseline New PiRA assessment, specifically whether that score fell in the lower half of the sample distribution.

The subgroup analyses focused solely on the primary outcome measure, and as highlighted in the SAP, all subgroup analyses conducted in this evaluation were exploratory. In summary, a series of adjusted regression analyses, including interaction effects, revealed small or relatively imprecise estimated interaction effects for all four groups. The evidence suggests that FSM pupils and lower attainers benefited from PALS-UK in terms of their reading relative to their counterparts in control schools but not relative to non-FSM and higher attaining groups in intervention schools. For pupils with SEND, there is some modest evidence that not only did their reading attainment improve relative to control school SEND counterparts, but that it improved to a greater extent than for pupils not with SEND in intervention schools. We discuss this finding and its implications in further detail in the 'Conclusion' section of this report.

Looking first at the interaction between FSM status and the intervention group indicator an adjusted analysis found an interaction of 0.02 (95% CI: -0.77 to 0.82; see **Appendix C**, **Table D2**). For the FSM-only sample, an adjusted mean difference in the New PiRA Summer Term Year 5 Test score between intervention and control groups of 0.76 (95% CI: -0.30 to 1.81) was found. This is equivalent to Hedges' g effect size of 0.09 (95% CI: -0.04 to 0.22, p = 0.16).

For pupils with SEND, an interaction between the intervention and SEND status variable of 0.97 was found (95% CI: -0.02 to 1.96; see **Appendix C**, **Table D2**). The p-value for the interaction was 0.054. This result does not reach statistical significance with respect to a formal test of the nill null hypothesis at the 95% confidence level but has a p-value that is nonetheless quite small. This may be interpreted as indicating a fair degree of discrepancy of our result with the nill null. Looking at just pupils with SEND, the adjusted mean difference in the New PiRA Summer Term Year 5 Test score is 1.84 (95% CI: 0.62 to 3.06). This is equivalent to Hedges' g effect size of 0.21 (95% CI: 0.07 to 0.34, p = 0.003).

For pupils within the lowest quartile on the New PiRA Summer Term Year 4 Test baseline (n = 872), an interaction effect of 0.22 was obtained (95% CI: -0.65 to 1.08; see **Appendix C**, **Table D2**). In this particular subgroup, the adjusted mean difference in scores on the New PiRA Summer Term Year 5 Test is 1.53 (95% CI: 0.37 to 2.69). This corresponds to a Hedges' g effect size of 0.22 (95% CI: 0.05 to 0.39, p = 0.01).

Among 442 EAL pupils scoring in the lower half on the New PiRA baseline test distribution, we obtained an interaction effect of 0.26 (95% CI: -1.00 to 1.51; see **Appendix C**, **Table D2**). In this specific subgroup, the adjusted mean difference in scores on the New PiRA Summer Term Year 5 Test is 0.98 (95% CI: -0.99 to 2.95). This translates to a Hedges' g effect size of 0.13 (95% CI: -0.13 to 0.38, p = 0.33).

Additional analyses and robustness checks

For the primary outcome, two further analyses were performed. The first sensitivity analysis employed a simplified regression model, excluding the pupil baseline measure of reading attainment. This approach aimed to evaluate the impact of including the baseline New PiRA Summer Term Year 4 Test score as a covariate on the accuracy of sample estimates in the main primary analysis. The adjusted mean difference in post-test scores between the intervention and control groups was 0.79 (95% CI: -0.44 to 2.02). The estimated Hedges' g effect size was 0.09 (95% CI: -0.05 to 0.23). In contrast to the primary analysis that incorporated the New PiRA baseline covariate, the p-value of 0.21 suggests that the results are somewhat compatible with the equality of average reading attainment between the two groups. Since there was relative balance between trial arms in the baseline New PiRA within the analysed sample (see **Table 12**), incorporating the baseline scores as a right-hand side variable did not account for observed variations between trial arms. Instead, it effectively reduced the outcome variance, thereby significantly enhancing the statistical power of our estimates.

The second form of sensitivity analysis mirrored the regression model used for the primary analysis but with the age-standardised New PiRA score obtained at follow-up as the dependent variable instead of the raw score. This specification omitted the age covariate previously included for the primary analysis. This second specification enabled us to assess how far direct age-standardisation of the outcome might influence results. The findings reveal that the adjusted mean difference in post-test scores between the intervention and control groups amounted to 1.63 (95% CI: 0.07 to 3.18). It is important to highlight that although this mean difference is greater than that observed in the primary analysis, the standardised New PiRA is measured on a scale ranging from 70 to 130, in contrast to the unstandardised New PiRA, which can range from 0 to 40. The estimated Hedges' g effect size was 0.11 (95% CI: 0.00 to 0.21). The p-value associated with this finding as p = 0.04, similar to that obtained in the main primary analysis.

Implementation and Process Evaluation

As noted above, the main change to the IPE design was to conduct the second round of case study data collection online. This meant that it was not possible to hold a second round of focus groups with pupils. Instead, we asked teachers at the case study schools to ask their pupils some questions we had prepared for them (similar to those asked during the focus groups) and to report back on this in their online interview. This decision was made to accommodate a change to endline data collection for the impact evaluation. We originally planned to recruit 20 pupils from our university to undertake this task for us. However, we became concerned about meeting this target. Instead, we subcontracted this activity to AlphaPlus and were required to make some adjustments to the budget in order to accommodate this expenditure. This was approved by the EEF evaluation manager.

Table 16 presents a summary of the planned samples for data collection tools and numbers of responses received.

Table 16: IPE data collection summary

Data collection methods	Participants / data sources (planned sample)	Participants / data sources (responses received)
Headteacher survey: pre-intervention	All, n = 120	86 of 105
Teacher survey: pre-intervention	All, n = >120	163 from 95 of 105 schools
Observation field notes of: initial training (two sessions); top-up training (two sessions); and twilight training (one session)	Delivery team (two persons) Teachers: initial training (all attending session); top-up training (all attending session); and twilight training (all attending session)	Delivery team (two persons) Teachers: Initial training (41); top-up training (33); and twilight training (seven)
Observation field notes of PALS-UK session	Case study class / field notes (at all six case studies on two occasions, n = 12)	Six observations
Observation checklists of four x PALS-UK sessions (two conducted by a researcher, two conducted by a peer review teacher)	Teachers (intervention only, all, n = >60)	First researcher observation: 51 of 53 schools First peer observer observation: 42 of 53 schools
Teacher logs of PALS-UK delivery	Schools (intervention only, all, n = 60)	47 schools completed logs, two confirmed full delivery via email
Interviews with peer observers in case study schools (second visit only)	Teachers (intervention only, n = <6)	Two
Interviews with teachers and headteachers (two visits)	Headteachers/senior leaders (n = 6 x 2) Teachers (n = 6 x 2)	Headteachers/senior leaders (5) Teachers (6)
Interviews with a group of pupils (two visits)	Case study school pupil sample (from six schools, six to eight pupils on two occasions, n = >36)	45
Headteacher survey (intervention): post-intervention	All, n = 60	31 of 53
Headteacher survey (control): post-intervention	All, n = 60	26 of 52
Teacher survey (intervention): post-intervention	All, n = >60	72 from 41 of 53 schools
Teacher survey (control): post-intervention	All, n = >60	50 from 38 of 52 schools

.

Usual practice

This summarises usual practice in all schools at baseline as well as any additional reading interventions implemented in intervention schools during the trial (greater detail is provided in Appendix F of the Technical Notes published alongside this evaluation report). It also examines control schools' usual practice during the trial, which answers the following research question:

2. What comparable initiatives were undertaken within control group schools and did they change over time?

The section draws on data from the staff surveys conducted at baseline and endline.

Usual practice in all schools at baseline

Teaching reading comprehension

Teachers reported that the three most frequently taught aspects were vocabulary, accuracy, and prosody for control schools and vocabulary, accuracy, and automaticity for intervention schools.[28] Many of them said that they used guided reading or shared reading sessions, often including activities that focused on VIPERS (Vocabulary, Inference, Prediction, Explanation, Retrieval, and Sequence or Summarise) skills. Teachers reported that reading lessons were delivered as a whole class, in small groups or individual and included giving pupils frequent opportunities to read in the classroom and listening to adults modelling reading. Activities focused on skills including inference, summarising, retrieval, vocabulary, prediction, comparing, skimming, and scanning, using background knowledge and answering test questions. Many teachers indicated that activities were often modelled, scaffolded, and supported with interventions. Many teachers reported that pupils were encouraged to read at home through homework tasks or the use of reading journals, and that parents were given support with this. Data indicates that most classes across all schools spent up to two hours reading plus up to two hours completing reading-related activities in class each week at baseline, with a minority of schools spending up to three and a half hours on each.

Reading programmes

A larger proportion of teachers in intervention schools (40%) compared to control schools (29%) indicated that they used literacy interventions. Accelerated Reader was the most commonly named intervention across both groups. Teachers reported that struggling readers were supported by specific interventions in 73% of intervention schools and 48% of control schools. Around two-thirds of headteachers reported that at least one reading programme was available for Year 5 pupils in their school (intervention 65%; control 70%). The majority of programmes lasted more than 20 weeks. Intervention school headteachers reported that a higher proportion of programmes were short (5–10 weeks) compared to control schools (intervention 20%; control 5%). Programmes across both arms of the trial were most commonly delivered at least once a week (n = 21 in intervention schools and n = 19 in control schools) or once a day (n = 16 in intervention schools and n = 14 in control schools).

Guided reading

The majority of teachers (intervention 65%; control 67%) and headteachers (intervention 73%; control 82%) reported the use of guided reading. Five headteachers said that their school used the reciprocal reading approach. Teachers reported that they delivered their guided reading lessons either in small groups or as a whole class. A smaller number of teachers mentioned pair work. Headteachers reported that guided reading was more commonly delivered as a whole class (intervention schools n = 19, control schools n = 18) than small groups (intervention schools n = 9, control schools n = 8). Teachers and headteachers both reported that guided reading was often delivered with the support of teaching assistants. Teachers reported that guided reading lessons would usually be delivered multiple times a week, ranging between daily and weekly. Lesson length varied between 20 minutes and 60 minutes. Overall, guided reading lessons were commonly delivered to similar extents and using similar methods by schools across both arms of the trial at baseline.

[28] Teachers' baseline survey question: 'Q16: How often do you teach the following aspects of reading comprehension in your lessons? Accuracy, Automaticity, Prosody, Re-telling, Summarising, Prediction, Vocabulary, Inference'.

Peer tutoring

Roughly even proportions of intervention (23%) and control schools (24%) used peer tutoring as a strategy to support the teaching of reading, although this was less common than guided reading across all schools. Teachers reported that this would usually be in mixed attainment pairs while fewer teachers mentioned that pairs would coach each other. Most teachers reported that these sessions were scheduled regularly, although some suggested that they were organised *ad hoc* based on teaching plans. Fewer teachers said that paired reading was used only with lower attaining readers.

Reading culture

A large majority of teachers agreed or strongly agreed that: i) there is an integrated whole-school approach to cultivating confident readers (intervention 87%; control 90%); and ii) there is an integrated whole-school approach to developing readers who are self-motivated to read (intervention 85%; control 87%). Strategies used to support a positive culture around reading for pleasure including providing school and classroom libraries and other spaces in the school to read, online resources, timetabled opportunities to read for pleasure, pupil choices of books, wall displays promoting reading, and other initiatives such as author visits and reward schemes for reading.

Continuing professional development (CPD) for teachers

Teachers and headteachers at intervention schools differed in their views when asked about the provision of formal CPD to support the teaching of reading (**Table 17**). Training was typically delivered in school and was usually designed in school (by the literacy lead and English team), occasionally using an outside provider or training for a particular programme. Teachers across intervention and control schools described a range of informal CPD activities including independently reading blogs, books, professional magazines, research and evidence-based publications, and attending webinars. Teachers also described engaging with social media, including Facebook groups, Twitter, and Instagram, and engaging with peers through conversations, meetings across Trusts, and teaching observations.

Table 17: Survey responses in relation to CPD[29]

	Intervention schools		Control schools	
	Teachers	Headteachers	Teachers	Headteachers
Formal CPD for teaching reading	34%	68%	52%	54%
Informal CPD for teaching reading	23%	N/A	26%	N/A

N/A, not applicable.

Future plans for teaching reading

Around a quarter of headteachers across both arms of the trial had planned changes to the ways in which reading was taught in Key Stage 2 in the academic year during which PALS-UK was delivered (2023/2024) (intervention 22%; control 24%). These included strategic changes such as curriculum review, having reading as a focus in school development plans, or having a whole-school focus on reading. Others referred to the introduction of particular reading programmes, materials, or strategies.

Usual practice in intervention schools during the trial

Intervention schools' practices in teaching reading were largely a continuation of those at baseline, with some schools making changes to strategy or practice while reducing time on other reading activities to make space on timetables for PALS-UK.

Teaching reading alongside PALS-UK

Teachers and headteachers reported that they continued to teach reading comprehension and fluency in a similar manner to that at baseline alongside PALS-UK, however to a reduced extent. Intervention schools maintained or reduced the number of programmes that they offered to Year 5 pupils, as well as reducing the time dedicated to reading and reading activities outside of PALS-UK lessons (from around four hours a week at baseline to two hours a week).

[29] Headteachers were not asked about informal CPD in the baseline survey. The data reported from headteachers here relates to the question: 'Q22. Were your Year 5 staff offered any CPD that relates to the teaching of reading in 2021/22?'

Fewer schools used guided reading with others reducing the time spent on it. Thus, PALS-UK typically replaced some of the existing methods of teaching reading to Year 5 pupils.

Around 43% of teachers (n = 31) used additional literacy interventions (typically Accelerated Reader, n = 21) in Key Stage 2 for all pupils (irrespective of ability) alongside PALS-UK, a similar proportion to baseline. Around 58% of teachers (n = 41) used existing literacy interventions in Key Stage 2 for pupils struggling to read, a lower proportion than indicated at baseline (73%). Around 77% of headteachers said that their school did not have any other reading programmes running for Year 5 pupils during PALS-UK, and 23% of headteachers reported one or two others running (a reduction from baseline).

CPD

A slightly lower proportion of teachers (24%) compared to baseline reported that they had undertaken formal professional development that year, aside from PALS-UK training. Around 26% said that they had undertaken informal professional development that relates to teaching reading, a figure roughly equal to baseline. A much higher proportion of headteachers (65%, n = 17) reported that CPD had been offered to Year 5 staff that related to the teaching of reading, a similar proportion to baseline.

Reading culture in schools

The majority of teachers remained positive about developing the reading culture in their schools with slight increases from baseline responses to questions on this. Teachers described strategies such as investments reading facilities and resources (e.g. libraries and books). Timetabling reading and in-school initiatives to promote reading (e.g. wall displays, assemblies, and competitions) were also commonplace. These comments were echoed by headteachers in their responses to questions on the development of reading culture.

Changes to practice

Around 22 teachers (31%) referred to changes in their practice including the new interventions (e.g. Accelerated Reader), new strategies (e.g. choral reading), changes to guided reading approaches, an emphasis on particular reading domains, and the promotion of reading for pleasure. In contrast, 48 teachers (69%) said they had not changed their practice. Around 22 teachers (31%) felt that changes during the academic year could have impacted on their pupils' reading development. Around 18 (69%) headteachers indicated that teaching reading had been a school development focus for resulting, for example, in the introduction new reading programmes, approaches or new initiatives, and/or increasing reading time in schools.

Usual practice in control schools during the trial

Teaching reading

Teachers in control schools reported that reading comprehension skills were taught frequently with vocabulary, inference, and prosody being taught particularly frequently. The aspects that PALS-UK lessons focus on, including fluency, retelling, summarising, and prediction, were all taught frequently by the majority of control schools. Teachers indicated that a range of approaches were taken to teaching reading comprehension. Of note, five teachers said that they used paired reading, including one who specified that the paired reading was mixed ability. A minority of control schools (17%, n = 8) had used peer tutoring to support the teaching of reading that year. This represents a slight reduction in the proportion of positive responses from baseline.

In control schools, the use of literacy interventions in Key Stage 2 for all pupils (irrespective of ability) was reported by 38% of teachers (n = 18). Around 62% indicated that they had used literacy interventions in Key Stage 2 for pupils who are struggling to read. In both cases, the proportion of teachers reporting the use of interventions had risen from baseline.

Headteachers in control schools indicated the reading programmes available to Year 5 pupils in their schools with just over half identifying at least one (52%, n = 16). The programmes were delivered by teachers (42%, n = 11), teaching assistants (23%, n = 6), or teachers with teaching assistants (27%, n = 7), either to a whole class (n = 9), small group (n = 5), or through independent work (n = 5). Three-quarters of the programmes were delivered for up to an hour per week (74%, n = 19), typically once a week (42%, n = 11), or once a day (39%, n = 10).

In terms of the training that staff in control schools had received to deliver these programmes, it was more commonly conducted in-house, often using training materials or delivered by staff who had received training. Slightly fewer headteachers responded that training was provided by the programme.

Overall, control schools provided a range of programmes and taught reading comprehension and fluency in a variety of ways. Teachers reported teaching the aspects of reading that PALS-UK focuses on, through diverse approaches, and involved less use of peer tutoring than at baseline. The way that control schools taught reading was then quite dissimilar to the PALS-UK approach. Responses were comparable to baseline and did not show any substantial changes over time.

Reading culture

Around 28% of teachers and headteachers in control schools said that there had been changes to the reading culture in their schools, including the purchase of new resources, new school initiatives, and a highlighted focus on reading and reading for pleasure. Around 64% of teachers perceived that their pupils' reading for pleasure had increased. The majority of teachers remained positive about developing the reading culture in their schools with little change from baseline perceptions.

Teachers in control schools described strategies such as investments in reading facilities and resources (e.g. libraries and books). Timetabling reading and in-school initiatives to promote reading (e.g. wall displays, assemblies, and competitions) were also commonplace. As for intervention schools, these comments were echoed by headteachers in their responses to questions on the development of reading culture.

CPD

In control schools, 47% of teachers at endline had undertaken formal professional development (in-house or external) relating to teaching reading and 30% had undertaken informal development activities (using social media, reading articles, and consulting colleagues). These were similar proportions to baseline. Around 65% (n = 20) of headteachers said that development opportunities relating to the teaching of reading had been offered to Year 5 staff, a small increase from baseline. Formal CPD related usually to particular reading skills, phonics, specific programmes or approaches, or reading for pleasure. Compared to intervention schools, the proportion of teachers in control schools that reported undertaking formal CPD was significantly higher, and informal CPD slightly higher, with an equal proportion of headteachers reporting that development opportunities had been offered. The difference in reports from teachers may be expected given that the question asked intervention schoolteachers to discount any CPD related to PALS-UK.

Changes over the year

Around 34% of teachers in control schools reported changes to their practice in relation to teaching reading including new approaches or programmes. One teacher said that they were using peer tutoring more. A minority of teachers (11%, n = 5) reported that changes during the year that could have impacted on the development of pupils' reading skills. Around 29% (n = 9) of control school headteachers felt that the way in which reading was taught in their schools had changed including a renewed focus on reading for pleasure, more whole-class reading and analysis, and a focus on lower attaining readers. Many teachers reported that there had not been any particular changes to how reading was taught that year, while some reported that their school had introduced a more consistent or formalised focus on teaching key reading skills. Around 48% (n = 15) of headteachers at endline also said that reading had been a school development focus in Key Stage 2, highlighting four main areas of development: reading resources; teaching particular reading skills or encouraging reading for pleasure; CPD for teachers; and a focus on improving and monitoring outcomes.

In sum, control schools had to a large extent continued with usual practice. While some survey respondents reported changes to how reading is taught or the reading culture in schools, few schools reported something happening that would impact on their pupils' reading skills.

Summary of key findings

- Schools in both arms of the trial were in similar positions at baseline, spending up to four hours per week on reading and reading-related activities. Many schools offered at least one reading programme delivered at least weekly. Reading programmes were more available in intervention than control schools according to teacher reports, although this difference was reversed and much

reduced as reported by headteachers. The majority of schools used guided reading delivered to the whole class, while about a quarter indicated that they used peer tutoring. Most schools supported their reading culture by investing in resources, dedicating space on the timetable and organising initiatives and events. About a quarter of teachers had recently engaged in informal CPD, while a higher proportion of control schoolteachers (52%) compared to intervention schoolteachers (34%) had engaged with formal CPD. Around a fifth of schools reported plans to make changes to the way reading was taught, including introducing particular strategic foci or new reading programmes.

- During the trial, intervention schools reduced the time spent on other reading activities and programmes while PALS-UK was being delivered, while continuing to teach reading comprehension and fluency and create a positive reading culture in a variety of ways, as they had at baseline. A slightly lower proportion of teachers had undertaken formal CPD than the year before. Nearly a third of teachers indicated that they had changed their practice in teaching reading and a quarter felt that other changes over the year could have impacted on their pupils' learning. A larger proportion of headteachers (69%) said that reading had been a focus for school development that year and mentioned examples of strategic and practical changes.
- Control schools had largely continued with usual practice during the trial. Schools continued to teach reading through a diversity of approaches, similar proportions of teachers continued to engage with CPD opportunities, and few schools reported something happening that would impact on their pupils' reading skills. A sizeable minority of survey respondents did however, report changes to how reading was taught, with half of headteacher respondents saying that reading had been a focus for school development. A significant minority of teacher respondents also stated that there had been changes to the reading culture in their schools, with a majority saying that they had seen evidence of their pupils' reading for pleasure increasing, suggesting an increased focus on this in control schools over the year.
- Compared to control schools, changes in intervention schools across the different measures were reported by similar or larger proportions of staff, suggesting that intervention schools may have been more active in making changes. It is important to note that these changes to usual practice that intervention schools mentioned implementing in addition to using PALS-UK might have impacted the outcomes of the trial.

Fidelity

This section examines the extent to which the PALS-UK intervention was delivered as intended to all in the treatment group and explains any areas where this was not the case. It answers the following research question:

1.1. To what extent did fidelity vary and why?

The aspects of the programme that will be examined are training and support for teachers, pupil training, pupil pairing, and the four structured PALS-UK reading activities. Data for this section comes from the delivery team's monitoring systems, the endline teacher and headteacher surveys, the evaluation team's observations of the training events, and observations of the delivery of PALS-UK lessons conducted by peer observers, research assistants, and the evaluation team.

Training and support

Attendance at the training

As reported in the compliance section of this report, attendance at initial training by the 53 intervention schools with valid baseline and endline reading attainment scores was very high: 51 schools attended; and two did not.[30] The two schools that did not attend were non-compliant and did not go on to deliver PALS-UK, thus all compliant schools attended initial training.

[30] When we include the one intervention school that is not included in the primary analysis, 52 attended initial training and two did not.

Attendance data from the delivery team suggests that 75 teachers from 47 schools attended follow-up training, while teachers from seven schools did not attend.[31] Those that did not attend told the delivery team that there were urgent matters in school, including safeguarding issues and parent meetings.

The delivery team also provided attendance data on the twilight sessions. Two sessions were organised, with 11 teachers attending all together. Data from the endline survey completed by intervention schoolteachers contradict this slightly: 16 people (22%) indicated that they did attend a twilight session while 56 (78%) suggested they did not. This might indicate some confusion between attending top-up training and twilight sessions when responding to the survey question.

Delivery of the training

The initial training was delivered across seven regional locations as full-day, in-person training. The days involved a combination of live presentations and opportunities for practising the PALS-UK activities, ensuring that schools left prepared to deliver the PALS-UK programme. Participants were seated on large tables, a layout that facilitated interaction between schools.

The days began with presentations giving a clear overview of PALS-UK and the trial, alongside the science of reading. This was followed by discussion of the four PALS-UK activities, each introduced with a video clearly showing how the intervention works in practice and a presentation on why and how to complete the activity, including practical advice. Teachers were then able to practice the activity first with the facilitator and then in pairs. Teachers were invited to ask questions and complete a quiz at the end of each of these sections to check their understanding. The day continued with guidance and discussion on a range of other issues including motivating pupils, pairing, and triads, and how to use the manual. It ended with teachers taking part in a full PALS-UK lesson.

A paper copy of the manual and pupils' Correction Cards and Question Cards were given to all schools at initial training. The Microsoft Teams channel was also used to distribute these resources alongside slides and videos from the training day.

Overall, we found that the initial training day was engaging and thorough, with clear guidance on how to implement PALS-UK and opportunities for practice and was likely to fully equip schools to deliver PALS-UK.

The top-up training was a 90-minute online session delivered after the completion of the four-week pupil training period. Six sessions were organised. Each session involved presentations that recapped and extended information from the initial training on the content and rationale for PALS-UK lessons and guidance on delivery. There was an in-depth focus on mini lessons, as well as pairing and book choice as noted in the logic model. The sessions were highly interactive with time for participants to ask questions, share experiences, and test their knowledge through quiz questions completed in the chat function on Microsoft Teams. The facilitators were responsive to participants' queries, giving advice on how to manage situations that were arising and providing reassurance. Overall, the top-up training provided an effective way to check in with teachers after they had experienced delivering PALS-UK, and teachers did not raise any major issues in implementing PALS-UK.

The twilight sessions were 45 minutes in length and delivered online at around the midpoint of the trial. Two sessions were organised. The delivery team gave reminders about the PALS-UK activities, the resources available, and the overall timeline of the trial. In the session that we observed, a small number of teachers shared experiences and asked questions around how PALS-UK was working in their school context. As well as these twilight sessions, ongoing support was provided through a dedicated Microsoft Teams channel in addition to email and phone.

Implementation of the PALS-UK activities: Pupil training; pairing; and structured reading in pairs

Dosage

The PALS-UK programme involves four weeks of pupil training followed by 16 weeks of PALS-UK lessons.

[31] When we discount the one intervention school that is not included in the primary analysis, 47 attended follow-up training and six did not.

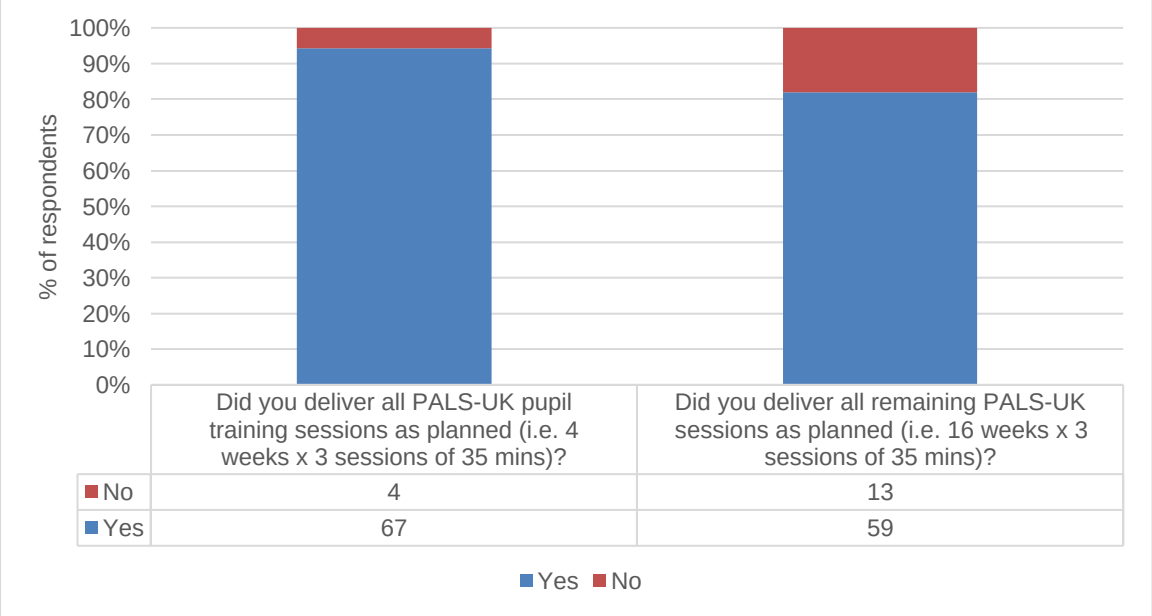
Monitoring data was provided by the delivery team on dosage. Data on pupil training has been reported in the 'Compliance' section of this report, indicating that in 51 schools one or more pupils completed all four weeks of training, with two schools being non-compliant.[32]

Monitoring data from the delivery team also provides evidence that 49 schools delivered one or more PALS-UK lessons per week for the full 20 consecutive weeks; there is no evidence of this from two schools[33]. **Table 18** reports on the number of sessions delivered by compliant schools during the 20-week period.[34] PALS-UK lessons are designed to be delivered three times a week over the course of the programme.

Table 18: Monitoring data on the number of PALS-UK sessions delivered in schools during the 20-week period

	Minimum number of sessions	Maximum number of sessions	Mean (SD)	Number of schools completing more than 45 sessions	Number of schools completing more than 50 sessions	Number of schools completing more than 55 sessions
Compliant schools (n = 51)	25	60	51 (6)	46	33	11
Schools delivering 20 weeks (inc. weeks with 0 sessions) (n = 47)	32	60	51 (5)	43	32	11

Figure 5: Programme fidelity dosage in the first four weeks and the remaining 16 weeks reported by teachers[35]



Across both rows of **Table 18**, the data suggests some variation in the number of lessons that schools delivered. The majority, however, did deliver more than 45 sessions, namely, at least two sessions a week, and the average number of lessons delivered (n = 51) was close to what would be expected if three lessons were delivered each week (n = 60).

Survey data on dosage from teachers at endline is reported in **Figure 5**. It indicates that fidelity was particularly high during the first four weeks of the programme, when teachers were delivering pupil training, with over 90% of respondents

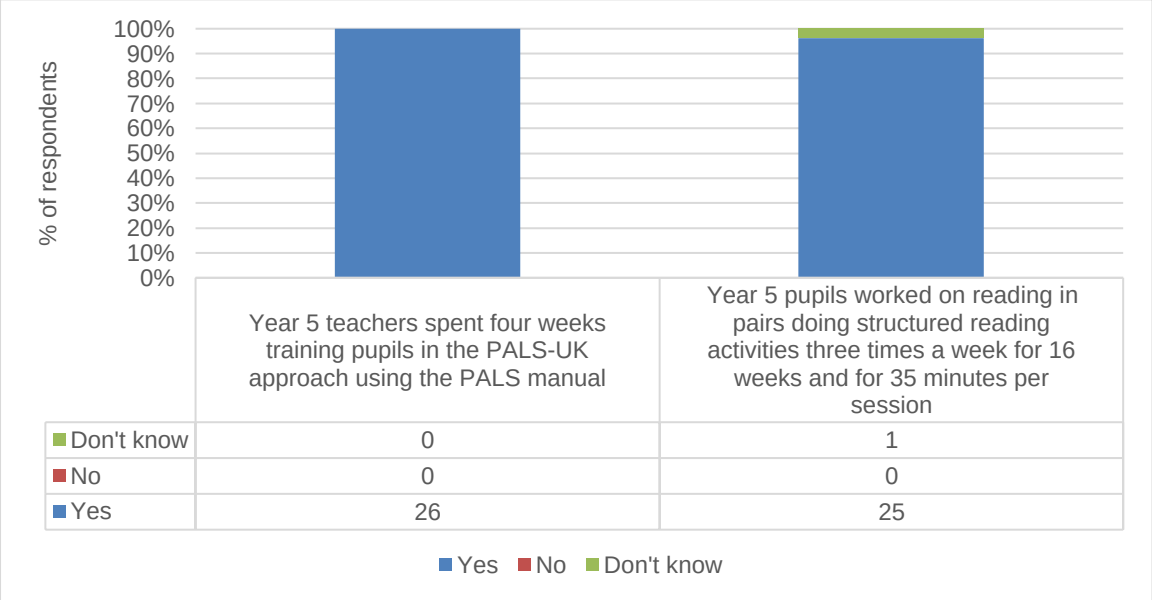
[32] Including the one intervention school that is not included in the primary analysis, there is evidence that 51 schools completed the pupil training, and no evidence from three schools.
[33] This is based on evidence from the 47 schools that completed weekly teacher logs detailing their delivery of PALS-UK, plus a further two schools that confirmed directly with the delivery team that they had completed 20 weeks.
[34] The first line of **Table 18** includes all compliant schools, including those who did not complete all of the weekly logs. The second line of **Table 18** only includes the 47 schools that completed all weekly logs, as the further two schools were not able to confirm the number of sessions that they had delivered.
[35] Intervention Teachers' Endline Survey questions Q34 and Q35.

reporting full fidelity. Fidelity appeared to fall slightly in the main 16 weeks of programme delivery, however, remained high at over 80% full delivery.

Teachers not delivering PALS-UK as intended were asked to describe and explain how their delivery differed. The main reasons given by schools for missed sessions were events on the school timetable, strike days, and absence of the staff that were trained to deliver PALS-UK. Some schools explained that they were able to catch-up with the sessions at the end of the programme, while others reported that they were not able to finish the full 60 sessions. One school reported that they '*shortened the [pupil] training and started the sessions in full as the children picked up the different tasks easily*' (T survey).

Headteachers were also asked about dosage at endline (**Figure 6**). Responses indicate perceived high fidelity during the four weeks of pupil training as well as the 16 weeks of PALS-UK activities.

Figure 6: Programme dosage as reported by headteachers[36]



Fidelity of programme delivery

Four observations were conducted in each school using a standardised observation schedule that included items on classroom set-up and the four activities (partner reading, retell, paragraph shrinking, and prediction relay). Two observations each were carried out by peer observers (during Weeks 5–9 and Weeks 15–20) and the delivery team plus two research assistants employed by the delivery team (during Weeks 1–4 and Weeks 10–14). Observers recorded whether a series of behaviours were observed, not observed, or not applicable. This results in an overall fidelity score.

Fidelity scores for the first peer observation and the first external observation were provided by the delivery team and are reported in **Table 19**[37]. The score is a weighted average percentage, computed by considering the mean score within each section of the observation schedule. Items marked 'not applicable' have been scored as observed. Each school score was calculated from the mean fidelity score across classes.

The meaning of these fidelity scores is translated by the PALS-UK delivery team into a marker of implementation quality as follows:

- 90%+: excellent implementation fidelity;
- 80–90%: very good implementation fidelity;
- 70–80%: adequate implementation fidelity; and

[36] Intervention headteachers' endline survey question: 'Q26. Please indicate whether the following activities occurred during the implementation of the PALS-UK programme in your school'.
[37] We are reporting data from the first observations here as a more reliable indicator of fidelity due to incompleteness of the data at the second observations.

- less than 70%: poor implementation fidelity.

Table 19: Distribution of schools by implementation fidelity score categories according to two observation points[38]

Observation point	Schools with excellent implementation fidelity (90%+)	Schools with very good implementation fidelity (80–89%)	Schools with adequate implementation fidelity (70–79%)	Schools with poor implementation fidelity (0–69%)
Research assistant observation 1	41	5	4	1
Peer observation 1	37	3	2	0

The data indicates that across these two observations, the majority of schools had excellent implementation fidelity. The overwhelming majority had at least adequate fidelity.

Observations were also carried out in case study schools by the evaluation team. We observed eight PALS-UK lessons, including one in each single-form entry case study school and two in the multi-form entry schools. We saw engaged pupils, expressive reading and supportive partnerships, and we noted that procedures for classroom set-up and classroom management were well followed. While most of the pupils we observed were reading the books provided with PALS-UK, in two schools we saw some pupils reading other books (at a lower reading level). In the PALS-UK training, teachers were instructed to use their existing books to supplement the books provided by PALS-UK to ensure a sufficient range of texts to match ability and interest. It is important to note that PALS-UK can be used with any texts; the book selection was provided to prevent limited book resources being a barrier to implementation.

The procedure that seemed to be followed with less fidelity in the case study schools was the 'check-it' procedure. Partners would often correct each other without following the full procedure or using the words 'check-it'. In one class the teacher was absent, and the class seemed less engaged, and there were two pairs in that class who were struggling to read the texts and so not completing the activities. In general, however, we observed that PALS-UK lessons were set-up and delivered with high fidelity across the case study schools.

Teachers who were peer observers in multi-form entry schools were asked the extent to which PALS-UK was delivered as intended in the classrooms they observed. The majority indicated that PALS-UK was delivered according to the manual. There were a small number of comments relating to reinforcing the 'check-it' procedure as a learning point from observations, and one respondent said that teacher turnover in one school reduced the fidelity of PALS-UK delivery. There were also two schools where a teacher had been unable to attend the initial training and was briefed on PALS-UK procedures in school; in one case, the peer observer reported that the effectiveness of the lesson delivery was reduced. This supports the importance of all staff who will be involved in PALS-UK being able to attend the training.

Data from headteachers on pupil pairing indicates that this was carried out with high fidelity (25/26 respondents, one respondent indicated 'don't know') and 22/26 respondents reported that Year 5 teachers changed pupil pairs approximately every four weeks (three responded 'don't know' and one responded 'no'). Most teachers reported that they had changed the pairs approximately every four weeks (72%, n = 52) compared to 28% (n = 20) who had not.

In terms of pairing fidelity, teachers reported several reasons for not changing pairs as often as including changing pairs when necessary or having a longer period for each pairing, usually a half-term. Maintaining existing partnerships was seen as a positive choice by some teachers. For example, they changed the pairs only according to friendships, new pupils arriving, or if pupils requested a change. Conversely, pairs were maintained because they worked well for those particular pupils, pupils were settled or did not want to change, or teachers felt that pupils' confidence and coaching skills were improving, and teachers wanted to maintain progress. Some teachers only changed pairs when they reached the end of a book, to increase engagement with the text and pupil satisfaction. One school had a longer initial run to give pupils the time to practice after the training while another maintained the pairs if sessions had been missed during the four-week pairing window. On the other hand, some teachers reported that their choices for changing were constrained by factors including smaller class size, the knowledge that some pairs would not work well together, or ability level, for example:

[38] N.B. No observations were carried out in three schools at research assistant observation 1. No observations were received from 12 schools at peer observation 1.

> *We could not find enough suitable pairs to be able to change regularly, due to the wide disparity in reading ability within the class.* (T survey)

Conversely, 31% (n = 22) of teachers said that they had made extra changes (i.e. more frequently than every four weeks) to pairs. In the case study interview data, all teachers reported broadly following the guidelines for the initial pairing and regular reassignment of pairs. However, as will be discussed in relation to research question 1.2 below (enablers and barriers), teachers needed to take into account personalities/relationships between pupils as well as attainment levels when pairing pupils.

A significant proportion of intervention schoolteachers made or considered making adaptations to PALS-UK lessons, with 32% saying that they had compared to 68% saying that they had not.[39] A common reason given for making adaptations was a reduction in the time spent on PALS-UK lessons, usually by shortening each activity, and explained in terms of school timetabling pressures. In other cases, teachers adapted their delivery of PALS-UK in response to pupils' behaviour and learning needs. One teacher, for example, altered the retell activity in response to pupil feedback so that both first and second readers completed it, while another said:

> *We reduced the prediction relay activity because the children found it most difficult despite mini lessons.* (T survey)

In other cases, teachers adapted or extended the PALS-UK lessons to enhance a particular aspect:

> *After a partnership completes a book, they were given some paper to write a summary of what they have read.* (T survey)

Overall, however, the survey data indicates that the main components of PALS-UK were delivered with high fidelity.

Similarly, the teaching staff interviewed at all six case study schools reported that they delivered the programme with high fidelity, being quite strict about adhering to the procedures set out in the training and the manual. Where adaptations were made, they were described as either minor in nature (e.g. making very minor changes to the wording of instructions for the pupil or having them displayed on a PowerPoint so that they did not need to be repeated quite as often or applying to just one or two pupils) or as affecting a small number of pupils (e.g. a pupil who was new to English was paired with a teaching assistant and used a book that was at a lower reading level than those books provided with PALS-UK but engaged with the same PALS-UK activities).

The logic model is based on a series of core activities taking place including the initial training and distribution of the materials, followed by the four-week pupil training, top-up training, and the 16 weeks of PALS-UK activities (including the four activities and correct procedures). Pupils should be paired by ability and pairs swapped every four weeks. This should be supported online throughout and through twilight support sessions. Although some teachers reported using professional judgement to make adaptations to how they delivered the programme, in general the findings indicate that these core activities were completed with high fidelity.

Summary of key findings

- All compliant schools attended initial training and the majority of schools attended top-up training. Around one-fifth of teachers indicated that they had attended a twilight session; however, the numbers of teachers that actually attended the twilight sessions was fewer than this (as recorded through attendance on Microsoft Teams), which suggests that teachers may have been confused between the different kinds of training sessions.
- The initial and top-up training was delivered in a very similar way across sites and provided a thorough and engaging introduction to delivering PALS-UK. Materials were effectively distributed to participating schools.
- Online support was available from the delivery team, including through the twilight sessions, throughout the trial.
- In terms of dosage, the delivery team's monitoring data indicates that all compliant schools completed pupil training. Monitoring data also suggests that most schools delivered 20 weeks of at

[39] Intervention teachers' endline survey question: 'Q36. Have you made or considered making any adaptations to the PALS-UK lessons as described in the manual?'

least two sessions a week, with some variation and a minority of schools that appeared to have delivered less than this. Survey data also indicates very high fidelity during the period of pupil training and relatively high fidelity for the remainder of the programme.

- Fidelity scores from the observation data indicate that a large majority of schools implemented PALS-UK with very good or excellent fidelity. Peer observers who responded to the survey indicated that lessons they observed were delivered with high fidelity. Survey and interview data indicates that pupil pairing was carried out with very good fidelity.
- While overall the programme was delivered with high fidelity, a significant minority of teachers who responded to the survey reported adapting or considering adapting the intervention and this was echoed in the case study data. In most cases, these adaptations were small reductions in the time spent on lessons, where necessary, due to timetabling, or other minor amendments to address particular learning needs that did not significantly change the intended delivery of PALS-UK.

Enablers and barriers

This section examines, which aspects of PALS-UK may have facilitated or limited its effectiveness. It answers the following research question:

> 1.2. What are the enablers and barriers to success with respect to training and support and the individual components of PALS-UK?

We start by considering survey data specifically on the enablers of and barriers to implementation. The aspects of the programme that are then examined further are: staff training and ongoing support from the delivery team; materials and resources provided to deliver the intervention; pupil training; and the process of pupil pairing. These themes were identified based on deductive coding shaped by the research question and logic model. Data for this section comes from the endline surveys completed by teachers and headteachers in intervention schools, and teacher, headteacher, literacy lead, and pupil interviews.

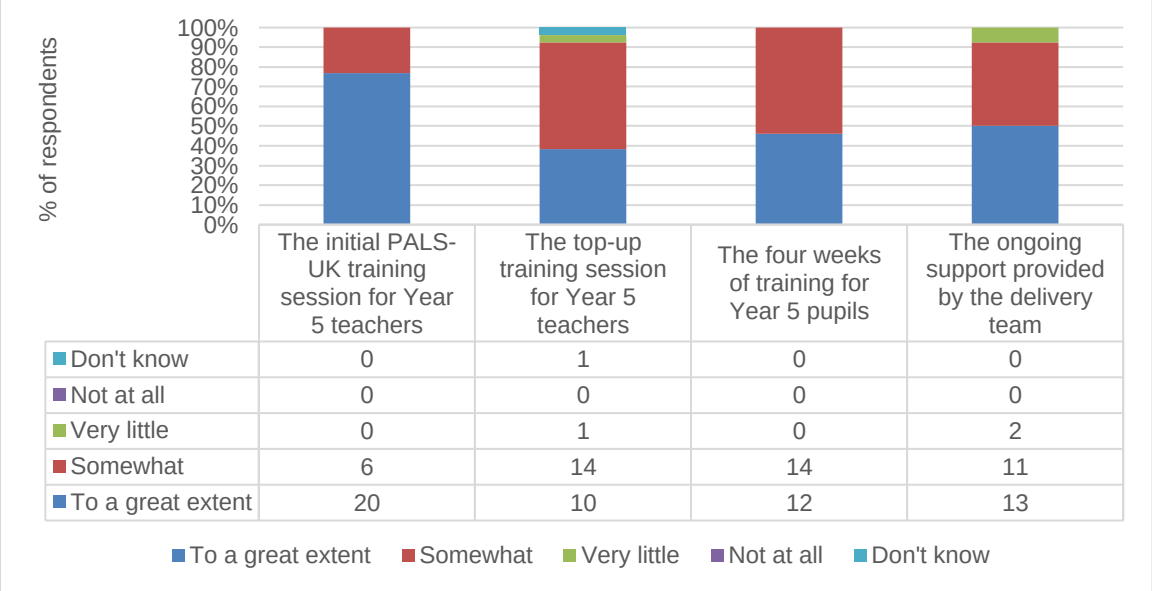
Enablers

Headteachers were asked to indicate the factors that they considered useful in facilitating effective implementation. **Figure 7a** describes the data relating to training and support, while **Figure 7b** focuses on resources, lesson observations, and other factors. As with the feedback from teachers, the training and support was very well received, particularly the initial training. The manual and selection of books were also considered to be enablers of implementation. The in-school peer observations were seen as a more powerful facilitator by a greater number of headteachers than the external observations, although a majority of them did find the external observations at least somewhat of a facilitator. Feedback on the video and online resources indicated that these were also perceived to be less powerful facilitators, however, again a large majority found these at least somewhat of a facilitator.

Teachers were also asked for their perspectives on the main factors that had enabled them to deliver PALS-UK. A combination of in-school factors and PALS-UK resources were identified. The most mentioned in-school factor was having protected time in the school timetable to deliver PALS-UK. Other factors mentioned by many teachers included support from colleagues such as phase leaders, senior colleagues, and teaching colleagues, and pupils' engagement and enthusiasm. PALS-UK resources were identified as giving teachers the knowledge and resources to successfully deliver PALS-UK. The initial training was the most frequently mentioned factor in survey responses due to its clarity, thoroughness, and time given to practice. Materials including the manual, the quality of contemporary books, and pupil prompt cards were also frequently mentioned. Further factors included support from the delivery team including the Microsoft Teams group, training videos about PALS-UK, and the highly structured nature of PALS-UK that established clear classroom routines and required little planning.

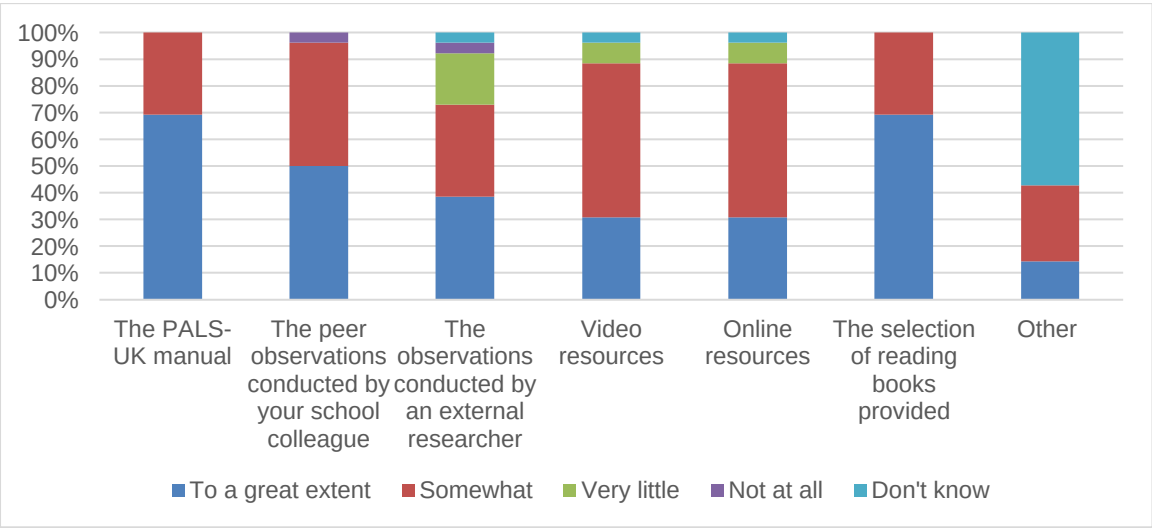
Figure 7: The facilitators of effective implementation relating to training and support, as reported by headteachers[40]

(a)



	The initial PALS-UK training session for Year 5 teachers	The top-up training session for Year 5 teachers	The four weeks of training for Year 5 pupils	The ongoing support provided by the delivery team
Don't know	0	1	0	0
Not at all	0	0	0	0
Very little	0	1	0	2
Somewhat	6	14	14	11
To a great extent	20	10	12	13

(b) Further facilitators of effective implementation reported by headteachers[41]



Barriers

We asked both teachers and headteachers about the perceived barriers to effective implementation of PALS-UK. Teachers' responses are shown in **Figure 8**. Pupil absenteeism was perceived to be the biggest barrier out of the choices that we gave them, followed by competing priorities. Unfortunately, we did not collect data on absenteeism from schools. A majority reported that limited capacity was not a problem or only a little problem (77%, n = 53), as was unrealistic expectations set by the PALS-UK programme (77%, n = 55).

[40] Headteachers' endline survey question: 'Q31. Please indicate the extent to which you consider the following factors to have been useful in facilitating the effective implementation of the PALS-UK programme'.

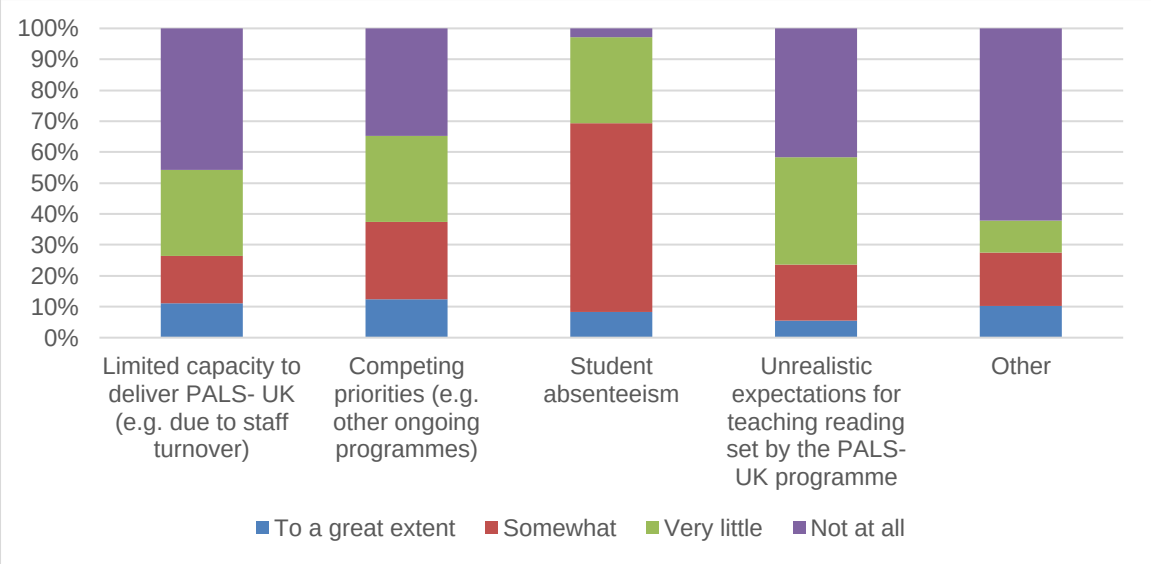
[42] Teachers' endline survey question: 'Q43. Please indicate the extent to which you consider the following factors to have been barriers or obstacles to the effective implementation of the PALS-UK programme'.

Teachers were asked to identify any 'other' factors. Most comments related to timetabling, including school events that led to timetable changes and more generally fitting the sessions into a crowded timetable. A typical comment on time pressures and competing demands was:

> *Given the amount of time PALS takes, it was often difficult to implement three sessions a week without it having a significant impact on other curriculum areas.* (T survey)

Declining pupil motivation and engagement, especially towards the end of the programme and the end of individual sessions, was a further barrier. Teachers also expressed some concern about the breadth of skills developed through PALS-UK and because of the length of PALS-UK lessons, the difficulty of then fitting in further reading lessons. Teachers' limited knowledge of the books, and pupil absence, and the need to form triads were also regarded as barriers to successful delivery.

Figure 8: The barriers to implementing PALS-UK reported by teachers[42]

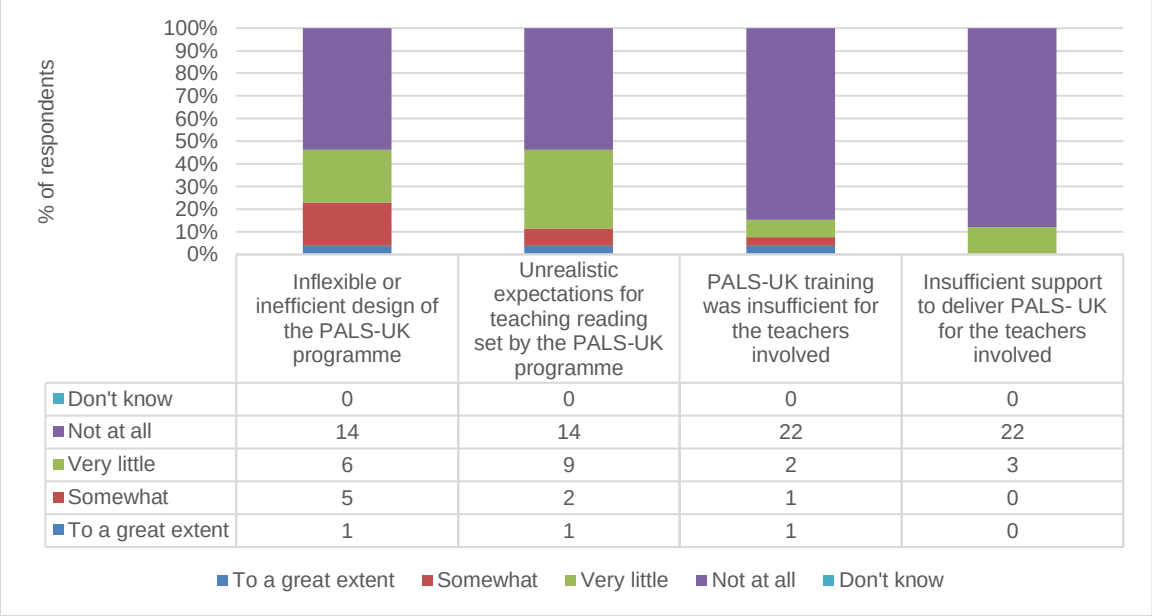


Almost a third of teachers responding to the endline survey (31%, n = 22) indicated that they faced any further challenges in delivery. Classroom management was the most common issue, in terms of keeping pupils motivated and challenged, particularly as the programme progressed and pupil engagement declined, especially among higher attaining readers. Another issue some respondents mentioned was the needs of lower attaining readers, for example: that pupils' cue cards could be overwhelming; that some pupils were less confident or less able to read out loud; that the books provided were too difficult for some readers; that phonics books did not work well with the PALS-UK activities; and that it was challenging to provide alternative activities for these pupils during PALS-UK lessons. Other comments related to resources and materials, particularly staffing issues resulting in PALS-UK being delivered by staff who had not been trained. One school mentioned lack of access to Microsoft Teams, which was used by the delivery team to share resources, deliver top-up training, and provide ongoing support. Two schools mentioned not having enough books in school to supplement the provided books. The final two issues reported were the time commitment for schools and its implications for other areas of the curriculum, and pupil absence.

Headteachers' perspectives on the barriers to implementation are shown in **Figure 9** and **Figure 10**. For respondents, the training was clearly not considered a barrier, nor was the support that teachers received or limited buy-in from teachers. A majority also considered that the design and expectations of PALS-UK alongside school capacity and competing priorities were either not a problem at all or posed very little of a problem, although the answers were more widely distributed on these questions. Two headteachers highlighted the time commitment of PALS-UK as an additional barrier, which had taken resources away from delivering other literacy lessons.

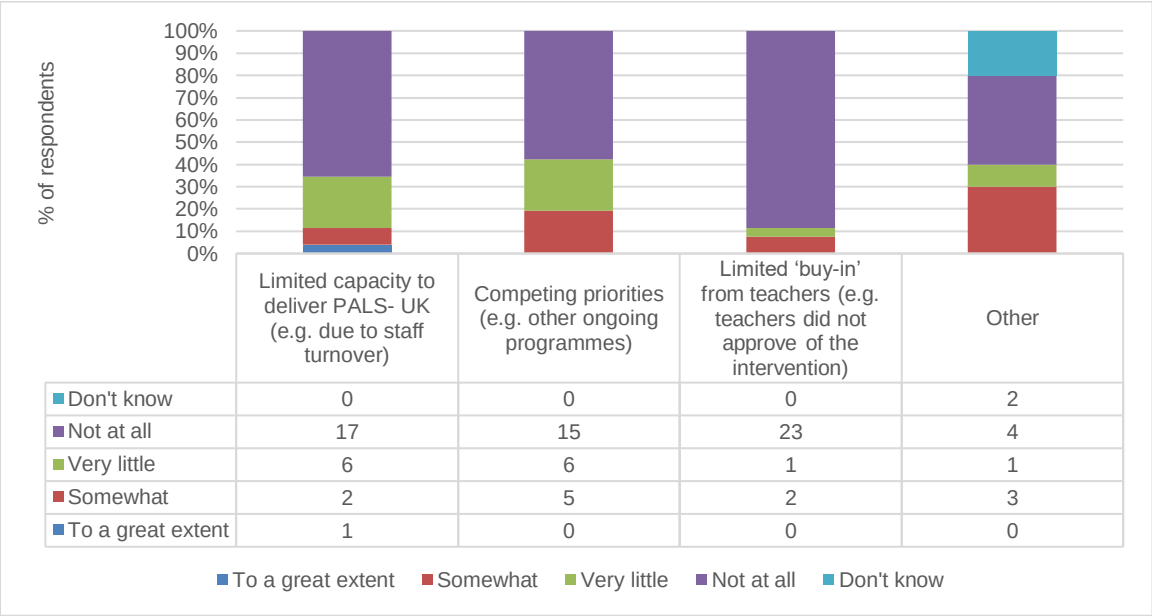
[42] Teachers' endline survey question: 'Q43. Please indicate the extent to which you consider the following factors to have been barriers or obstacles to the effective implementation of the PALS-UK programme'.

Figure 9: The barriers to successful implementation reported by headteachers[43]



	Inflexible or inefficient design of the PALS-UK programme	Unrealistic expectations for teaching reading set by the PALS-UK programme	PALS-UK training was insufficient for the teachers involved	Insufficient support to deliver PALS- UK for the teachers involved
Don't know	0	0	0	0
Not at all	14	14	22	22
Very little	6	9	2	3
Somewhat	5	2	1	0
To a great extent	1	1	1	0

Figure 10: The barriers to successful implementation reported by headteachers[44]



	Limited capacity to deliver PALS- UK (e.g. due to staff turnover)	Competing priorities (e.g. other ongoing programmes)	Limited 'buy-in' from teachers (e.g. teachers did not approve of the intervention)	Other
Don't know	0	0	0	2
Not at all	17	15	23	4
Very little	6	6	1	1
Somewhat	2	5	2	3
To a great extent	1	0	0	0

Time also emerged as a key barrier in the case studies, although teachers seemed to find ways to fit PALS-UK into the curriculum effectively overall. Staff from three schools mentioned that it was sometimes tricky to fit PALS-UK in when there were one-off events affecting school timetables (such as in-service training days, voting days, and assessment week). Staff from two schools mentioned that they found it hard to fit in the rest of the curriculum around PALS-UK. Staff from one school also mentioned that the timing of the activities within PALS-UK was constraining and did not leave enough space for teachers to respond to in the moment learning opportunities, in relation to exploring new vocabulary. A teacher from another school mentioned that she sometimes had to shorten the PALS-UK activities to allow time for moving between classrooms. Teachers from two schools talked about having PALS-UK lessons ring-fenced on the

[43] Headteachers endline survey question: 'Q30. Please indicate the extent to which you consider the following factors to have been barriers or obstacles to the effective implementation of the PALS-UK programme'.
[44] Headteachers endline survey question: 'Q30. Please indicate the extent to which you consider the following factors to have been barriers or obstacles to the effective implementation of the PALS-UK programme'.

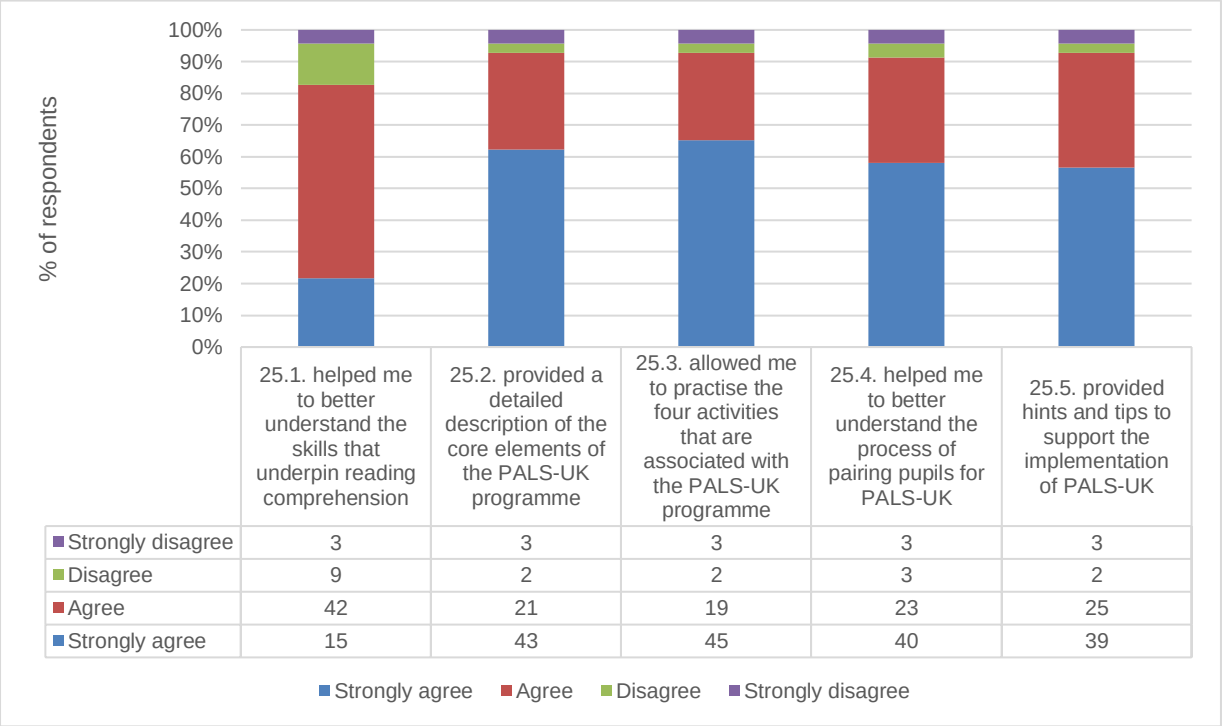
timetable within regular slots being an enabling factor. In addition to the theme of time, one case study teacher talked about the challenge of fitting PALS-UK in from a spatial perspective if you had a big class. In this school, they were able to spread out into another classroom, but the teacher suggested that conducting the intervention in just their classroom would have felt crowded and with all the pairs reading at once being distracting.

Training and support

Initial training

When asked about the initial training, the vast majority of teachers agreed or strongly agreed that the initial training had given them the knowledge and skills to deliver PALS-UK, although the general agreement was not so strong on whether the training had improved their understanding of reading comprehension skills (**Figure 11**).

Figure 11: The benefits of the initial training, as reported by teachers[45]



	25.1. helped me to better understand the skills that underpin reading comprehension	25.2. provided a detailed description of the core elements of the PALS-UK programme	25.3. allowed me to practise the four activities that are associated with the PALS-UK programme	25.4. helped me to better understand the process of pairing pupils for PALS-UK	25.5. provided hints and tips to support the implementation of PALS-UK
Strongly disagree	3	3	3	3	3
Disagree	9	2	2	3	2
Agree	42	21	19	23	25
Strongly agree	15	43	45	40	39

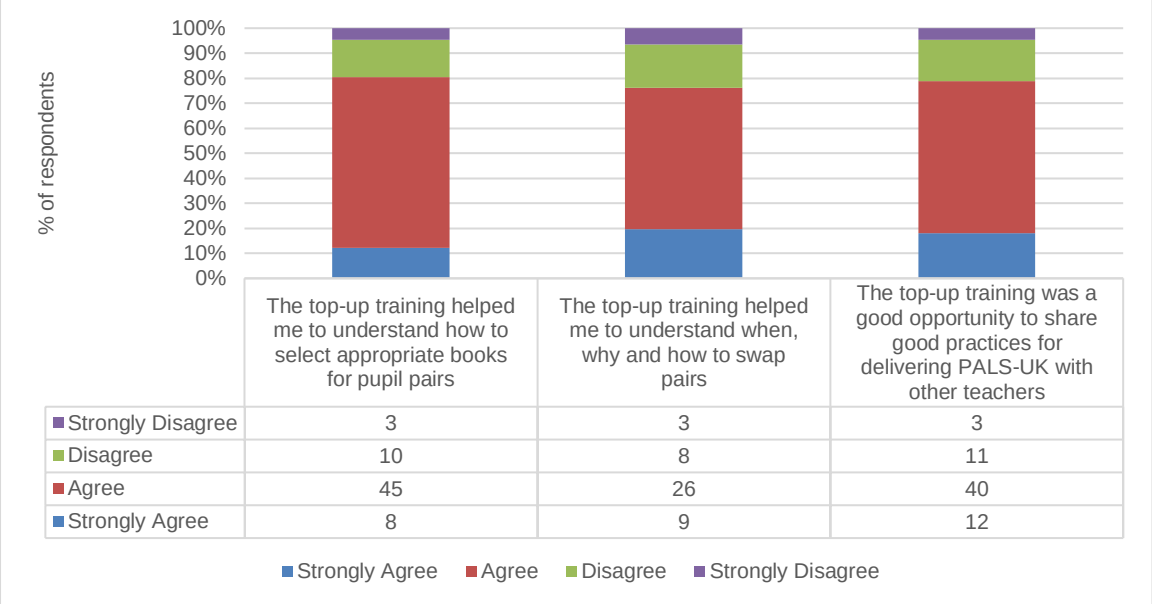
Interview data suggests that all teachers and other staff (headteachers/teaching assistants) who attended the initial training were overwhelmingly positive about its role as an enabling factor in the effective delivery of PALS-UK. The training was described as clearly explained and well structured, helping teachers to feel well-prepared for delivering the intervention. Five of the six case study schools mentioned that they found it particularly helpful being able to practice the activities as if they were a pupil, so that they got a strong sense of the skills that pupils would need to develop to carry out the activities. Teachers also valued the opportunity to discuss how the intervention might work in their particular context with the delivery team. Interestingly, staff from two of the schools mentioned that they especially liked the research focus of the training as they wanted to feel confident that they were taking on an intervention that already had research evidence behind it. One teacher mentioned that the training felt a little overwhelming at first but that once she took the knowledge back to school, she realised that she felt well-prepared. Another suggested that there could perhaps have been more of a focus on reassuring staff about how to manage potential issues with pairings in relation to pupils who may be less motivated, etc.

[45] Intervention teachers' endline survey question: 'Q25. Thinking about the initial full day training session, please specify the extent to which you agree with the following statements. The initial training…'.

Top-up training and further support

Feedback on the top-up training is similar (**Figure 12**). Overall, most respondents strongly agreed or agreed that the top-up training gave them a more detailed understanding of PALS-UK and the opportunity for peer support, although the numbers in disagreement were slightly higher than on the initial training question.

Figure 12: The benefits of the top-up training, as reported by teachers[46]



	The top-up training helped me to understand how to select appropriate books for pupil pairs	The top-up training helped me to understand when, why and how to swap pairs	The top-up training was a good opportunity to share good practices for delivering PALS-UK with other teachers
Strongly Disagree	3	3	3
Disagree	10	8	11
Agree	45	26	40
Strongly Agree	8	9	12

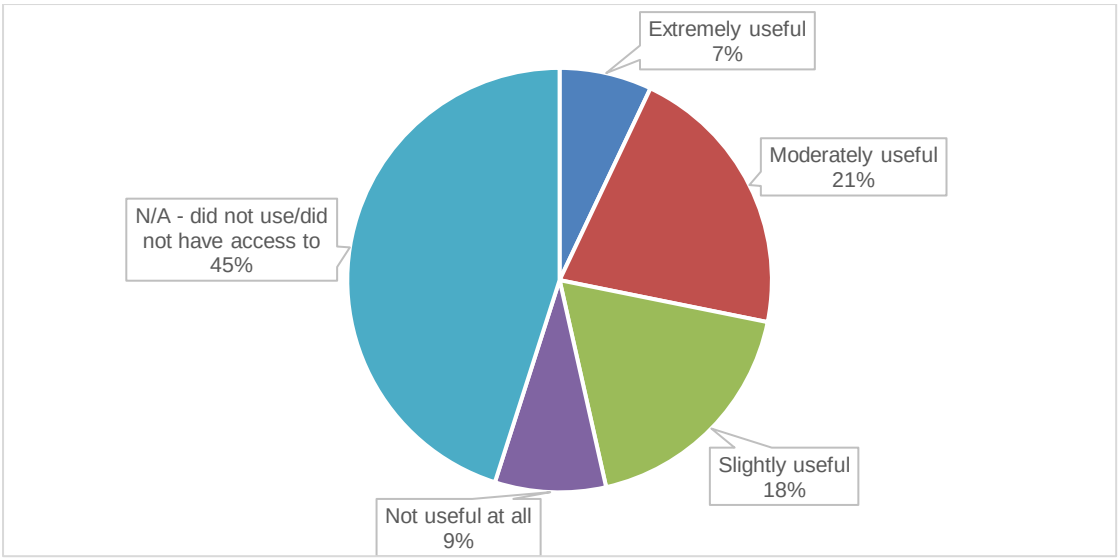
Similarly, the responses from the six case study teachers to the top-up training were more mixed than those on the initial training. Some teachers had not managed to attend. Those that did found it to be clear, but some felt that they were already confident about how to deliver PALS-UK by that point. One teacher mentioned that she liked being able to ask questions in the Microsoft Teams chat at the online top-up training as this provided opportunities to ask questions easily, as opposed to feeling that you might be interrupting the speaker in a face-to-face training session.

Twilight sessions

Teachers were also asked about the usefulness of the twilight sessions (**Figure 13**). Nearly half of the respondents did not attend a twilight session and roughly even numbers found it extremely/moderately useful compared to slightly/not at all useful.

Figure 13: The usefulness of the twilight sessions, as reported by teachers[47]

[46] Intervention teachers' endline survey question: 'Q27. Thinking about the online top-up training, please specify the extent to which you agree with the following statements'.
[47] Intervention teachers' endline survey question: 'Q31. In your view, how useful were the following resources for delivering the PALS-UK programme? The twilight sessions for sharing practice with other teachers'.



N/A, not applicable.

Eleven teachers responded to a question asking what benefits they got from the twilight session. Teachers found the session useful in two ways. First, as a recap of key points about PALS-UK delivery, clarification from the delivery team on any queries that had emerged through practice, for example, on areas that pupils were struggling with, and reassurance on fidelity. Second, in terms of peer learning, teachers found it helpful to informally share experiences of delivery with other schools.

We asked teachers how they thought that the training could be improved. There were a range of comments including: that the teachers could have been grouped by school demographic at initial training to allow for more peer learning; that the delivery team could run a training session in schools for pupils; and that the twilight session could have been more structured and participatory (e.g. with teachers asked to prepare a contribution). One teacher suggested more guidance on pupil pairing, there were two comments on the books being challenging, and one teacher suggested more input from the delivery team on selecting books at initial training.

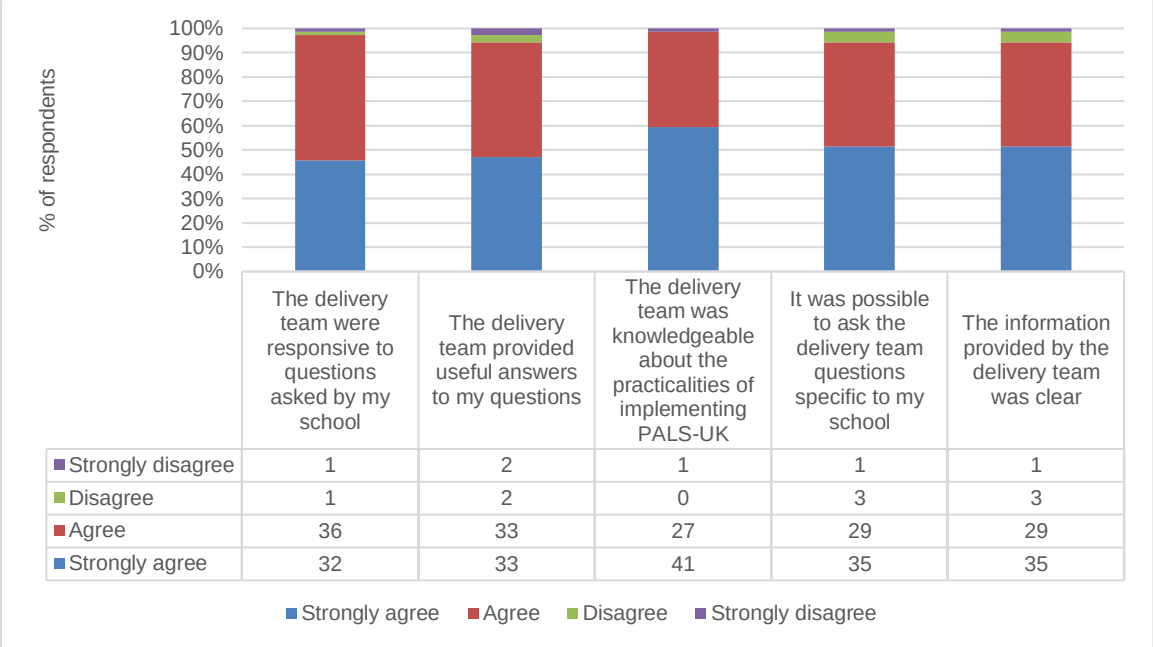
There were some comments on the timing of the top-up training. Two teachers thought that it would have been more useful later in the programme, and three thought that it was unnecessary or could be delivered on request. One teacher found the after-school timing of the twilight sessions difficult. Two thought the training would be better delivered in-person than online and one school had serious difficulties accessing Microsoft Teams.

Support

Teachers were asked about the support available to them from the delivery team. When asked about the usefulness of the 'just-in-time' support provided by the delivery team, while a significant proportion (40%) had not used it, the largest portion of respondents had and found it extremely or moderately useful (45%) and fewer (15%) found it slightly or not at all useful. We also asked for more detailed feedback on the responsiveness of the delivery team (**Figure 14**) and the appropriateness of communication channels (**Figure 15**). Most respondents agreed or strongly agreed that the delivery team provided relevant, clear, timely, and useful support. Email was the communication method that was preferred by most respondents to this question.

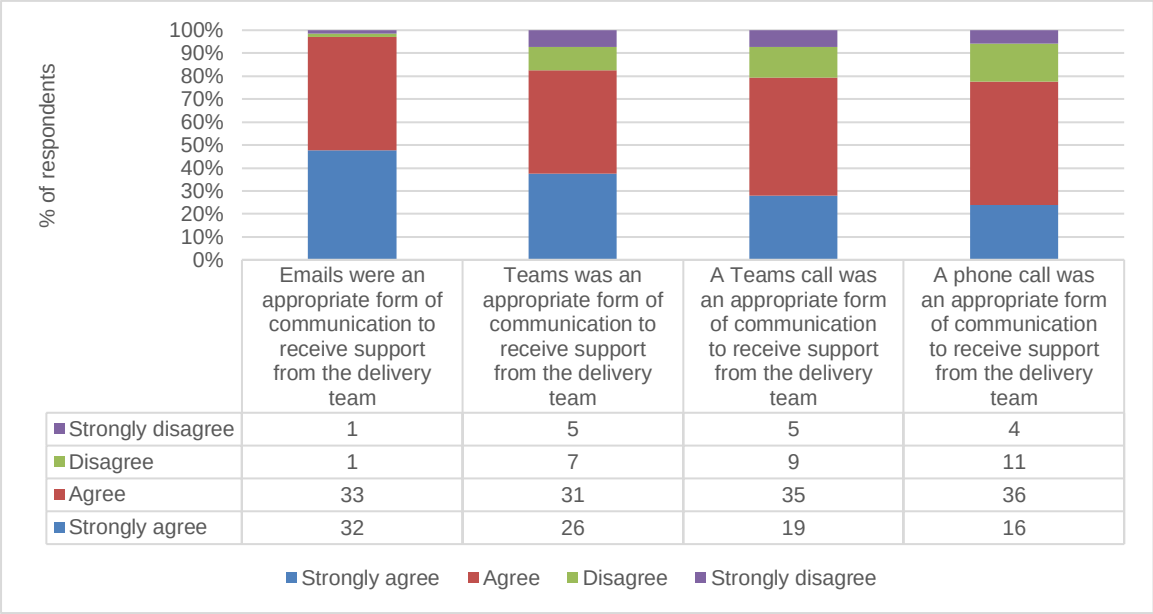
Figure 14: The responsiveness of the delivery team, as reported by teachers[48]

[48] Intervention teachers' endline survey question: 'Q32. Please indicate the extent to which you agree or disagree with the following statements about the support given by the PALS-UK team (Dr Emma Vardy, Dr Helen Breadmore, and Dr Luisa Tarczynski-Bowles and the research assistant staff that conducted the observations) to deliver PALS-UK'.



	The delivery team were responsive to questions asked by my school	The delivery team provided useful answers to my questions	The delivery team was knowledgeable about the practicalities of implementing PALS-UK	It was possible to ask the delivery team questions specific to my school	The information provided by the delivery team was clear
Strongly disagree	1	2	1	1	1
Disagree	1	2	0	3	3
Agree	36	33	27	29	29
Strongly agree	32	33	41	35	35

Figure 15: The appropriateness of communication channels for communication with the delivery team, as reported by teachers[49]



	Emails were an appropriate form of communication to receive support from the delivery team	Teams was an appropriate form of communication to receive support from the delivery team	A Teams call was an appropriate form of communication to receive support from the delivery team	A phone call was an appropriate form of communication to receive support from the delivery team
Strongly disagree	1	5	5	4
Disagree	1	7	9	11
Agree	33	31	35	36
Strongly agree	32	26	19	16

In relation to the ongoing support provided via Microsoft Teams and email, the case study teachers really valued the fact that the delivery team were approachable and were always on hand to respond promptly to questions. Some people had issues accessing or finding the materials on Microsoft Teams, but any issues were quickly resolved via email.

A minority of teachers (15%, n = 11) sought support to implement PALS-UK outside of the delivery team (i.e. within their school or with other schools/colleagues). These teachers did so mainly with other Year 5 teachers to confirm correct implementation as well as discuss, for example, reading materials and changing pairs, as well as senior colleagues to discuss timetabling. Two respondents had not attended training, indicating the importance of peer support in these cases.

Overall, the data suggests that the training, particularly the initial training, was very well received by teachers and was itself a facilitator of successful implementation. The twilight training, while moderately helpful to some attendees, was

[49] Intervention teachers' endline survey question: 'Q32. Please indicate the extent to which you agree or disagree with the following statements about the support given by the PALS-UK team (Dr Emma Vardy, Dr Helen Breadmore, and Dr Luisa Tarczynski-Bowles and the research assistant staff that conducted the observations) to deliver PALS-UK'.

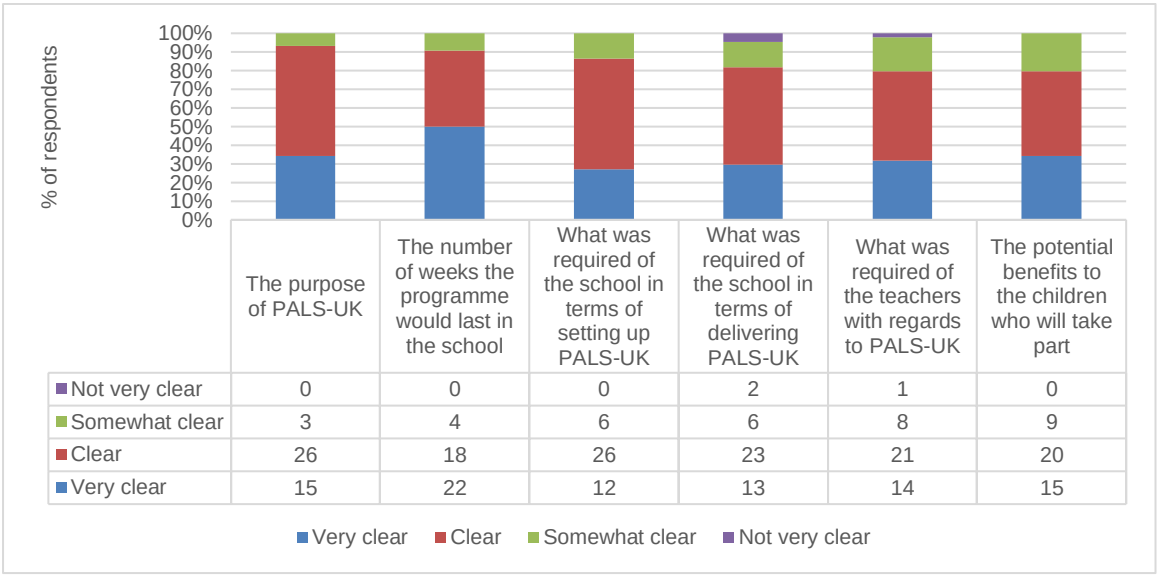
perceived as less important. The support provided by the delivery team and the methods of delivering that support were similarly helpful to schools delivering PALS-UK.

Programme implementation

Initial information about PALS-UK

At baseline, headteachers at all schools were asked about the clarity of information they had received before the start of PALS-UK (**Figure 16**). Across all measures, the headteachers that responded felt that they had been provided with clear information. In addition, most headteachers (96%, n = 22) felt that their school was appropriately prepared for implementing PALS-UK by the delivery team.

Figure 16: The clarity of information provided about PALS-UK, as reported by headteachers at baseline[50]



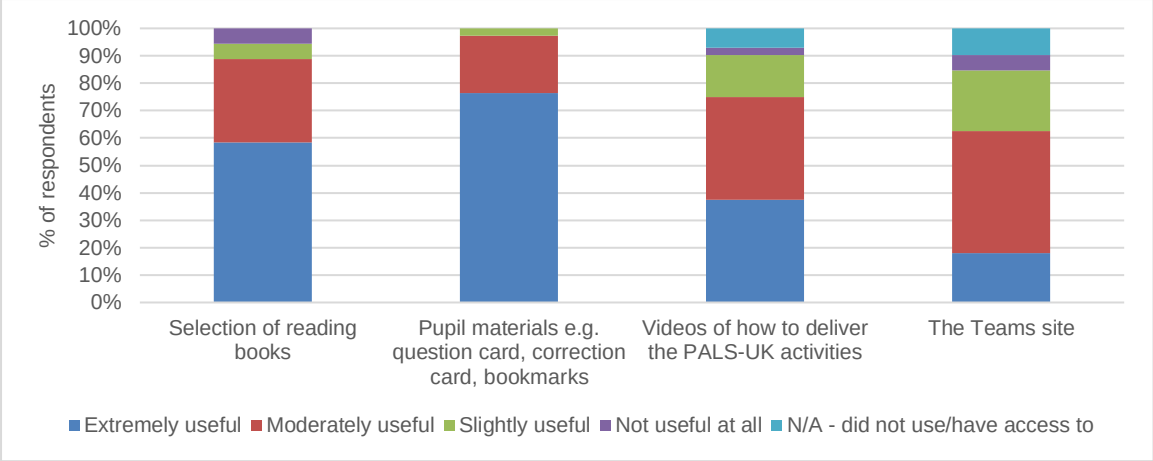
	The purpose of PALS-UK	The number of weeks the programme would last in the school	What was required of the school in terms of setting up PALS-UK	What was required of the school in terms of delivering PALS-UK	What was required of the teachers with regards to PALS-UK	The potential benefits to the children who will take part
Not very clear	0	0	0	2	1	0
Somewhat clear	3	4	6	6	8	9
Clear	26	18	26	23	21	20
Very clear	15	22	12	13	14	15

PALS-UK materials and resources

Teachers were asked about the resources provided to support their delivery of PALS-UK (**Figure 17** and **Figure 18**). **Figure 17** suggests that the pupil materials and the reading books were regarded as particularly useful. Feedback on the videos and Microsoft Teams site was more mixed, although in both cases most respondents still agreed that they were at least moderately useful.

Figure 17: The usefulness of PALS-UK resources, as reported by teachers[51]

[50] Headteachers' baseline survey question: 'Q25. Based on the information your school has received before the start of PALS-UK, please indicate how clear the following details are to you'.
[51] Intervention teachers' endline survey question: 'Q31. In your view, how useful were the following resources for delivering the PALS-UK programme?'

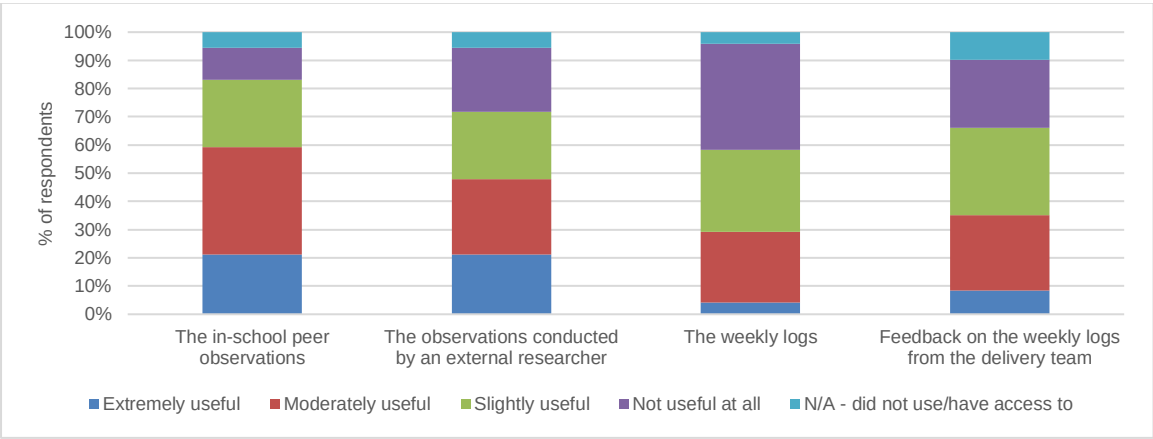


Staff from all six case study schools spoke very positively about the books, commenting on the excellent range for pupils to choose from and reporting that the pupils seemed to really enjoy reading them. It was noted, however, by staff from four schools that sometimes these needed to be complemented with other books for lower attaining readers or pupils new to English. Staff from two schools mentioned that while the books provided with PALS-UK were a fantastic resource, it was also important for the teachers delivering the intervention to have a good knowledge of them. One teacher felt this was important so that they could ensure that each pair reads a book, which is both at the right level and is likely to be of interest to them. Another noted that having a good knowledge of the books enables the teacher to monitor the success with which pupils are completing the activities (i.e. it is hard to know if they are predicting accurately if you are not familiar with the book they are reading). At four of the case study schools, most of the pupils commented that there was a good variety of books to choose from. This included titles they had not come across before, which they enjoyed because '*you don't really know what's going to happen in them*' (P interview). These pupils talked about the diversity of characters and themes in the choice of books as well as the different lengths and levels of difficulty, and they named particular books that they had enjoyed. The variety enabled them to choose books that they thought they would like or were motivated to try. Four teachers reporting feedback from their pupils also noted that they enjoyed the new books that PALS-UK introduced to them.

One case study teacher described how the prompt cards for pupils were useful as they helped the pupils to remember their role as coach.

Figure 18 focuses on the observations and weekly logs. Most teachers found the in-school peer observations at least moderately useful, while they did not find the observations conducted by an external researcher as useful. Feedback on the weekly logs was very mixed, with 38% finding them not useful at all. However, respondents found the feedback on these logs slightly more useful.

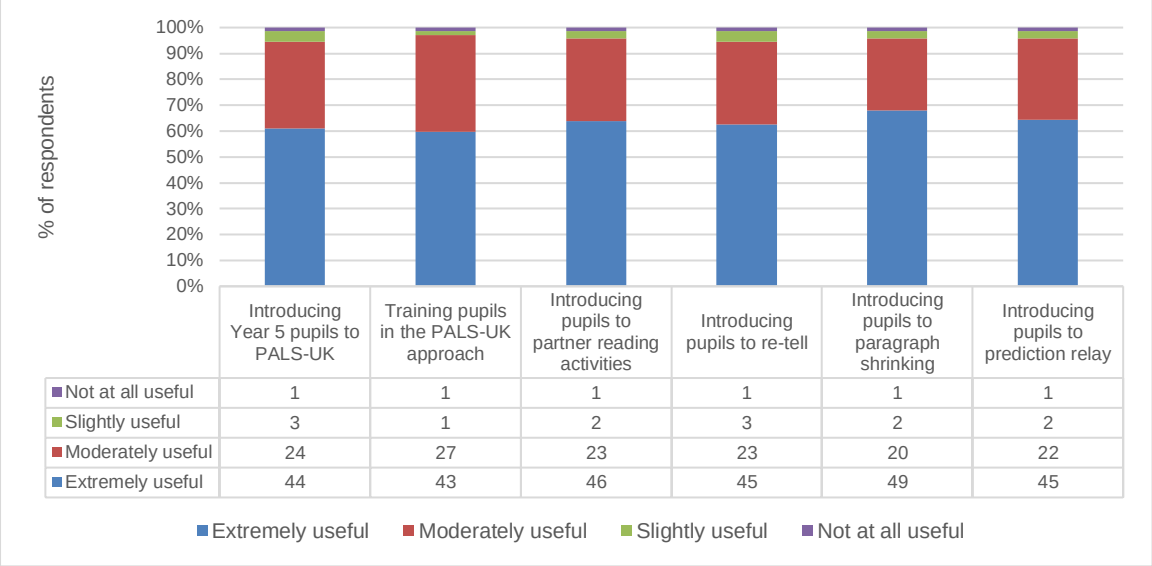
Figure 18: The usefulness of PALS-UK lesson observations and weekly logs, as reported by teachers[52]



[52] Intervention teachers' endline survey question: 'Q31. In your view, how useful were the following resources for delivering the PALS-UK programme?'

Teachers found the manual was very useful in supporting them to deliver PALS-UK with their classes (**Figure 19**).

Figure 19: The usefulness of the PALS-UK manual, as reported by teachers[53]



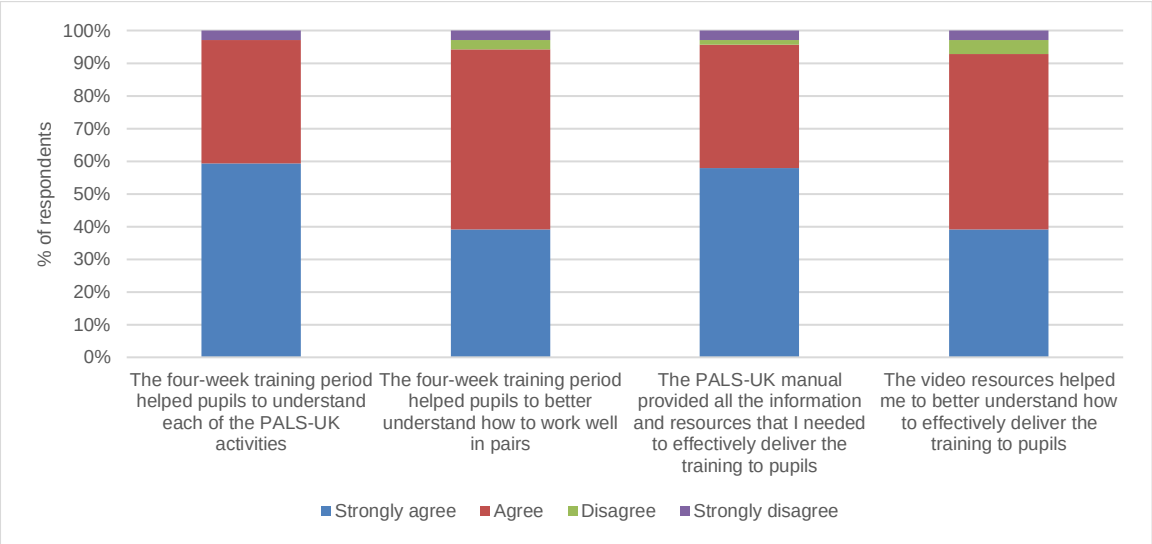
	Introducing Year 5 pupils to PALS-UK	Training pupils in the PALS-UK approach	Introducing pupils to partner reading activities	Introducing pupils to re-tell	Introducing pupils to paragraph shrinking	Introducing pupils to prediction relay
Not at all useful	1	1	1	1	1	1
Slightly useful	3	1	2	3	2	2
Moderately useful	24	27	23	23	20	22
Extremely useful	44	43	46	45	49	45

Feedback from case study teachers about the PALS-UK manual was also very positive. It was described as *'fool proof'* (T interview), and easy to navigate in the moment. One teacher valued the balance between a script being provided to support consistency, alongside *'tips'* (T interview), which could be used to make minor adaptations.

PALS-UK pupil training and mini lessons

The majority of teachers were positive about the four-week training period for pupils (**Figure 20**). When asked for feedback on how the training could be improved, four teachers commented that the period for pupil training could be condensed into fewer sessions.

Figure 20: Views on the four-week training for pupils, as reported by teachers[54]



The case study teachers indicated that pupil training was straightforward although pupils found some of the PALS-UK activities harder than others (see section 'PALS-UK lesson structure and activities' below). One teacher suggested that the pupil training was effective because of its staged nature with multiple opportunities for pupils to practise each activity.

[53] Intervention teachers' endline survey question: 'Q30. We would like to know your views on the usefulness of the PALS-UK manual. Please indicate how useful (if at all) the PALS-UK manual was for the following aspects of programme delivery'.
[54] Endline teachers' survey question: 'Q26. We would like to know more about your views on the four-week training for pupils as part of PALS-UK. Please indicate the extent to which you disagree or agree with the following statements'.

The use of mini lessons varied between case study schools. Some schools said they found them helpful to use after a half-term break; others said they had not needed to use them.

Pupil pairing process

The fidelity section of this report has shown that this was an aspect of PALS-UK delivered with high fidelity, and here we give teachers' perceptions on the process.

Many teachers found changing the pairs a positive and useful process. Some teachers reported that they made extra changes to pairs to manage classroom behaviour, where pairs were not getting along, to improve the coaching relationship and level of support, to accommodate new pupils or pupil absences, or where attainment level or specific needs were having implications for the partnership. Teachers mentioned that changing the pairs kept PALS-UK '*fresh*' (T interview) and often led to an improvement in pupil behaviour:

> *It helped to re-focus children, keep them engaged and helped with some of their enjoyment of PALS.* (T survey)

It was also an opportunity to change pairs that were not working so well together:

> *Those who didn't always have an enthusiastic partner got an opportunity to experience it with others.* (T survey)

Changing the pairs also allowed the pupils to work with a range of peers and access different types of support. Teachers mentioned the importance of using their knowledge of their class alongside New PiRA baseline results to inform the pairings:

> *Knowing my pupils helped me adapt some pairings to maximise the effectiveness of the reading collaboration.* (T survey)

Teachers also reported some challenges and barriers to pairing and changing the pairs. Large attainment gaps posed a challenge with pairing, for example, with finding a book that suited both partners, having lower attaining readers struggling in roles as the first partner, or frustration among pupils at being paired with pupils with the lowest reading attainment, or with a lack of challenge for higher attaining readers. Teachers also noted that forming effective pairings for pupils with SEND and EAL was a challenge.

> *We also had to accept that in situations like this, reader 1 didn't get much out of that pairing, because the reader 2 wasn't able to spot any of reader 1s mistakes when reading.* (T survey)

> *Difficult to pair the children who were reading below key stage—they needed very reliable and patient partners.* (T survey)

Managing the behaviour between pairs and maintaining the coaching relationship when the pair were not getting on was a challenge. This meant that taking into account pupils' friendships, relationships, and behaviours was important, which was more difficult for teachers new to the class who did not have that knowledge. Teachers in small classes and mixed year group classes found it increasingly difficult to form balanced pairings as the intervention progressed—although even in larger classes the difficulties in forming new and effective pairings increased with time, particularly for pupils with more challenging behaviour or lower reading attainment. A small number of teachers reported pairing pupils with similar abilities due to large differentials in attainment. Finally, dealing with absence was a challenge. There was also more general feedback from teachers that pairs were rarely able to finish their books due to pairs changing, leading to some frustration among pupils. The main challenges with pairing related then to ability, behavioural needs, and absences.

Data from the case study schools gives us further insight into school experiences of pairing. The pairing process was considered to be central to the intervention's potential success. Staff at two schools noted that the method for initial pairing of pupils (rank ordering the class in terms of reading attainment, splitting the class in half. and pairing the reader at the top of the top half with the reader at the top of the bottom half, etc.) was broadly effective as it did not result in too big an attainment gap between first and second readers. However, staff at one school noted that this was challenging in their school were there was a very broad attainment range. For them it was hard to create pairings where *'no one gets a bad deal out of it'* (T interview). As well as consideration of the gap in attainment between the two readers, staff at this school also talked about the difficulties associated with ensuring that both pupils' accessibility needs would be met, while maintaining motivation for both readers (e.g. where one reader might require larger print/dyslexia-friendly

texts). Staff from one school reported that they paired a child who was new to English with the teaching assistant and that this adaptation was very effective, allowing the pupil to benefit from the volume of reading and structured activities associated with PALS-UK, while reading books that were suited to their current level of English proficiency (provided by the school to supplement those provided with PALS-UK).

All schools reported that a broad range of factors needed to be considered when pairing pupils in addition to their reading attainment/English language proficiency levels. These included personality, relationship dynamics, SEND status, patience, and levels of motivation. Staff at two schools noted that balancing these considerations became increasingly difficult as PALS-UK progressed. However, staff at one school commented that the changes in pairings were needed to *'freshen things up'* (T interview) and all schools reported that they found some way to manage the pairing process in a way that felt satisfactory overall. Interestingly, while one teacher talked about the importance of changing a particular pairing if the dynamic is not working, another teacher purposely stuck with an initially problematic pairing (from an interpersonal perspective) and took this as an opportunity to support the pupils in developing their social skills.

Overall, while pairing pupils is an essential activity in the PALS-UK logic model and was implemented with high fidelity, the process posed some challenges to teachers in creating effective pairings, more so as the programme progressed.

Thus, key enablers of successful implementation were PALS-UK resources, particularly the initial training, manual, and books, alongside school support, particularly in the form of timetabled space for the delivery of PALS-UK. The barriers included contextual factors, particularly pupil absenteeism and competing priorities on class time, and there were some reports of challenges in maintaining pupils' motivation and engagement with the programme as it progressed.

Summary of key findings

- Teachers perceived that the training and support provided by the delivery team was extremely helpful to the delivery of the programme, although the twilight training may be an area where improvements could be made.
- Teachers and headteachers perceived that their schools were well informed about the intervention and well supported in the initial stages, although there were a small number that perceived the four-week pupil training was too long.
- While pupil pairing was seen as an essential component of the programme, teachers did identify barriers to creating effective pairings that became more challenging as the programme progressed.
- Pupil absenteeism and competing priorities were identified as particular barriers to implementation, while the initial training, manual, and books, alongside timetabled delivery were regarded as enablers.

Stakeholder experiences and perceived impact

This section explores what teachers, headteachers, and pupils thought about the PALS-UK activities and the extent to which they felt they had an impact on reading comprehension, fluency, vocabulary, and overall reading attainment. We also report on pupils' and teachers' thoughts on any effects of PALS-UK on pupils' feelings about reading, including self-efficacy, motivation, and identity in relation to reading. We also explore teachers' perceived impacts of PALS-UK on their knowledge and confidence in relation to teaching and reading, their reflections on how they experienced the shift towards a peer-led pedagogy, and the extent to which they perceive PALS-UK to be accessible/effective for all pupils. Pupils at case study schools were asked to reflect on their experiences of PALS-UK either directly (during our visits to the six case study schools) or by providing their feedback to their teacher prior to the second case study teacher interview.

The data for this section examines how PALS-UK was experienced by teachers, headteachers, and pupils in the research question:

1.3. How do different stakeholder groups experience PALS-UK and what is the perceived impact?

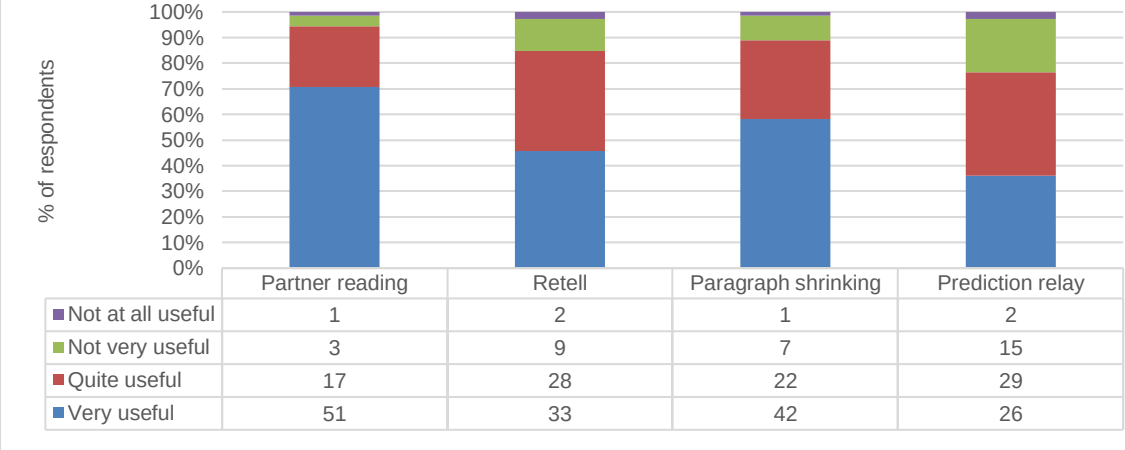
The data for this section comes from the teacher and headteacher surveys and the teacher, headteacher, literacy lead, and pupil interviews from the case studies.

PALS-UK lesson structure and activities

Case study teachers were asked to describe, which aspects of the PALS-UK lessons might be effective in promoting reading attainment and, which aspects might be less effective/present challenges. One of the enabling characteristics of PALS-UK lessons, which most teachers spoke about was the fact that it has a clear and consistent structure. Teachers reported that pupils liked '*knowing what was coming*' (T interview) and that the predictable routine helped keep them on task, especially when schools were able to timetable PALS-UK for the same time of day each week. One pupil commented to their teacher that the structure of PALS-UK helped them to keep focused on the book. Staff at three of the six schools also mentioned how the repetitive structure of the lessons and the ready-made resources helped to reduce teacher workload. On the other hand, staff at four schools suggested that the repetitive nature did lead to some pupils becoming bored with it towards the end of the intervention, with one school suggesting that they might use the intervention in the future but over a shorter period of 12 to 14 weeks. Staff at two of the schools mentioned that the use of a timer helped pupils to stay on track, although one teacher reported not using a timer that was visible as it was making one of the pupils anxious.

Across all of the activities, at least 75% of teachers found the four PALS-UK activities quite useful or very useful (**Figure 21**). Partner reading was regarded as the most useful, with 95% of teachers agreeing that it was quite or very useful. This was followed by paragraph shrinking (89%), retell (85%), and then prediction relay (76%).

Figure 21: Usefulness of the four PALS-UK activities reported by teachers[55]



	Partner reading	Retell	Paragraph shrinking	Prediction relay
Not at all useful	1	2	1	2
Not very useful	3	9	7	15
Quite useful	17	28	22	29
Very useful	51	33	42	26

Teachers were asked to explain their responses to the question of how useful the four activities were. A few commented that all the activities worked well together to enable pupils to practice and develop reading skills and achieve positive outcomes for reading.

Partner reading

Teachers reported that partner reading was popular among their pupils, leading to high levels of engagement. The safe environment of reading to a peer was mentioned:

> *The partner reading was good as children felt more confident and comfortable in reading in smaller groups.* (T survey)

[55] Teachers' endline survey question: 'Q40. We are interested in your experience delivering the four different activities—partner reading, re-tell, paragraph shrinking, and prediction relay. How useful did you find each activity in terms of developing your class's reading skills?'

Teachers liked the peer support aspects, and noticed improvements in comprehension, confidence, stamina, and fluency, especially for lower attaining readers.

> *The ability to have a substantial amount of solid reading time has been hugely beneficial as many don't read for any period of time. It also helped them to be more familiar, and less daunted, by longer texts.* (T survey)

While some teachers said that the modelling aspect of partner reading helped to increase pupils' understanding of the text, it had the potential for frustration from higher attaining readers and in general this activity was felt to have less impact on confident readers. And while teachers valued the focus on listening skills, they noticed that not all mistakes were caught by peers or correctly explained.

The case study teachers spoke of how the partner reading activity provided pupils with regular opportunities to read long chunks of text, thus developing their stamina. They discussed the value in the second reader having access to a good role model in terms of '*expression and tone and fluency*' (T interview)—although, one teacher noted that the first reader may not always model good use of expression and that pupils still need access to role modelling from adult readers. The process of requiring pupils to check for mistakes was valued by teachers; however, one teacher reported that this was '*done with varying levels of success*' (T interview). This teacher noted that pupils '*don't always give their partners the time to actually check it for themselves*' (T interview). Another teacher from a different school raised the issue of pupils '*zoning out*' (T interview) if their partner did not make many mistakes, especially as they got further into the intervention. They suggested that perhaps in these cases, pupils should be encouraged to focus on expression/prosody rather than just accuracy of reading. At one school, the teacher spoke of how for some second readers, reading the same text that had been read by the first reader enabled them to feel safe and boosted their confidence, but other second readers found it demotivating because they knew what was coming. A teacher from another school mentioned a proficient reader who did not like the partner reading aspect, because she did not like reading aloud and saw reading as *'an introspective thing you did in your head'* (T interview).

Retell

Teachers identified retell as a more challenging part of PALS-UK sessions. Some pupils were unable to fill the time, one partner was less engaged, and more teacher prompting was required to elicit the required level of detail. These challenges were more pronounced for lower attaining readers whose focus was often on decoding an often challenging book. Those teachers who valued retell appreciated the support that it gave second readers to be able to give detailed consideration to smaller sections of the text and consider what they had understood, and the opportunity that it gave the first reader to support the second.

One case study teacher felt that the retell activity was appropriately *'snappy'* (T interview) after pupils had just engaged in a relatively long chunk of reading. It was also suggested by staff at two schools that the retell activity supports the skill of retrieval, by helping pupils to focus on the details of what they have read in short chunks. One teacher described the value of only second readers being required to retell:

> *…the first readers will be able to follow the books, but the second readers might not be able to. I think what is important is that they are able to pick out those major points in the story and major characters, which a lot of children don't have that skill.* (T interview)

In terms of barriers relating to the retell activity, the case study teachers spoke of pupils initially finding it difficult to remember what they had read because they had been focusing on decoding and that they sometimes found it hard to focus on the level of granular detail required to retell the key details from a small chunk of text. They described pupils getting better at this over time and staff at one school mentioned the (variable) coaching ability of the first reader as a key enabling factor in supporting the second reader to focus on the granular detail. These same staff also suggested that for the weaker readers two minutes might not be long enough for them to quickly summarise all the text that they have just read, which can be a sizeable amount, given that weaker readers may read quickly but struggle more with the comprehension side of things.

Five pupils from three schools felt that the first reader should also do retell. One teacher from a fourth school also said that her pupils felt the same and that she had explained to them why it was this way.

Paragraph shrinking

Within the survey, data paragraph shrinking was similarly valued because it supported pupils to explore the text in detail:

> *I loved this activity and will use it to teach summarising in the future. It was a clear sequence of identifying the main idea and then saying it in 10 words or less. Pupils really understood this idea and it made summarising clear.* (T survey)

Like retell, this activity gave teachers an understanding of how much pupils had understood in their texts. Difficulties included attempting to summarise too much of the text, or not having enough information in one paragraph to summarise, however, some respondents said that the activity became easier over time. One teacher commented that some books were not appropriate for paragraph shrinking, for example, because they were too short.

One of the key enabling aspects of paragraph shrinking reported by the case study teachers was their pupils' enjoyment of it. Staff at one school described this as '*by far their favourite bit*' (T interview), reporting that pupils get excited about the challenge of summarising the paragraph in ten words, which in turn get them to '*think about language*' (T interview). While the teachers described this task as challenging, staff at four schools described this challenge positively as it pushed pupils to develop the skills of inference and summarising. Barriers to successful engagement with this task included pupils struggling:

> *…to identify what a paragraph is, or they know what a paragraph is but they just don't stop reading at the end of the paragraph; they want to read more*. (T interview)

One teacher mentioned that the wording for this activity was '*a little bit cumbersome'* (T interview) and that this activity did not '*get the depth from the children that the other bits did*'(HT interview). This was described as being due to the fact that sometimes '*not a lot has happened in that paragraph*' (T interview) and/or there is some ambiguity in terms of what the '*most important who or what*' (T interview) is within that paragraph. A teacher from another school described paragraph shrinking as having the '*highest cognitive load*' (T interview) and noted that it can be tricky to monitor pupils' accuracy on this activity while they circulate as they are all reading different books. They expressed concerns that while the pupils had become skilled at getting down to ten words, they were not always '*finding the most important bit*' (T interview).

Prediction relay

Within the survey responses, prediction relay was also identified as a difficult task for pupils:

> *Lots cannot predict the events in the next half a page and either predict an event which would occur much further into the story, or a 'silly' or an unlikely event. Other children are not always able to suggest that this would not be appropriate.* (T survey)

Teachers fed back that predicting every half page was too often. In addition, there could have been more discussion of the evidence for predictions and their veracity, as well as prediction based on authorial elements or character intent to stretch learning. The nature of the text seemed to have an impact on how manageable this task was for pupils. However, it was valued for supporting pupils to specify their predictions and ground them in their understanding of the text:

> *Prediction relay helped us refine the skill of predicting in a very useful way as it is often not fully successful, being too generic or vague.* (T survey)

Staff at the case study schools felt that prediction relay was a useful activity overall. Staff at three schools suggested that this activity was helpful in encouraging pupils to think more carefully about their predictions, making them less '*Iwild*' (T interview) and more '*rooted in the text*' (T interview). Staff at one school suggested that in order to make sensible predictions, pupils needed to develop skills of inference/deduction, which were deemed to be important aspects of reading comprehension. Another benefit of the activity identified by staff at one school was that it gave some of the weaker readers, who might struggle with the decoding side of things, a 'real *moment for them to shine"*. In terms of barriers, one school described it as a '*massive cognitive task to try and unpick*' (T interview), which involves pupils developing the ability to '*nail it down into just the next bit*' (T interview). Staff from another school noted that even adults find it difficult as when predicting '*we make such long arches in time and space*' (T interview). Staff from two schools suggested that five minutes might be too long for this activity: one reported that they sometimes reduced it to three minutes; another wondered if it might have been better to have replaced some of the time with some vocabulary input instead, for example: '*find an interesting word and figure out what it means*' (T interview). One final barrier that was

identified by staff at one school was the fact that it is difficult to monitor if pupils have made a feasible prediction when they are all reading different books and that it would perhaps be better to use a more manageable number of books.

PALS-UK and Key Stage 2 national assessments

Aspects of reading not covered by PALS-UK, including preparation for Key Stage 2 national assessments

More than half of the teachers (59%, n = 42) felt there were aspects of reading skills that were not covered by PALS-UK. The most common reason given for this was vocabulary, with some teachers concerned that the lack of time to discuss vocabulary was restricting pupils' learning of new words and the skills of using context to understand new words. Inference skills was also widely mentioned by teachers as well as retrieval. There was some concern that PALS-UK did not cover all the skills required for Key Stage 2 national assessments, including these vocabulary, inference, and retrieval plus understanding and answering questions, and providing written evidence. Further areas that teachers mentioned included authorial intent, non-fiction texts, comparison and contrast, structure and sequencing, language and tone, reading shorter texts, reading at speed, modelling fluency, and reactions to the text.

Feedback from a school that withdrew from the trial partway through raised similar concerns that PALS-UK did not cover all of the necessary reading skills, particularly inference, and that it did not prepare pupils for answering test questions.

Survey data suggests that schools responded in different ways to these perceived gaps. Some did additional reading lessons or interventions that focused on specific areas or incorporated these skills into reading or English lessons that were running alongside PALS-UK. Two teachers did mention that it was difficult to find the time to teach these skills alongside PALS-UK. One said that towards the end of PALS-UK they introduced a way of recording new vocabulary between partners.

Within the case study interviews, teachers and headteachers spoke specifically about the gaps that they perceived PALS-UK to have in relation to preparation for the Key Stage 2 national assessments. Teachers from all six case study schools perceived that PALS-UK does not prepare pupils for responding to 'SAT-style questions' in written form. For example, one teacher noted:

> *It is quite far removed from a SATs paper or a test paper, I think you do, as a teacher…you have to make sure that they are still seeing those and learning to apply it in a written context.* (T interview)

In particular, teachers talked about pupils needing opportunities to practice explaining/elaborating on their points and providing evidence, which they felt PALS-UK did not provide any space for. However, staff at one school noted '*that comes so much after the verbal anyway, if they've got the verbal intent then the written will come*' (HT interview) and that the verbal approach '*appeals to boys who don't like to write*' (HT interview). Staff from four of the six schools felt that inference, which is one of the content domains assessed within the Key Stage 2 reading test, is not explicitly addressed within PALS-UK. While it was noted that some of the PALS-UK activities might develop inference implicitly (e.g. prediction relay), staff felt that there was a need for this to be addressed more directly through discussion and teacher modelling. Similarly, teachers from three schools talked about vocabulary not being explicitly addressed, and a teacher from another school felt that impact and structure (two further content domains covered within the national assessments) are not addressed at all by PALS-UK. Another concern raised in relation to preparation for national assessments was the fact that the programme does not allow the teacher to tailor their teaching to particular domains that pupils are weaker in (instead, all pupils do the same thing). Despite these concerns, overall, the six case study teachers were positive about PALS-UK but reported needing to address the gaps outlined above in literacy/English sessions outside of PALS-UK sessions. Two teachers noted that while PALS-UK worked well within Year 5, they felt that it would not be appropriate for Year 6. Staff at one school suggested that PALS-UK could be improved further by the inclusion of an activity to develop inference and/or vocabulary. It is important to note that teachers comments related to what they perceived PALS-UK to cover/not cover in relation to the overall aims of the reading curriculum in England (i.e. working towards pupils being successful across each of the reading content domains, which are assessed in the end of Key Stage 2 national assessments); however, PALS-UK is not designed to address all these domains, just those specified within the logic model.

Elements of PALS-UK, which support preparation for national assessments

The case study teachers also spoke specifically about the ways in which they felt PALS-UK might help prepare pupils for the national assessments in Year 6. One of the advantages identified was that PALS-UK supports pupils' confidence

in dealing with longer texts, which staff from three schools suggested might mean that they will be less thrown by the texts that they need to engage with in the tests. Staff from two schools mentioned that the gains in fluency that they had observed would help pupils to tackle national assessments with speed, allowing them to focus on retrieving key information. The practice that pupils have had from PALS-UK in terms of focusing on the meaning of the text, paraphrasing, and being concise were also noted as helpful elements for national assessments preparation and staff from one school noted that PALS-UK provides a more interesting approach to developing these skills: '*I think it is good. It is really involving for them. it is not just sitting, looking at questions and things like that*' (T interview). Staff at another school described PALS-UK as providing '*a really good foundation for them for Year 6*' (T interview).

Teacher and pupil experiences of PALS-UK

As described in response to research question 1.2, there were many enabling aspects of PALS-UK, which meant that all case study teachers found it easy to deliver overall. One teacher valued the fact that it provided high impact for minimal workload in terms of administration. Staff at three schools noted that the planning and day-to-day preparation for PALS-UK was minimal and that you could '*just get the box out*' (T interview). It was noted that the teacher needed to put considerable time into thinking about the pairings and, which book would be most appropriate for each pair, but that '*once that's there, […] it's a really straightforward system to run*' (T interview). Staff at one school mentioned that the nature of the intervention made it feasible during periods without '*high levels of staffing*' (T interview), which they noted was normally required for English lessons given the range of ability within the class.

Most teachers responding to the survey said that their classes enjoyed PALS-UK and that it was overall positive for pupils, including reluctant readers. Pupils were engaged with the activities and procedures, enjoyed working in pairs, keen to read the books provided, and talk about them with classmates.

Most pupils interviewed said that they had enjoyed participating in PALS-UK and perceived that it was beneficial. One teacher reporting feedback from pupils said that they were very positive and none of them said that PALS-UK was boring. Another two teachers said that most of their pupils said that they enjoyed PALS-UK.

However, some teachers felt that pupils with negative attitudes towards reading (unsurprisingly) tended to be less motivated to participate. Some teachers reported that pupils' level of reading attainment affected the extent to which they enjoyed PALS-UK. There were some comments that those with low reading attainment were more likely to lack confidence in reading out loud and being corrected by their peers and experience frustration at trying to keep up with their partner, while those with high reading attainment potentially experienced a lack of challenge in the activities, disliked reading books aimed at lower attaining readers, and listening to their partner, found the programme too slow paced, and became disengaged. Some teachers reported that their classes' interest waned at certain points, particularly towards the end of the programme, due to the repetitiveness of the sessions, which two pupils also referred to. This was also included in feedback from the school that had stopped delivering PALS-UK partway through the trial. Some teachers commented specifically on the length of individual sessions:

> *Some were starting to lose enthusiasm by the end of the 20 weeks. Some of the children who struggle to remain focused on a day-to-day basis struggled to focus for the length of time required in PALS.* (T survey)

Teachers who responded to the survey also pointed to partner dynamics as an important factor in whether pupils enjoyed the programme or not. For example, having a dominating or disinterested partner, or where partners found it difficult to work together. More generally, pupils who dislike partner work and compromising with peers (e.g. on book choice) were less likely to enjoy the programme, as were those that disliked coaching.

Of the different PALS-UK activities, paragraph shrinking was the most popular with 14 pupils saying that they enjoyed it the most. Ten pupils particularly enjoyed prediction relay and one teacher commented that the pupils also said that they really enjoyed this activity. Five pupils (from five different schools) said that they liked retell the most and four pupils (from three different schools) said that they enjoyed partner reading the most. Pupils were also asked, which activity was the hardest and 13 pupils from five of the case study schools said prediction relay. Eight pupils from four case study schools felt that paragraph shrinking was the hardest activity. Three pupils from two schools said retell, two pupils from one school said that coaching was the hardest activity and one said that reading was the hardest.

In contrast, five pupils from four of the case study schools said that they did not like retell, with three of these feeling that it is boring and repetitive for the second reader to read the same bit of text. Another said that it is difficult to remember

what has happened. At one school, a pupil found prediction relay boring and at another school a pupil said that paragraph shrinking should be replaced with more reading.

Many of the pupils at one school perceived that there should be more time to do the PALS-UK activities. One suggested that there should be more time for retell and two said that because prediction relay is the last activity, there was not always time to check whether or not the prediction is true, and you do not necessarily remember what it was the next time you pick up the book. In contrast, one teacher said her pupils felt that too much time was allocated to paragraph shrinking and prediction relay.

Of those pupils who expressed an opinion, seven pupils from three schools said that they preferred PALS-UK to other approaches to learning reading, three pupils from two schools said that they preferred guided reading, and two pupils from two schools said that they preferred to read to themselves.

Over two-thirds of teachers said they had pupils who did not appear to enjoy the PALS-UK approach (70%, n = 50). Four teachers reported that most pupils in their class disliked or struggled with PALS-UK indicating that some pupils displayed frustration with the PALS-UK activities. Sources of frustration included wanting to read on rather than do the activities, wanting to continue reading a book rather than change it, disliking the book selection (one teacher commented that the level was too high), disliking changing partners, the repetitiveness and formality of the programme, the length of the programme and of individual sessions, and the lack of opportunity to discuss texts in class.

Only one pupil who was interviewed explicitly said that they did not like PALS-UK:

> *I don't really like PALS. I don't mind it, but if you had a choice, it wouldn't be the one to go to. It just gets a little bit boring.* (P interview)

Teacher as facilitator

One aspect of PALS-UK, which some of the teachers found valuable, but at times disconcerting, was the emphasis on the pupils' leading the activities rather than the teacher. One teacher perceived PALS-UK lessons provide the opportunity to take a step back and observe the pupils: '*I'll go and sit with the group and just stay with them and just listen to them read the book*' (T interview). Staff from four of the schools described how PALS-UK encourages pupils to be more independent, leading to a shift in the dynamics within the classroom. One teacher described how the structured nature of the activities allowed her to use a more pupil-led, facilitative (rather than directive) approach to teaching that might not otherwise be possible in what she reported to be a 'challenging class'. The same teacher described how as a result of PALS-UK, the pupils were becoming *'more independent in their reading'* (T interview) and were more willing to try to work things out or ask each other rather than going straight to an adult for help. One teacher described PALS-UK as providing 15 mini teachers rather than just one adult teacher to support the learning. One school leader reported that PALS-UK had opened up '*a different teaching style and a different way to work*' (H interview) to teachers, which was well received by them. And staff at another school reported plans to build more peer reading into reading pedagogy throughout the school, describing it as '*a lower stakes sort of thing, and it also guarantees that everybody is reading*' (T interview).

Most pupils said that they enjoyed working with a partner for a variety of reasons including being able to help others, receiving help and encouragement, learning vocabulary and improving comprehension, and the benefits of reading out loud including building confidence. Two teachers reporting feedback from their pupils echoed these comments. For example, one pupil said:

> *I really like reading with a partner because I think it's not as awkward as reading in front of the whole class, but you're not just reading on your own, like you're reading with someone but not in front of a large group.* (P interview)

Three pupils from one case study school noted that being a coach can be challenging. For example:

> *So, yes, the part I struggle with is being the coach, because when someone does something wrong and I have to say, 'Check it', I feel like I don't want to because it is going to disturb the person's flow of reading.* (P interview)

Three pupils from three separate schools felt that it can be awkward when they make mistakes reading in front of their partner.

> *I enjoy reading out loud because it makes it easier but I also find it a bit annoying when you make a mistake that you know it was right but you just stuttered on the word or something and your partner corrects you and sometimes it makes you feel a little bit embarrassed.* (P interview)

One aspect of the peer-led nature of PALS-UK, which some teachers expressed concerns about was monitoring and assessment of progress. As mentioned in response to research question 1.2, when discussing some of the barriers associated with specific PALS-UK activities, teachers reported that it was difficult to get a full sense of how successfully pupils were completing the different activities. One teacher explained this predicament as:

> *I'm listening to bits and I'm trying to pick them up, but you can't be everywhere at the same time, can you? […] How would they know and how do we know that they are always doing it correctly?* (T interview)

Pupil pairing

Three pupils from one school and one pupil from another noted that it was enjoyable working with peers that they did not usually sit next to or work with. One pupil noted that it was frustrating not finishing a book before changing partners.

Two pupils from two different schools described incidences where there was disagreement between themself and their partner in choosing a book. For example:

> *It is, kind of, hard to choose a book because you don't know which one…You and your partner might not agree on one that you really like but your partner might like a book that you don't like. Yes, it's just difficult picking a book.* (P interview)

Pupils from two of the schools commented that it was a challenge if the chosen book was not enjoyable.

Three pupils from one school said that it was difficult being partnered with someone you did not like or who did not engage fully with PALS-UK. Two teachers, summarising pupil feedback, also echoed this comment. One teacher said that her pupils noted that there was some frustration on the part of the first readers when the second readers were slow and found it difficult to undertake the tasks. Another teacher reported mixed comments from her pupils about partners due to disagreements about word recognition and pronunciation. In some cases, different pronunciations were due to different dialects and actually both correct.

Three pupils from three different schools said that staying with the same partner for four weeks was boring and that partners should change more frequently. One teacher explained that one pupil said that he would rather change partners weekly. A very small number of pupils from two schools noted that some partners were easily distracted and this caused problems. For example, the partner did not listen or did not intervene when they should have done.

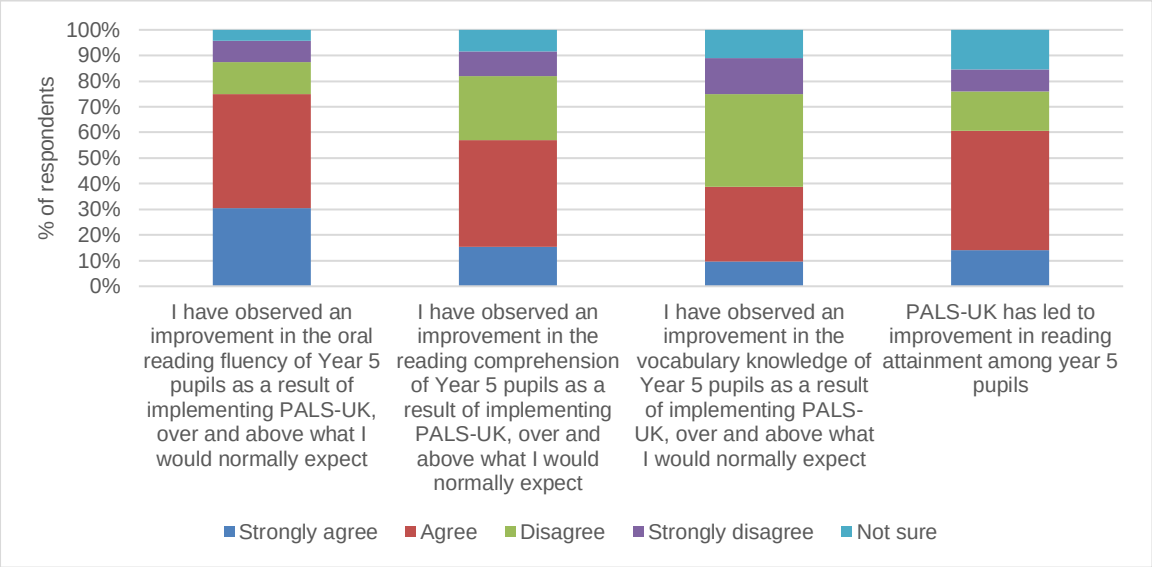
Perceived impact of PALS-UK

In the subsections that follow, we consider the impact that teachers and headteachers felt PALS-UK might have had on pupils and teachers. It is important to note, however, that the case study teachers spoke about the difficulties with trying to evaluate the impact of PALS-UK within their schools. Staff at two schools spoke about the difficulties of isolating the impact of PALS-UK from the impact of all the other things that are going on in schools, the influence of individual teachers, and other factors within pupils' broader lives. And staff at another school highlighted the difficulty of assessing the ongoing impact of PALS-UK as it happens due to its '*fast and furious*' (T interview) nature.

In terms of pupils' reading skills and attainment (**Figure 22**), fluency was the area that teachers were most confident that they had seen an improvement in (75% agreed or strongly agreed). Over 60% agreed that PALS-UK had led to improvements in reading attainment, and over half (57%) had observed an improvement in reading comprehension.

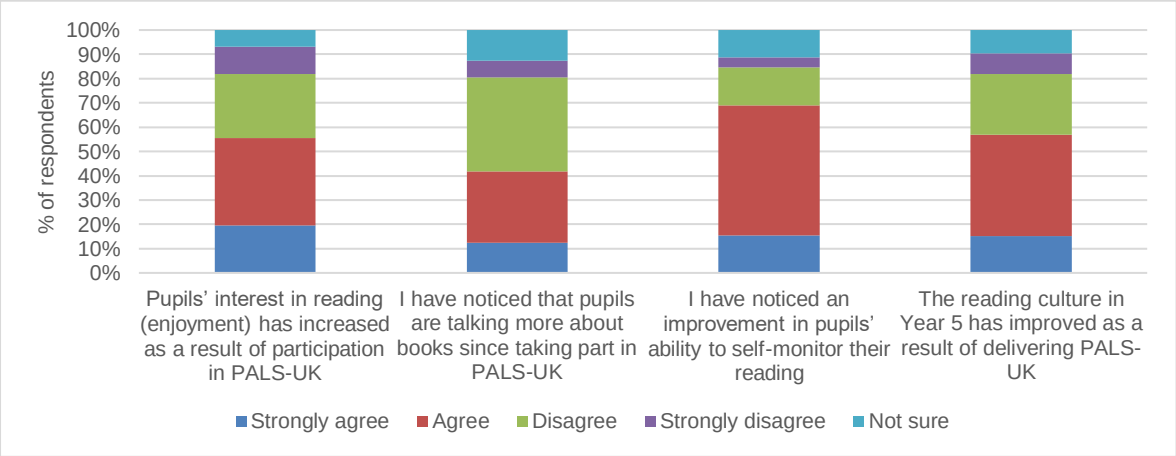
Vocabulary knowledge was the area that teachers were least likely to agree that they had seen an improvement in (39%).

Figure 22: The impact of PALS-UK on reading skills and attainment, as reported by teachers[56]



In **Figure 23**, over two-thirds of teachers (70%) strongly agreed or agreed that there had been an improvement in pupils' ability to self-monitor their reading. Over half believed that the reading culture in Year 5 had improved (57%) and that pupils' enjoyment of reading had increased (55%). Less than half (42%) of teachers however, reported that pupils were talking more about books.

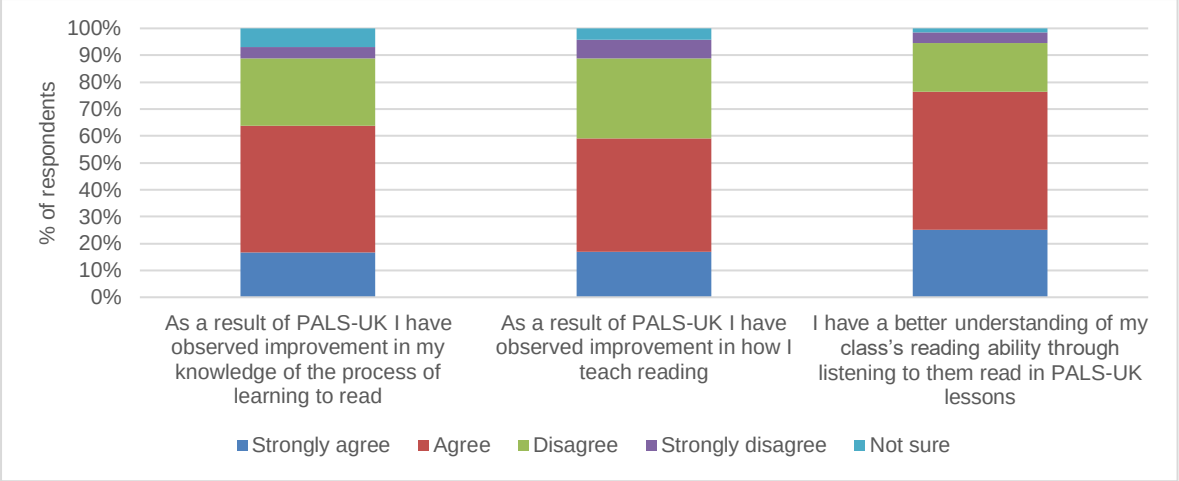
Figure 23: The impact of PALS-UK on reading culture and pupils' reading self-efficacy and motivation, as reported by teachers



Most teachers agreed or strongly agreed (**Figure 24**) that PALS-UK led to an improvement in their understanding of their class's reading ability (76%), their knowledge of the process of learning to read (64%), and how they teach reading (59%).

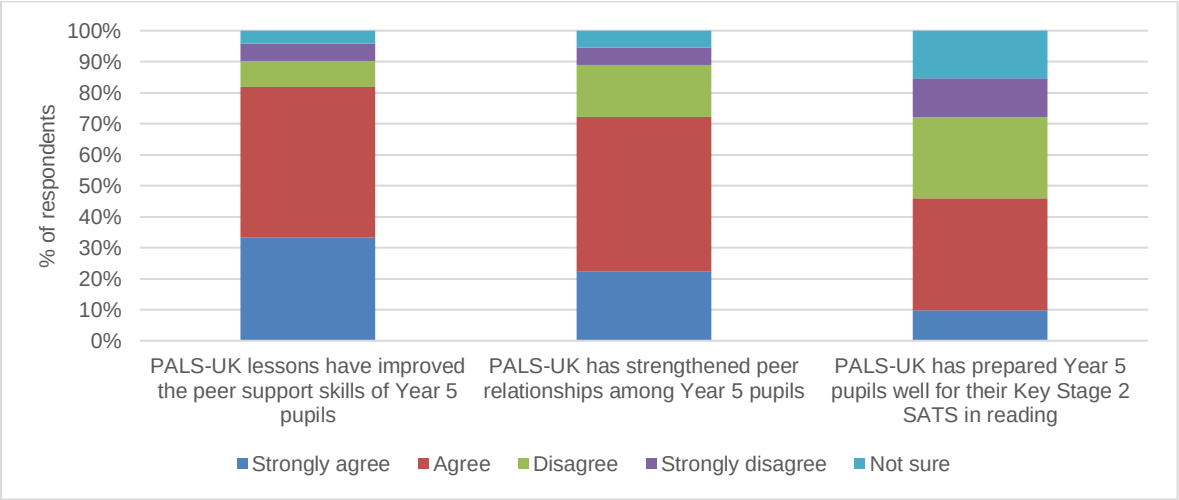
Figure 24: The impact of PALS-UK on teachers' confidence in teaching reading, as reported by teachers

[56] Teachers' endline survey question: 'Q46. Please indicate the extent to which you disagree or agree with the following:'



As shown in **Figure 25**, most teachers agreed or strongly agreed that PALS-UK was beneficial for peer relationships (72%) and peer support skills (82%). There was less agreement on the role of PALS-UK in preparing pupils for Key Stage 2 SATS, with less than half (46%) agreeing or strongly agreeing that this was the case.

Figure 25: The impact of PALS-UK on peer relationships and readiness for Key Stage 2 national assessments, as reported by teachers

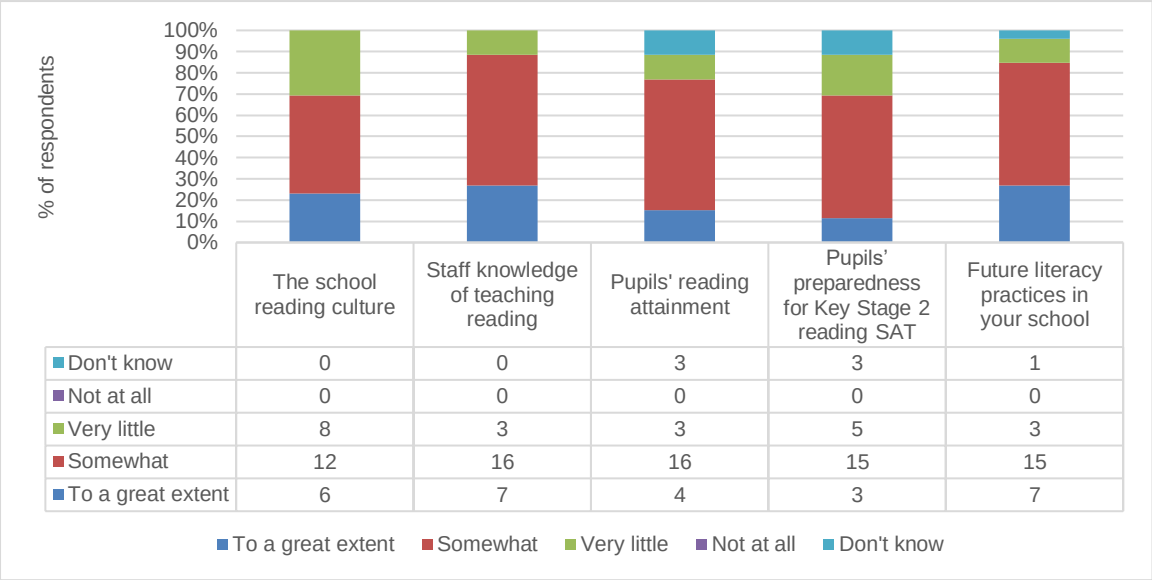


Teachers were asked to explain what the main benefits of PALS-UK are. Improvements to reading skills were frequently mentioned by respondents. Fluency was the most commonly mentioned benefit to reading skills, especially for lower attaining readers. Others included summarising and retelling, predicting, comprehension, pronunciation and expression, reading stamina, self-assessment in oral reading, retrieval, and vocabulary. Peer support skills were also mentioned by many respondents; a typical example includes: '*being a supportive listener*' (T survey). Building peer relationships within the class was another benefit perceived by teachers who responded. Self-efficacy was the next most frequently cited benefit; respondents said that the programme had improved pupils' confidence in reading, particularly for lower attaining pupils. Pupils' enjoyment of reading was also felt to be a benefit by respondents, as was the provision of a range of high-quality books. Teachers also mentioned other features of PALS-UK that they felt were beneficial: regular reading; peer support; having the text modelled by higher attaining readers; the repetition of clear activities; high levels of pupil engagement in lessons; and the ability of teachers to work with a range of pupils in each lesson with little preparation.

More than half of the teachers had seen any evidence of PALS-UK positively impacting on their pupils' reading for pleasure (58%, n = 41). Some teachers reported that pupils were keen to finish the books they read in their PALS-UK sessions, to continue sets, or find books by the same author of a PALS-UK book. Others said that pupils talked among each other about the books they had read and made or sought recommendations to/from peers. Respondents reported that pupils chose to read longer and more challenging books more during free time in the classroom, that they were asking to visit the school library more and to read novels as a class, that they were more willing to try books that were new to them, and that pupils reported enjoying reading more outside of school. Three teachers reported that some of those with less interest in reading became active readers, and that lower attaining readers were more likely to read for pleasure.

Headteachers perceived positive impacts of PALS-UK (**Figure 26**) particularly in relation to teaching practices (staff knowledge of teaching reading and future literacy practices in school). While the impacts on pupils' reading attainment, readiness for national assessments, and school culture was more contested, in all cases a majority of headteachers agreed that PALS-UK had had at least somewhat of an impact.

Figure 26: The impacts of PALS-UK, as reported by headteachers[57]



	The school reading culture	Staff knowledge of teaching reading	Pupils' reading attainment	Pupils' preparedness for Key Stage 2 reading SAT	Future literacy practices in your school
Don't know	0	0	3	3	1
Not at all	0	0	0	0	0
Very little	8	3	3	5	3
Somewhat	12	16	16	15	15
To a great extent	6	7	4	3	7

Headteachers reported a range of other positive impacts of PALS-UK, including on reading stamina, comprehension, self-assessment of oral reading, and reading for pleasure, alongside fluency and confidence particularly for lower attaining readers. Some headteachers reported that PALS-UK had shown the school new ways to teach reading, particularly for pupils who do not particularly enjoy reading, and that their schools would continue to use aspects of PALS-UK. Others said that it had exposed pupils to a range of texts and that pupils had begun to read more widely and frequently. Two responded that no impacts had been seen.

In the subsections that follow, in order to attempt to explain and elaborate on the survey findings presented above, we report teachers' and headteachers' perspectives on the mechanisms behind these perceived impacts.

Perceived impact on reading comprehension and reading attainment

Staff at all six case study schools perceived that PALS-UK improved pupils' reading comprehension skills. Staff from three of the schools felt that PALS-UK had had an impact on pupils' ability to make accurate predictions based on the text, with a more granular proximal focus (rather than making vague predictions about what might happen at the end of a text). Staff from four of the schools perceived that PALS-UK had improved pupils' ability to summarise key points from the text. Staff at one school were impressed by how well the pupils were now able to *'get the gist'* (T interview) of a passage and effectively paraphrase it. Staff at another school suggested that the paragraph shrinking activity had helped pupils to progress in this area. Beyond these isolated skills, staff at one school reported an overarching shift for some of their weaker readers in terms of *reading for meaning*: self-monitoring their understanding as they read and going back if something does not make sense or sound right. This same school suggested that while PALS-UK does not have a specific focus on vocabulary development, by encouraging pupils to read for meaning, it helped pupils to notice and question new words within the texts. Staff at one school mentioned that they had seen gains in pupils' reading assessment scores in relation to prediction, summarising, vocabulary, and impact (although impact—consideration of how authorial choices enhance meaning—was not something targeted by PALS-UK or mentioned in relation to it within the discussions).

Pupils from all but one of the case study schools perceived that PALS-UK had a positive impact on their comprehension. Three pupils commented that paragraph shrinking improves comprehension while a further three felt that retelling the

[57] Headteachers' endline survey question: 'Q32. Please indicate the extent to which you believe PALS-UK has had or will have a positive impact on the following'.

story improved comprehension. One pupil felt that reading out loud makes it easier to understand the text while a further two commented that working with a partner helped comprehension. For example:

> *And when you're reading in your head, sometimes you make mistakes and that's okay but you don't correct yourself, you just carry on reading, but with a partner they correct you and they help you with all of the understanding of the words and the meanings.* (P interview)

Perceived impact on fluency

Staff at all six case study schools perceived that PALS-UK improved fluency. Staff at one school reported that their Year 5 pupils performed better on a fluency test that the school was using than the Year 6 pupils. Staff at another school felt that the gap between weaker and stronger readers in relation to fluency had narrowed since implementing PALS-UK. Gains in fluency were reported to include improvements in use of expression and intonation, as well as accuracy. Staff at one school suggested that these gains might have been due in part to pupils having an audience in their partner; it was suggested that reading to a classmate encouraged them to try to be accurate and to use expression to hold their partner's interest. Staff at another school suggested that pupils read more fluently because the paired reading opportunities had helped to boost their confidence. Finally, staff at three schools suggested that the key driver behind the apparent improvements in fluency was the volume of reading that the pupils engaged in due to PALS-UK, although staff at one of these schools felt that these gains might be limited slightly by the lack of an adult role model in terms of expression and intonation.

One of the most common impacts that pupils referred to in all but one of the case study schools was learning new words or vocabulary. In particular, they felt that having a partner, reading out loud, and listening really helped them to identify errors and increase their knowledge. Six pupils from four of the case study schools also perceived that partner support and reading out loud improved their fluency:

> *I don't really mind reading out loud that much, even though I'm used to just reading in my head, but reading out loud, it allows my partners to correct my mistakes and I can do the same for them.* (P interview)

> *I feel having a partner does help because as someone already said, when you're reading on your own you can miss words or pronounce them incorrectly, but if you read them out loud you can hear that its incorrect or the partner can spot it for you.* (P interview)

> *I really like it because if you're struggling on a word, the other person can help you but if they don't know a word, you can try and work it out together. Then, it just makes it more fun because it's fun working with each other.* (P interview)

One teacher also noted that some of the pupils in the class had *'moved onto harder books with harder words'* (T interview). The same teacher also said that the pupils perceived that their vocabulary, fluency, and stamina had improved as a result of PALS-UK. Three other teachers noted that the pupils in their classes had reported that PALS-UK helped them to read better and develop their vocabulary.

In three schools, pupils were asked if they felt that participating in PALS-UK had made them a better reader compared to how they were learning to read before. The majority of pupils in two of these schools felt that PALS-UK had made them better readers. In the third school, most pupils felt that learning to read in a variety of ways was the best approach, and that this would address the perceived limitations of PALS-UK (such as the perceived repetitive nature, and not being able to complete long books within four weeks, etc).

In contrast, three pupils from one school felt that their reading level had stayed the same.

Perceived impact on feelings about reading

The logic model predicts that PALS-UK will improve pupils' reading self-efficacy, motivation, and attitudes towards reading. It also predicts that PALS-UK will develop pupils' confidence in reading and their tendency to read for pleasure. Staff at all six case study schools felt that PALS-UK may have had a positive impact on one or more of these constructs. Staff at one school described how the low stakes nature of PALS-UK supported reading self-efficacy, using the following example:

> *…having the time alone, just reading to one person helps that self-belief of 'actually, yeah, I can. I can read out loud to someone else who's just one person. I don't have to shout. I can just use my normal voice and know that that person's going to support me'.* (T interview)

Most pupils who participated in focus groups felt that PALS-UK had made them better readers for a variety of reasons including support from a partner, additional practice, increased vocabulary, being more aware of making mistakes (especially through reading out loud), improved comprehension, and fluency. One of the teachers also said that their pupils felt that PALS-UK helped them to identify mistakes, particularly the 'check-it' procedure. One pupil said:

> *I think [working with a partner] is very good, so you can share your ideas of the book and if you enjoy it or not. You can see if they like reading or if they could improve that reading or if they're struggling with words, we could help them out.* (P interview)

One pupil added that because they had become a better reader, they were reading far more than previously:

> *It made me read a lot more because when I get books now, I read them a lot more. I've got four different books on my bedside table and it's like made me enjoy it a lot more because you know that you can read a little bit better. So, it makes you want to read more because you can do it better.* (P interview)

Another teacher described the apparent confidence boost following engagement with PALS-UK as being due to the way that it scaffolds the overall task of reading comprehension into manageable pieces:

> *I think definitely [an impact on] confidence […] You don't have to understand what everything is straight away. You can read it […] and you can begin to work out what words mean because you read them in context. Just a bit more confidence in reading, I think.* (T interview)

Teachers also described the impact that pupils' growing confidence had on their identities as readers, perceived to be in part due to them being able to access longer, more difficult texts. For example, one teacher remarked:

> *It's that sort of reading identity isn't it, like, 'I can read big books'. You know that's a big thing.* (T interview)

Four pupils from two different case study schools perceived that PALS-UK, and particularly reading out loud to a partner, had a positive impact on their confidence in reading. A teacher commented that pupils' self-belief had improved although at this school there was a strong emphasis already on *'cultivat[ing] strong confident readers'* (T interview). A teacher at another school reported that three of her pupils had said that their confidence in reading had improved.

The experience also made them more confident about reading out loud because reading in a pair is less pressured than reading to the whole class, a point also made by two of the teachers reporting on feedback gathered from their pupils.

Staff from some of the case study schools also mentioned that some aspects of PALS-UK might differentially impact first readers versus second readers. One school reported that some first readers appeared to be losing engagement or motivation as they were less challenged by the activities and that this might have a knock-on effect on some second readers as a result of the first reader's lack of investment in the process.

Staff from all six case study schools commented on the value of pupils being exposed to a range of different genres and authors. One teacher commented on the potential for these opportunities to transform pupils into readers:

> *I think some of the people who have taught them previously, might be quite shocked at some of the children who now would be saying I'm a reader and I like this, and I would like this type of book.* (T interview)

As well as access to a broad range of texts, teachers also spoke of the social dimension of peer reading as something, which seemed to promote enjoyment of reading. As argued by one teacher:

> *There is a lot more enjoyment of reading when you're sharing it with somebody else, so they can have those natural discussions between their carefully matched pairings and they're reading for pleasure almost without realising it.* (T interview)

Some teachers talked about pupils choosing to read more outside lessons, including wanting to carry on with the PALS-UK books. One teacher talked about having conversations about authors with her class that had not happened previously. It was noted however that, PALS-UK does not provide a panacea for developing reading for pleasure universally.

> *For some children, I don't think it's had probably much of an impact because they were good readers and they read a lot anyway, they loved reading. So they've just sort of continued that love for reading. And others have started to read more often or have changed maybe the type of books that they're reading. And then there's another group who probably still lack some enthusiasm and I think will always lack enthusiasm, no matter what we do with them. I think they're kind of just slightly switched off reading, which is a shame.* (T interview)

Several pupils perceived that participating in PALS-UK meant that they engaged in more reading (themselves) in the classroom than they had done previously.

Four pupils from three of the six case study schools commented that after participating in PALS-UK they were reading for pleasure (at home) more than they had done previously. Three teachers summarising feedback from their pupils made the same comment. Three pupils at one case study school felt that they now enjoyed fiction whereas previously they would have preferred non-fiction books. Again, one teacher made a similar comment based on conversations with the class after PALS-UK had finished.

> *We had pupil[s] saying that they are more likely to take books on holidays or if they were going on a train journey. And some of them also commented about being more included to go to the library to try and find other books.* (T interview)

Two pupils from the same case study school said that they did not enjoy reading prior to participating in PALS-UK, and that now they choose to read in their own time. One teacher echoed these comments, noting that their pupils had said that they now *'know more about what sort of books they like'* (T interview). Similarly, another teacher reported that their pupils felt that they liked reading more than they did before. She also provided examples of pupils reading longer texts, harder books, and a wider range of texts than before.

In relation to school reading culture, one pupil said:

> *I notice more people reading at lunchtime and there seem to be more books around.* (P interview)

One pupil felt that the partnering role had helped to develop their teamwork skills. Another from a different case study school felt that helping someone else also helped them to improve their own reading as well. Two teachers reported that pupils in their class perceived that PALS-UK had helped them to work better with others. As an interesting aside, one of these teachers commented that she might mix up groups more in the future and not just specifically in reading activities.

Perceived impact on peer relationships

Staff at five of the six case study schools talked about PALS-UK having an apparent impact on peer relationships within the classroom. Staff at one school stated that the *'social side of it has been massive, absolutely massive'* (T interview). They talked about how PALS-UK supported pupils to develop a range of social skills including turn-taking, cooperation, and compassion. Within this particular school, the class was described as challenging' but it was reported that PALS-UK helped some of the pupils who had tended to become involved in conflicts (e.g. on the playground) settled really well during PALS-UK, which helped them to forge stronger relationships with their peers. A teacher at another school talked about how PALS-UK served as a vehicle for teaching pupils who were able (with guidance from the teacher) to learn to work together and be civil, despite initially struggling to get along. This same teacher warned, however that there is a danger that *'dysfunctional pairings'* (T interview) might be missed and not addressed, where pupils are not engaging fully (e.g. feeling too scared to point out the other pupil's mistake) because of the particular relationship dynamic. Teachers sometimes talked about how the pupils surprised them in relation to the way that they interacted with each other, and that PALS-UK helped them to see their peers in a new light.

Perceived impact on teacher knowledge and pedagogy

Staff at five of the six case study schools talked about how PALS-UK had impacted on teacher knowledge or the way that teachers thought about the teaching of reading. Several teachers spoke of how they had developed a broader repertoire of strategies and activities to use to support reading, including techniques they *'would never have thought of'* (T interview). One teacher reported that the paragraph shrinking activity made her realise *'how important it is for children to actually realise what the main theme of that paragraph is and what are the most important things'* (T interview). Another teacher noted that the structured, fast-paced nature of PALS-UK had led her to see the value in being more succinct with her explanations. Two teachers reported that the prediction relay activity helped them to think differently about how to support pupils' prediction abilities:

> *Even my own, you know, as a reader, I'm thinking when somebody asked me this random, 'let's predict what this text is about', you always do these overarching predictions that are way too vague.* (T interview)

However, one teacher also spoke of missing the depth of engagement involved in her previous method of teaching reading through whole-class discussion.

Overall perceived impact of PALS-UK by teachers and headteachers

While staff at all case study schools raised both advantages and limitations in relation to PALS-UK, they all appraised it positively overall. Staff from all six schools suggested that they were considering taking it forward in some form and/or they would recommend it to other schools. However, it was highlighted that it is important for other elements of reading to continue to be taught in sessions outside of PALS-UK in order to fully address all of the reading content domains. It is important to note that PALS-UK is not designed to address all aspects of reading but to focus on oral reading fluency and development of specific reading comprehension skills (as set out in the logic model).

Teachers were asked whether, if the programme was available to their school in the future, they would like to continue implementing PALS-UK with almost two-thirds saying yes (60%, n = 42). Seven (27%) headteachers felt that PALS-UK led to changes in how reading is taught in Year 5 and seven (27%) felt that the changes applied to other year groups. Respondents were able to choose more than one answer.

Summary of key findings

- Staff perceived that PALS-UK is easy to deliver, with some time required for considering pairings and supporting book selections. The consistent structure worked well although some pupils were less engaged towards the end of the 20 weeks, which attributed to the repetitive nature of the programme.
- Staff perceived that PALS-UK is a useful approach to teaching reading, particularly partner reading and paragraph shrinking.
- Staff and pupils perceived a positive impact of PALS-UK on fluency, attainment, and comprehension, partly attributed to the volume of reading undertaking and reading in pairs.
- In relation to the development of vocabulary, fewer staff perceived that there had been a positive impact, but the majority of pupils frequently talked about learning new words. While PALS-UK does not involve explicit teaching of vocabulary, the logic model suggests a reciprocal relationship between vocabulary and oral reading fluency, and vocabulary and reading comprehension.
- Staff and pupils also perceived that PALS-UK had a positive impact on self-monitoring (developing awareness of mistakes), enjoyment of reading, and reading for pleasure.
- Staff also perceived that there was a positive impact on preparing pupils to read longer and harder texts.
- Staff perceived that there had been a positive impact of PALS-UK on pupils' development of peer support skills (including turn-taking, cooperation, and compassion) and peer relationships.
- However, there were a number of partnership issues including not liking the partner, disagreeing on book choices, understanding when to intervene, not intervening when required, one partner not engaging, and disagreements on words and pronunciations.
- Staff had some concerns that first readers might not benefit much from PALS-UK and were more inclined to experience frustration. They also found it difficult to monitor what was happening in the classroom and to identify partnerships that were not fruitful or where PALS-UK was not being implemented correctly.

- Staff identified some gaps that might be required for ensuring preparation for Key Stage 2 national assessments, particularly developing inference, impact, and structure.
- Both senior leaders and teachers felt that there had been a positive impact on teachers' pedagogy and understanding of teaching reading.

Contextual factors

This section examines, which contextual factors may influence the effectiveness of PALS-UK from the research question:

1.4. What contextual factors contribute to or inhibit PALS-UK effectiveness?

The factors that will be examined are school context and resources, senior leadership support, classroom management, and pupil characteristics. Data for this section comes from the teacher and headteacher surveys, and the teacher, headteacher, and literacy lead interviews from the case studies.

School context and resources

The case study teachers identified reasons why working in a bigger school with multi-form entry may act as an enabling or limiting factor to the effectiveness of PALS-UK. One teacher mentioned how they valued the support from their other Year 5 colleagues, noting that being in a bigger school facilitated the peer observation process. The same teacher also noted, however, that it could be harder to maintain consistency of practice within a bigger school. A Year 5 teacher from a one-form entry school talked about the peer observation process feeling less useful because the colleague observing them had not been teaching PALS-UK and so it was less familiar to them. While this colleague had attended the training it was a while back and she had not had the opportunity to teach PALS-UK since then. Another issue associated with being in a small school is that the changing of pairs becomes harder as the intervention progresses as there are fewer combinations of pupils to select from (as discussed in, in relation to research question 1.2).

Some of the case study teachers spoke of dynamic contextual factors, which might affect the effectiveness of PALS-UK. On the positive side, one teacher reported that PALS-UK was particularly helpful to them at a difficult time of transition between sites as its routine nature provided consistency for the pupils during a time of change. On the other hand, a teacher at another school talked about the difficulties of trying out a new approach to teaching reading while under pressure to *'get results up rapidly'* (T interview) following an Ofsted inspection. This teacher expressed concerns over the potential for PALS-UK to leave gaps that needed to be filled in order to make the required progress.

In general, case study teachers praised PALS-UK for coming as a ready-made package, which for most pupils did not require any additional resources. As discussed in previous sections, there were often few pupils who needed to use different books than those included with PALS-UK, because they required books below the reading proficiency level of those provided; however, as mentioned previously, the books provided with the package were not intended to be a comprehensive selection. Teachers were told in the training to supplement these books with the books they have in school. It was suggested that teachers played an important role in selecting an appropriate book for each pair. One teacher noted:

> *That's a school factor in the fact that we've got a lot of books to choose from. Teachers are very well read and know generally what to recommend and what book will suit that pair.* (T interview)

This suggests that to maximise the success of PALS-UK schools need access to additional high-quality pupils' literature and staff need to have the expertise to select books appropriately.

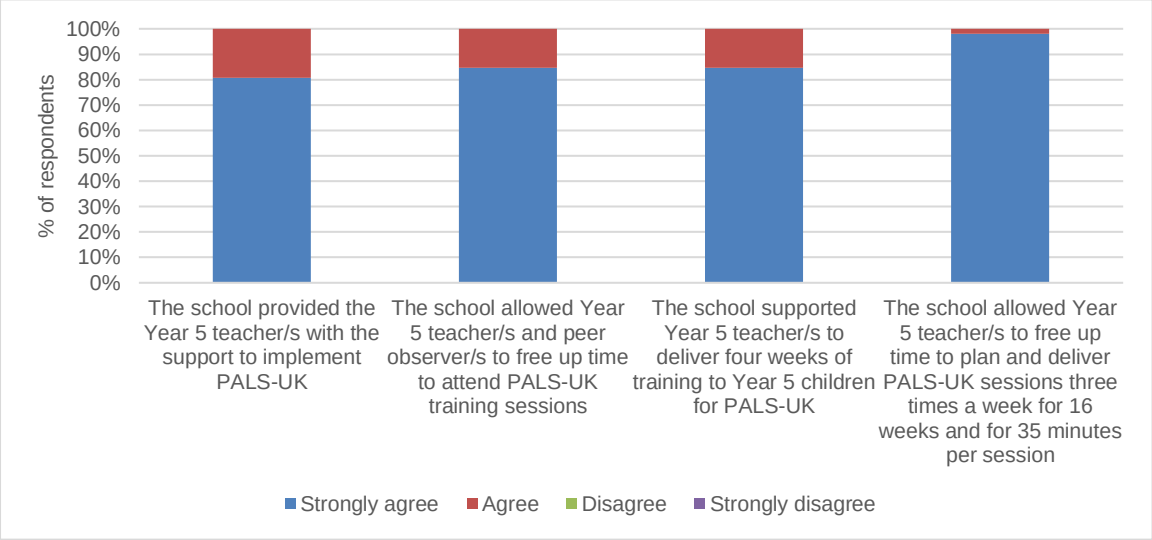
Senior leadership support

Most teachers (80%, n = 58) indicated that they had support and engagement from senior colleagues in school to deliver PALS-UK. Some teachers explained that leaders were supportive with the initial stages of the trial: they instigated the schools' participation; attended training and acted as a peer observer; made time for PALS-UK on the school timetable; supported changes to school reading programmes to accommodate PALS-UK; and supported the informal training of additional staff such as higher level teachings assistants to deliver PALS-UK if necessary. Others referred to support given throughout the project; for example, providing teaching cover in the case of mixed year group classes, visiting classrooms to observe implementation, and giving teachers time to observe each other, advising on any queries, for

example, about class dynamics, assigning time to discuss PALS-UK in school meetings, asking for updates, and giving encouragement and praise about PALS-UK and how it was being implemented.

Headteachers also indicated that teachers were well supported in the delivery of PALS-UK (**Figure 27**).

Figure 27: Support provided by schools for the delivery of PALS-UK, as reported by headteachers[58]



Around 92% (n = 24) of headteachers also told us that they received regular (at least fortnightly) updates from the lead teacher in their school about how the project was progressing.

Staff at all six case study schools also talked about the importance of the SLT playing a supporting role in ensuring the successful implementation of PALS-UK. These teachers noted that they valued the headteacher or literacy lead accompanying them on the training (as well as releasing them to attend), discussing how it might work in their school with them, and their support to troubleshoot any issues, for example, in terms of pupil pairings. Teachers highlighted the importance of senior leadership facilitating protected time for PALS-UK on the timetable, and the need for PALS-UK to be aligned with the school's existing priorities. Some of the teachers also talked about the importance of staff in school being openminded and prepared to try out new approaches. One teacher noted that it was important for all staff involved to feel enthusiastic about the changes to practice rather than experiencing PALS-UK as being imposed '*top-down*' (T interview). One headteacher described the ideal scenario as:

> *We're doing this, we're excited about doing it. We've got the training, we've got the resources, it's timetabled, we've got all the literacy leads on board, senior leaders on board.* (HT interview)

Classroom management

Many teachers indicated that they taught PALS-UK according to the manual. They confirmed that they carried out key tasks including setting out resources and controlling the timer, circulating the room and observing pupils, providing support with vocabulary or focus, motivating pupils, giving feedback, introducing the activities, and reminding pairs of procedures where necessary. Teachers from two case study schools highlighted that an enabling factor was a sense of knowing what you are doing, facilitated by the high-quality of training and materials (see above). One case study teacher also emphasised the importance of teachers having good knowledge of the books and their class so that they could ensure appropriate pairings of children and allocation of books. Some survey respondents said that they aligned PALS-UK with existing classroom practices, for example, amending a vocabulary display or behaviour charts, giving rewards, or making notes for class records. One teacher led the class in a book discussion following each session, while another reviewed new vocabulary.

Staff at two case study schools said that for the intervention to be successful, teachers need to be committed to the trial, which involves a willingness *''to try new things'* (HT interview). As part of this commitment, one school identified that it

[58] Headteachers' endline survey question: 'Q27. Please indicate the extent to which you disagree or agree with the following statements'.

is important for teachers to play an active role in monitoring the children during the activities and three schools highlighted the teacher's role in getting the children *''on board'* (T interview) through their own enthusiasm. Staff from one school noted that seeing the pupils enjoying PALS-UK helped them to sustain their own enthusiasm and commitment.

Teachers said that they had different strategies for supporting pupils. For example, spending two to three minutes per pair, visiting all pairs within two lessons or focusing on one group, table or pair of pupils per lesson, focusing support on lower attaining readers and pupils with behavioural issues, or partnering readers where their partner was absent. Some reported that the implementation of this strategy would depend on how much support the rest of the class needed with reading and behaviour. A few teachers reported that they had a teaching assistant for some PALS-UK lessons. Some teachers reported that pupils were increasingly able to independently and correctly follow procedures through the programme.

Staff at four of the six case study schools talked about pupil absence being disruptive to the implementation of PALS-UK. These staff reported that it was difficult to amend pairings on the spot and that triads did not work as effectively as pairs because pupils are *'not as engaged in a task all of the time'* (T interview). One teacher described the problem with pairings having to be shuffled around due to absence as follows:

> *Obviously one of them doesn't then know the book so they spend time telling them what the story's about and obviously with the time being so structured it doesn't work very well. I mean, we made it work, but it's not ideal.* (T interview)

While all four of these schools described pupil absence as a challenge, this was much more significant for schools that had ongoing issues with attendance. Some teachers commented that pupils who were persistently absent, easily distracted, or had difficulty working cooperatively with others were less able to engage with the programme.

Pupil characteristics and suitability for PALS-UK

Around one-third of teachers (38%, n = 27) indicated that PALS-UK had more impact on particular groups of pupils. The number of teachers perceiving PALS-UK to have a particularly positive impact on high, middle, and low ability readers were ten (37%), seven (26%), and six (22%), respectively (N.B. This includes one teacher who said that the programme particularly benefited both lower and middle ability readers). It was suggested that the particular benefits of PALS-UK for less able readers related to frequent opportunities for them to develop fluency and confidence at reading out loud. Having a partner to help them and correct their errors was also reported as being important for lower ability readers, providing opportunities for them to become more precise and more confident at reading out loud. Teachers who talked about the benefits of PALS-UK for middle ability readers suggested that the modelling provided and knowing what to expect were particularly helpful for this group. Where teachers reported PALS-UK being helpful for higher ability readers they suggested this was due to developing skills to cope better with the activities than other learners, and that they were able to develop their reading skills further through coaching their partner. Two teachers suggested that EAL pupils responded particularly well to the programme because of the opportunities to listen to their partner read, which helped them to learn new words. They also noted that opportunities to recount key events from the story and check their understanding with a partner may also have been particularly helpful for EAL pupils. One teacher did not differentiate between low/middle/high ability readers but suggested that the second reader tended to benefit more than the first reader. Another suggested that PALS-UK worked particularly well for those children who read regularly at home. Finally, another teacher suggested that those children who already had a love of reading were the ones that gained the most from PALS-UK, as they enjoyed the opportunity to explore new books. Taken together, this suggests that teachers perceived PALS-UK to be most beneficial for lower ability readers, but that particular benefits for EAL pupils, middle and higher ability readers were also noted.

Case study teachers also talked about the potential of PALS-UK to develop the confidence of weaker readers; however, teachers from five of the six case study schools noted that PALS-UK was not suitable for all pupils. Case study school staff provided examples of particular pupils who were low attaining readers, were new to English, or had autism spectrum disorder, who they felt were not able to access PALS-UK unless it was adapted significantly. For example, teachers mentioned using a teaching assistant as the first reader, and also using much simpler texts. Staff at one school reported that two pupils with autism *'couldn't handle being focused for that long'* (T interview) and so were provided with an alternative reading activity. Another school described a child with autism, who was a strong reader, asking: *'Why am I doing this? I don't need to read with somebody else. I can read' (*HT interview*).* These findings, therefore suggest that

schools with high proportions of SEND, lower attaining readers, and/or high numbers of pupils new to English might experience challenges implementing PALS-UK as a whole-class intervention. We return to this issue in the 'Conclusion' and 'Interpretation' sections below of this report. For example, one teacher noted:

> *I would imagine that managing PALS in a school where there are lots of variations of English, due to something like EAL or SEND, would be a real challenge.* (T interview)

Less than half of the teachers (44%, n = 32) felt that there were pupils who struggled to adapt to the PALS-UK approach. Pupils with SEND were the most commonly mentioned group, suggested by nine respondents; teachers explained that additional needs might lead to difficulties with the texts, with maintaining focus, with the social aspects of PALS-UK, or with effectively supporting their reading partners. Six teachers also felt that lower attaining readers were more likely to struggle to adapt to PALS-UK, especially those still requiring support with phonics, who respondents said struggled with fluency and confidence. Staff at one case study school reported that they had a small number of pupils who were very low attainers in reading, who they felt would not be able to access PALS-UK. They took part in a phonics intervention instead. Pupils who found it challenging to work well with others, pupils learning EAL, and pupils displaying challenging behaviour were also mentioned by four, two, and two respondents, respectively as groups of pupils who had experienced the most difficulties in adapting to PALS-UK. This was echoed by one of the case study teachers who said that an EAL new arrival also engaged in alternative activities as it was felt that PALS-UK was *'not the best choice'* (T interview) for them. Feedback from a school that stopped implementing PALS-UK partway through the trial also referred to concerns that the programme did not support EAL pupils with learning new vocabulary. Just one teacher mentioned that some higher ability pupils found it hard to adapt to PALS-UK because they didn't like the slow approach.

At first glance, these IPE findings might appear to contradict the (weak) interaction effect observed for pupils with SEND (which suggests that PALS-UK may have been particularly beneficial for this group). This apparent inconsistency may come from treating pupils with SEND as a homogenous group. It is worth noting that from the very beginning, teachers were permitted to identify and exclude pupils with SEND from the study who had more pronounced reading challenges. Based on these qualitative observations, we tentatively suggest that there may be a threshold of capability among pupils with SEND, where those above it might derive greater benefits, while those below it might struggle to engage altogether. It is also important to note that more than half of the respondents said that they did not have any groups of pupils who struggled to adapt to the approach.

Just over one-third of teachers (37%, n = 26) felt that PALS-UK had less impact on particular groups of pupils. Again, this suggests that overall teachers found PALS-UK to have broad suitability for the different groups of learners within their classes. Where teachers did identify groups of learners for whom they perceived PALS-UK to have had less impact, the most frequently mentioned group was pupils with higher levels of reading attainment (ten respondents). Some teachers felt that pupils who were already fluent readers were not adequately challenged or stretched by PALS-UK. Pupils with SEND (eight respondents) and lower attaining readers (seven respondents) were also mentioned frequently. It was felt that some of these pupils had trouble finding suitable texts and would have benefited from more teacher input. One teacher suggested that PALS-UK had the least impact for EAL pupils, where the books might not be at their current English proficiency level, and another teacher suggested that PALS-UK may be less beneficial for boys than girls. This is in contrast to the comment reported earlier made by another teacher who felt that PALS-UK may be particularly appealing to boys due to its verbal nature.

Summary of key findings

- The size of the school can impact on the effectiveness of PALS-UK, and particularly for smaller schools in which (unsurprisingly) it is more challenging to organise peer observation and pupil pairing.
- Senior leadership support is perceived to be very important, particularly in relation to releasing staff for training and facilitating peer reviewing. The data suggest that it is important for all staff who will be involved in teaching PALS-UK to attend the training.
- Teachers' commitment to and active participation in delivering PALS-UK is an important contextual factor that contributes to effectiveness of the intervention.
- Pupil absence was described as a significant challenge to teachers' delivery of the lesson as it caused difficulties with pairings and pupils not knowing the book that their partner was midway through. They also described a range of different strategies for supporting pupils while they engaged with PALS-UK, which varied depending on the needs of the class.

- Staff reported a range of perceived differential impacts of PALS-UK depending on pupils' individual characteristics. While the findings were mixed, teachers suggested that PALS-UK had the most impact on lower ability pupils. More than half of teachers reported that there were no groups of pupils who struggled to adapt to PALS-UK, but others reported that some children with SEND found it hard to adapt to the programme and/or required significant adaptations from the teacher (and in some case alternative provision entirely). While some teachers reported that PALS-UK has potential benefits for higher ability readers, this was the group of learners that they perceived to benefit least from PALS-UK.

Unexpected outcomes

This section examines any unexpected outcomes that arose from the implementation of PALS-UK addressing research question:

1.5. Were there any unexpected outcomes?

One teacher mentioned in response to the survey that PALS-UK has influenced how they teach writing:

> *When planning writing, we will build up from the most important who/what for a paragraph, what we want to say about that thing and then build detail from there.* (T survey)

In contrast, one teacher in the case study interviews mentioned that one of the gaps associated with PALS-UK is that there is no explicit focus on the transition from reading to writing, that is, exploring a text that you are reading as a model for writing. One case study teacher identified that an incidental outcome of PALS-UK was that it generated *'book talk'* (T interview) outside the PALS-UK lessons:

> *They do genuinely talk about the books when they go and pick them up. They have conversations: 'Oh, we read this in PALS. Check it out'.* (T interview)

A teacher in another case study school noted that she had been impressed with how PALS-UK seemed to have developed pupils' metacognitive awareness:

> *They were talking about not just what they were doing, but why they were doing it.* (T interview)

This was also evident in the pupil interview at the same school, where the researcher was very impressed with the pupils' ability to reflect on the nature of the learning that PALS-UK facilitated. The pupils talked in-depth about the complexities of the coaching role and how it required skilled formative assessment skills. For example, one pupil noted:

> *Something that always makes me think, if someone gets a word very close, mispronounces a letter but it sounds similar, so I think if maybe you just said that one letter wrong, because they knew it was the right letter but they just couldn't say it right. That always makes me think for a second and stop before I say, 'Check…'.* (P interview)

Pupils were also able to talk in detail about how the different activities supported different aspects of their reading, suggesting that perhaps PALS-UK made the process of reading comprehension more transparent to them. For example, one pupil reflected on the learning that takes place when completing the paragraph shrinking activities as follows:

> *So, you need to explain the paragraph that you just read, no matter how long it is, and you need to get it down to at least 10 words. Basically, you need to simplify something that you just read so you need to understand it to simplify it.* (P interview)

Another positive outcome perceived by teachers to have occurred, which was not included within the logic model is the transfer of learning from PALS-UK to other areas of the curriculum beyond literacy. One case study teacher talked about how their pupils had found it helpful to use the ten-word summarising strategy from PALS-UK to help them to focus on key ideas in other subjects such as science and religious education.

One potentially negative unexpected outcome mentioned by a case study teacher, is that they needed to take five/ten minutes off their mathematics lesson because of PALS-UK taking 40 minutes once you took into account needing to move seats, etc.; however, the teacher suggested that they still had '*enough time for maths in the rest of the week*' (T interview).

Cost

As part of the trial, schools assigned to implement the PALS-UK programme were exempt from covering the expenses related to training fees and the training manual. Here, we present the actual costs associated with implementing the programme outside the trial context.

The average cost per pupil per year for schools implementing the PALS-UK programme was estimated following the EEF costing guidance released in 2023 (EEF, 2023). A year is defined as a year of implementation, aligning with the academic year. The calculated costs are based on the mean number of eligible pupils per class included in the evaluation (n = 25). Since PALS-UK is implemented at the classroom level, costs are reported per class as well as per pupil. In the case of multi-form entry schools, the class-level cost should be multiplied by the number of classes or forms in the school. **Table 20** details the resources required to implement the programme, following the ingredients method (Levin *et al.,* 2017). The primary cost for programme implementation is the training cost, with additional costs identified as staff cover and supplementary materials, categorised as mandatory.

Table 20: List of resources (ingredients)

Category	Item
Personnel for preparation and delivery	Classroom teacher for delivery of the intervention
	Teaching assistant or headteacher (optional)
Personnel for training	Classroom teacher for training
	Teaching assistant or headteacher (optional)
	Classroom teacher (cover for training)
Training costs	Training fee
	Travel fees for training
	Teacher cover
Programme costs	Manual
Facilities, equipment, and materials	Books[59] (optional)
Other programme inputs	Printing training materials
	New books (optional)

Pre-requisites

During the implementation of the programme, schools were expected to leverage their current learning resources. A survey conducted during the evaluation aimed to understand the existing resources that schools were using, aside from basic teaching facilities like classrooms. Most respondents reported that there were no pre-existing material resources necessary for implementing PALS-UK.

To participate in the PALS programme, schools were supplied with necessary teaching materials (e.g. books, and manual) as part of the programme package. Therefore, no significant pre-requisites were required other than utilising existing facilities for teaching purposes. Under usual circumstances, schools are obliged to independently provide the books used for the intervention. As per the delivery team, an average of 20 books per class (with an average class size of 25) is essential for effective intervention delivery, which was taken into consideration below to estimate the implementation cost.

Time

PALS-UK was intended to be delivered to the entire class during normal school hours. Additionally, all Year 5 teachers in the intervention arm were mandated to participate in the initial training sessions, which were conducted as full-day, in-person sessions, along with a half-day top-up session conducted online. The majority of time-related costs were front-loaded, primarily associated with start-up training. According to the teacher survey, the average time dedicated to training by staff was 15 hours. Typically, Year 5 teachers attended these sessions, sometimes accompanied by the

[59] In this trial books were provided for intervention schools to use, however, PALS-UK can also be implemented using existing books in schools

headteacher, teaching assistants, or a school reading lead. The majority—22 out of 25—schools responding to the survey also indicated an average of seven hours teaching cover was needed during the training period.

The online twilight sessions, lasting 45 minutes each, took place approximately midway through the trial, with two sessions scheduled. The teacher survey indicated low attendance at these sessions (22.2%).

Regarding preparation for the programme delivery, teachers delivering the intervention spent an average of 29 hours preparing. Some preparation by other staff was occasionally noted, where 20.6% of the schools reported an average of 19 hours of preparation by other staff (i.e. teaching assistant, headteacher). Around 71.4% of the respondents indicated that preparation was not considered extra time, while the remaining 29.6% considered around eight hours of the preparation as additional.

The intervention was designed to be delivered in 35-minute sessions three times a week over 20 weeks, suggesting an overall delivery time of around 35 hours per class. Staff survey responses confirmed this expectation (mean = 36). Five out of 25 survey respondents also mentioned that teaching assistants were involved in the programme delivery, averaging 29 hours. More than half of the respondents (59.2%) stated that the delivery was not considered additional to normal reading intervention delivery (mostly replacing existing reading lessons), while others reported PALS-UK as requiring hours of additional time beyond the usual time spent teaching reading. Two schools noted that they regarded it all as additional, and some mentioned that it demanded some extra time. On average, considering all respondents, the additional time required to deliver PALS-UK amounted to seven hours.

An additional aspect of peer support observation was part of the programme agenda, with 26 out of 29 respondents claiming that a staff member was involved in this activity. Most commonly, it was the SLT (n = 11) or a teaching assistant (n = 8), with an average time of three hours.

Table 21: Total time devoted by personnel for training, teacher cover, as well as for preparation and delivery

		Year 1		Year 2		Year 3	
		Number of staff per class	Mean number of hours (min, max)	Number of staff	Mean number of hours (min, max)	Number of staff	Mean number of hours (min, max)
Training	Classroom teacher	1	15 (6, 36)	0	0	0	0
	Other staff	1	4 (0, 15)	0	0	0	0
Teacher cover	Classroom teacher	1	7 (0,13)	0	0	0	0
Preparation	Classroom teacher	1	29 (3, 60)	1	22 (2, 45)	1	14 (1, 30)
	Other staff	1	19 (0, 40)	1	14 (0, 30)	1	9 (0, 20)
Delivery	Classroom teacher	1	35 (20, 60)	1	35 (20, 60)	1	35.4 (20, 60)
	Other staff	1	20 (0, 60)	1	20 (0, 60)	1	20 (0, 60)
Peer observation	Other staff	1	3 (0, 16)	1	3 (0, 16)	1	3 (0, 16)
Twilight session	Classroom teacher	1	0 (0, 1,5)	0	0	0	0

Financial costs

Training

The primary training sessions were conducted in seven regional locations as full-day, in-person sessions. The training cost approximately £200 per teacher, although schools were exempt from paying this fee during the trial.[60] Attendance by all Year 5 teachers was mandatory, emphasising the importance of maintaining fidelity in implementation. Group training was conducted at the school level. Nearly all schools in the programme arm of the trial underwent the initial training, and subsequently, a half-day top-up session was conducted online for those schools. Ongoing support, including twilight sessions, was provided through a dedicated online 'Teams' channel, supplemented by email and phone communication. No extra fees were incurred beyond the initial training period.

[60] Intervention schools were only asked to contribute a nominal amount of £100 per school towards the costs of the training and materials they received.

Staff cover

As detailed above, cover was used variably by schools. On average, across all the schools that provided data, an expenditure of seven hours or £175 was allocated for cover. Due to the large majority of schools highlighting the necessity of covering costs, this is regarded as an essential expenditure. Schools considering implementing the PALS-UK programme should factor this into their considerations.

Materials

Additional materials

Schools did not have to pay for the training manual during the trial. Outside of the study conditions, the manual costs £40 per teacher. In the trial, the delivery team supplied books to intervention schools at no charge to ensure that they had minimal resources needed for implementation. Schools were not only asked about their utilisation of existing materials but were also queried about the necessity of acquiring additional resources. In cases where schools obtained extra materials, they were requested to disclose the financial costs involved. Overall, 31 out of 34 schools (91%) reported that no supplementary resources were needed for programme implementation. The remaining three schools (9%) acknowledged the need to purchase additional books but did not specify the associated monetary costs.

Under normal circumstances, schools are not supplied with books and are expected to independently procure the necessary books to implement the programme. According to the delivery team, an average of 20 books per class (with an average size of 25) is necessary for intervention delivery. While schools can use their own books, it is generally recommended for them to make sure that they have a good mix of material, and to refresh existing stock with the introduction of new books. As reported by the delivery team, schools still running PALS-UK spend an average of £100 per class each year on these books, which is equivalent to approximately 15 books annually.

Additionally, schools needed to print training materials for each pupil, which encompassed the Activity Booklet (11 pages), Question Card (2 pages), Correction Card (2 pages), and the Bookmark (1 page), see **Table 22**.

Table 22: Costs of the implementation of the programme, per ingredient per class[61]

			Year 1			Year 2			Year 3		
Category	Cost ingredient	Start-up or recurring?	Mean quantity required (min, max)	Price per unit required (min, max)	Mean cost (min, max)	Mean quantity required (min, max)	Price per unit required (min, max)	Mean cost (min, max)	Mean quantity required (min, max)	Price per unit required (min, max)	Mean cost (min, max)
Personnel for training	Teacher cover	Start-up	7 hours (0 – 13)	£25 per hour	£175 (0 – 325)						
Training costs	Training fee	Start-up	1	£200 per teacher	£200						
	Travel fees for training	Start-up	1 (1 – 2)	£75 per trip	£75 (75 – 150)						
Programme costs	Manual	Start-up	1	£40	£40						
Facilities, equipments, and materials	Books	Pre-requisite	5 books	£7 per book	£35						
Other programme inputs	Printing training materials	Recurring	400 pages per class[62]	0.1 per page	£40	400 pages per class	0.11 per page	£44	400 pages per class	0.12 per page	£48
	New books	Recurring	15	£100	£100	15	£110	£110	15	£120	£120

[61] Multi-form entry schools should multiply figures by the number of classes at Year 5.
[62] Assuming 25 pupils per class.

Overall costs

The comprehensive training, materials, and support for relevant personnel amount to an average cost of £952 per class over a three-year period. The largest component of this expense is the start-up cost, encompassing training fees and associated teaching cover. Other essential costs, such as manuals, printing materials, and books, are also incurred.

Table 23 outlines the total cost per pupil per year for three years based on the trial costs. It is important to note that this calculation assumes an average of 25 pupils per class, and that no additional top-up training is needed after the first year. With a total cost of £12.69 per pupil-school year over three years, PALS-UK is classified as a 'very low-cost programme' (EEF, 2023: p. 28).

Table 23: Recurring costs of the implementation of the programme, per ingredient

Category	Cost ingredient	Start-up or recurring?	Nominal values			
			£ Year 2023	£ Year 2024	£ Year 2025	Total
Personnel for training	Teacher cover	Start-up	£175			£175
Training costs	Training fee	Start-up	£200			£200
	Travel fees for training	Start-up	£75			£75
Programme costs	Manual	Start-up	£40			£40
Other programme inputs	Printing training materials	Recurring	£40	£44	£48	£132
	New books	Recurring	£100	£110	£120	£330
Cost per class per year			**£630**	**£154**	**£168**	**£952**
Number of pupils per class per year			**25**	**25**	**25**	**75**
Total cost per pupil-school year						**£12.69**

Conclusion

Table 24: Key conclusions

Conclusions
1. Children in PALS-UK schools made the equivalent of two months' additional progress in reading, on average, compared to children in other schools. This result has a moderate to high security rating.
2. Children in PALS-UK schools made the equivalent of two months' additional progress in reading comprehension, on average, compared to children in other schools.
3. There were mixed results in relation to the impact of the programme on oral reading fluency, with children making additional progress on one measure of fluency (oral reading fluency rate) but not on another broader measure.
4. Pupils eligible for FSM in PALS-UK schools made the equivalent of one months' additional progress in reading, on average, compared to children eligible for FSM in the controls schools. This result is less secure than the findings based on all children in PALS-UK schools. Further analysis suggests that the impact of PALS-UK on children eligible for FSM was very similar to the impact for other children.
5. Overall, the programme was delivered with high fidelity, although a minority of teachers reported making minor adaptations or considering adapting the programme. In some cases, this was due to concerns around the accessibility of PALS-UK for certain groups, such as children with lower attainment.

Impact evaluation and IPE integration

This study is a two-armed, cluster randomised controlled trial, with whole schools randomised to intervention and control groups. In total, some 114 schools were recruited to the trial and randomised to conditions—57 to the intervention group and 57 to the control group. At analysis, 103 schools remained in the sample—53 in the intervention group and 50 in the control group. At randomisation, there were a total of 4,840 pupils in the sample, 2,450 in the intervention group, 2,390 in the control group. The as 'analysed sample' comprised 3,628 pupils, or 75% of the 'as randomised' sample. In the intervention group 1,907 pupils were retained at analysis, or 78% of the randomised intervention sample. In the control group 1,721 pupils were retained at analysis, 72% of the control group at randomisation. The loss of 1,212 pupils from randomisation to analysis was mainly due to schools leaving the trial (n = 414), or pupils being absent on the day of testing, either at baseline (n = 270) or follow-up (n = 295).

Sample loss gives rise to concerns that results may be biased, while smaller sample sizes also have consequences for the precision of sample estimates. As can be seen, the rate at which pupils were retained in the sample for analysis differed somewhat in the trial arms. Some reassurance can be taken from the fact that in the two arms of the study the samples were relatively well balanced at randomisation, with any observed differences between them at that stage the product of chance rather than systematic bias. Further, the broad pattern of balance in observed covariates by the trial arms in the 'as randomised' and 'as analysed' samples look relatively similar, except for some reduction in the imbalance between the two arms for EAL and FSM pupils, but even here the degree of difference is slight. In general, then, we conclude that the results presented here can be interpreted with a reasonable degree of confidence.

This trial was an efficacy study, and the sample of schools recruited to the trial were a convenience sample. This means caution should be exercised in drawing conclusions as to the likely effectiveness of PALS-UK based on the result of this study for schools and pupils beyond the sample.

The primary impact analysis reveals that pupils in schools assigned to receive PALS-UK saw an improvement in their reading scores equivalent to two months' additional progress relative to their counterparts in control schools.

In the secondary analysis, for fluency outcome measures specifically, results are more mixed. For the oral reading fluency rate measure, obtained from the WIAT-III UK-T assessment, pupils in intervention schools made two months' additional progress relative to pupils in control schools. In contrast, for the reading fluency measure from the MDFS, there was no evidence of pupils in intervention schools making greater progress than their counterparts in control schools. However, WIAT-III UK-T measures rate/accuracy whereas the MDFS measures rate/accuracy combined with expression and volume. This may account for the different findings for the two fluency measures. The MDFS was included as an additional measure of fluency in addition to the fluency rate measure from the WIAT-III UK-T, so that we could explore whether broader aspects of fluency in addition to reading rate (i.e. prosody) were impacted by PALS-UK. The results do not provide any evidence for this.

A third secondary outcome measure was a reading comprehension score derived from the WIAT-III UK-T assessment. Pupils in intervention schools again made two months' additional progress on average in reading comprehension over pupils in control schools.

Turning to subjective measures capturing the relationship with reading among pupils, there is some weak evidence to suggest that pupils in intervention schools reported an improvement in their reading motivation over control school pupils but no evidence that they saw a relative improvement in their reading self-efficacy.

Finally, looking at subgroups within the trial sample, there is evidence that FSM and lower attaining pupils in the intervention group benefited from PALS-UK but not differentially so when compared to non-FSM and higher attaining pupils also in intervention schools. By way of contrast, there is weak evidence that while pupils with SEND in intervention schools benefited from PALS-UK, they may have benefited to a greater extent than pupils not with SEND.

The IPE results are generally consistent with the impact evaluation results. Intervention teachers and headteachers perceived PALS-UK to have a positive impact on pupils' reading attainment, which they felt was in part mediated by gains in reading fluency as well as gains in development of specific reading comprehension skills such as summarising and prediction. Some teachers reported that less able readers may have benefited more from the intervention than the more able readers. This is partially in line with the weak evidence for a differential impact of PALS-UK for pupils with SEND, although it is important to note that teachers did not talk about a differential impact for pupils with SEND specifically, but about the broader category of lower ability readers. Clearly, some lower ability readers will not have SEND and some pupils with SEND will not struggle with reading. It is also important to note that the IPE data suggests that adaptations needed to be made for some pupils with SEND. As discussed in the 'Subgroups' section below, it is difficult to assess the extent to which these adaptations impacted the effectiveness of the intervention for these pupils. The fact that teachers perceived PALS-UK to have had a positive impact on pupils' confidence in reading was not reflected in pupils' self-efficacy scores within the impact evaluation.

Evidence to support the logic model

The findings of both the impact evaluation and the IPE support most of the components of the logic model. In this section we consider what light evidence from both strands of the evaluation shed on various elements of the logic model: training and support activities; teacher and pupil engagement; contextual factors hypothesised to be important for implementation and delivery; as well as outcomes and impacts. The supporting factors that were theorised to be important in enabling outcomes and impacts are also explored.

Training, support, and provision of materials

Key activities of PALS-UK that are inputs into the programme are the training received by teachers both initial training and top-up training, as well as the training in PALS-UK for pupils delivered by teachers.

With this in mind we find:

- Attendance at both initial and top-up training was high and in line with expectations—52 from 57 schools in the intervention group had representatives at initial training and 47 at follow-up. A small number of teachers attended further 'twilight' training.
- The IPE analysis found that the initial and top-up training provided was thorough and equipped schools and teachers to deliver PALS-UK with fidelity, providing teachers with a good grasp of reading pedagogy.
- Online support was also available to teachers.
- All schools that engaged with the teacher training elements of the intervention also delivered the PALS-UK pupil training with high fidelity.
- Teachers that attended training were overwhelmingly positive about it.
- Teachers generally found the ongoing support provided by the delivery team useful, relevant, and timely.
- Further evidence from teacher surveys suggests that over half of teachers in intervention schools agreed that their knowledge of reading instruction had improved, that their teaching of reading had got better, and that they better understood pupils' reading when listening to them read.

Alongside the training and support, teachers were provided with the PALS-UK manual, high-quality contemporary books, and pupil materials such as prompt cards:

- teachers claimed in the main to have taught PALS-UK consistent with the manual and the other materials provided;
- teachers were positive about the manual, regularly using it to guide their teaching, perceiving it to be 'fool proof' and very useful. This helped to ensure that PALS-UK was delivered with fidelity;
- when teachers were asked to identify enablers, they frequently mentioned the materials provided with PALS-UK, particularly the manual and the books;
- teachers felt that they were provided with an excellent range of books, which their pupils had not seen before;
- pupils enjoyed the books; some pupils developed a greater interest in fiction books; and
- pupil materials were considered to be very useful.

In terms of limitations of the materials and barriers in relation to the effectiveness of PALS-UK, teachers reported that:

- the books included within the PALS-UK programme were perceived to be too difficult for some low attaining readers; however, as noted above, teachers are advised in the training to use the books that they already have in school to supplement those provided;
- the intervention was not suitable for all pupils; and
- it was difficult to monitor the pupils' activities when they were all reading different books that they themselves were also not familiar with.

Teacher and pupil engagement (fidelity and full participation in PALS-UK)

For the trial to represent a test of the efficacy of PALS-UK, the programme needs to be delivered by teaching staff with fidelity and to the required dosage. Pupils also need to engage with the programme consistent with programme developers' intentions.

With this in mind, aspects of PALS-UK that promoted engagement included:

- The majority of teachers felt that all four main activities of PALS-UK were either quite or very useful, with partner reading judged to be most useful, followed by paragraph shrinking. Prediction relay was broadly judged to be more difficult and complex to deliver.
- In general, teachers felt that PALS-UK was easy to deliver due to the comprehensive training and materials they received and the structured content of the programme.
- In general, teachers and schools implemented PALS-UK with high, if not excellent, fidelity. Some teachers did mention, however, that they had or were considering adapting some aspects of PALS-UK, though these adaptations in the main were minor.
- Four-fifths of teachers reported delivering the full required weekly sessions over the entire 16-week period of instruction. The pairing and re-pairing of pupils was also adhered to.

Aspects of PALS-UK, which may limit engagement:

- Some teachers expressed concerns over their ability to manage the pairing process effectively.
- One barrier to the effective implementation of PALS was pupil absence from school. Due to the paired nature of the intervention, this was felt to be particularly disruptive when one pupil in a pair was absent.
- Pupils also raised problems with the partnering process, such as not liking their partner, disagreement over the choice of book, and difficulties where one partner refused to engage.
- Timetabling issues were a concern for some teachers particularly where school events occasionally led to a lack of time to deliver PALS-UK.
- There was some evidence that pupil motivation and engagement started to wane towards the end of the study.
- There was also some concern for lower attaining readers and other groups. Evidence from case study schools suggests that low attainers, and pupils with SEND and EAL could not access the programme unless it was heavily adapted by the teacher. Though this does conflict with evidence from the impact evaluation that the intervention might be particularly effective for pupils with SEND. Some children also struggle with the social aspects of PALS-UK. In contrast, some respondents suggested that PALS may not be sufficiently stretching for the highest achievers and that there was a tendency for them to become frustrated.

Context

Teachers and headteachers were asked to reflect on some of the contextual factors that might act as barriers or enablers to effective delivery of PALS-UK. Contextual factors are important as key programme mechanisms might lie dormant without necessary contextual factors present. Factors felt to be important by teachers and school leaders were the general school context and resourcing, senior leadership buy-in, and classroom management.

We find that:

- In bigger schools, staff valued the support they received from their colleagues while in smaller schools teachers found some PALS-UK processes difficult to manage.
- Having protected time in the school day to deliver PALS-UK was widely mentioned as a key enabler of effective implementation.
- Support from the senior management team was a key factor set out in the theory of change. Generally, the level of engagement of senior leaders was mentioned as an important facilitator of the programme by many of our interviewees.
- Headteachers generally said that PALS-UK was well supported. That schools supported teachers in the implementation of PALS-UK, allowed peer observation, time for training, and the planning and delivery of sessions. Eight in ten teachers that delivered PALS-UK and responded to the survey, agreed that they had received support and engagement from senior colleagues.

Outcomes and impacts

Outcomes and impacts need to be interpreted in the context of what other concurrent support for reading was available in intervention schools, and crucially, what support with reading was available in control schools, and how this may have changed over the course of the trial. Any impacts we see are relative to what constituted 'business as usual' support for reading received by pupils in control schools.

With this in mind, we find that:

- Schools in both arms of the trial engaged in similar reading instruction practices at the outset of the study. Most teachers used guided reading, with a lesser proportion, about a quarter already using a form of peer reading.
- There was evidence that teachers in intervention schools substituted PALS-UK for some of their existing practice in line with guidance from the delivery team to replace guided reading (or similar reading activities) with PALS-UK. Outside of PALS-UK lessons, the teaching of reading in intervention schools continued in much of the same way as it did prior to introducing PALS-UK. However, in some of the case study schools, teachers spoke of supplementing PALS-UK in other parts of the week with aspects of reading that they would normally have covered within their guided reading sessions (or equivalent) that were not covered by PALS-UK (e.g. teaching the content domains: inference; impact; and structure).
- 'Business as usual' practice in control schools remained reasonably consistent over the period of the study. There was some evidence that the teaching of reading in control schools bore more than a passing similarity to aspects of PALS-UK. This may have weakened the experimental contrast between intervention and control conditions. In other words, the evaluation implies a test of PALS-UK against a highly active control group; this is unsurprising given the importance of reading in the primary curriculum and its relevance for Key Stage 2 national assessments. Thus, this can further be interpreted as representing a 'high bar' against which we are comparing PALS-UK.

The theory of change hypothesises that certain support factors should be present for PALS-UK to be effective and for the anticipated outcomes and impacts to emerge. These include teachers' knowledge of reading, a positive reading culture across the school, better peer support, and that pupils enjoy reading for pleasure.

With this in mind, we find that:

- The delivery of both initial and follow-on training suggests that teachers delivering PALS-UK had the necessary knowledge and understanding of reading instruction for them to deliver the intervention effectively with fidelity.
- There was evidence that both control and intervention schools took a number of steps and engaged in a wide range of activities to support a positive reading culture in their schools—there is some evidence of this in schools in the intervention group, where PALS-UK was being used. Evidence from the IPE suggests that 57% of teachers agreed that the 'reading culture' in the school had improved as a result of PALS-UK.

- We found that there was a range of activity to encourage reading for pleasure in both control and intervention schools. Survey data from teachers does suggest that this was a particular area of focus for control schools during the intervention.
- Teachers perceived that PALS-UK improved both peer support skills in reading and peer relationships in general. Over 80% of teachers in the survey perceived that PALS-UK had improved pupils' peer support skills.

The theory of change suggests that exposure to PALS-UK will raise attainment in reading through improvements in reading fluency, reading comprehension, reading self-efficacy, and motivation to read.

With this in mind, we find that:

- There is evidence to support the hypothesis that PALS-UK can lead to improvements in pupils' reading comprehension relative to usual practice. We find that pupils in intervention schools made on average two months' additional progress in reading comprehension compared to pupils in control schools. Just over a half of teachers either agreed or strongly agreed with the notion that reading comprehension had improved. It is perhaps surprising that this percentage is not higher given the observed gains. The fact that a significant minority of pupils did not perceive PALS-UK to have improved their reading comprehension may perhaps be explained by the fact that the PALS-UK activities do not explicitly cover all of the strands of reading comprehension, which are measured within the Key Stage 2 reading test.
- Within the impact evaluation, pupils in intervention schools made on average two months' additional progress on the WIAT-III UK-T reading fluency rate measure relative to pupils in control schools. In contrast, we found no effect on reading fluency as measured using the MDFS. As noted above, the MDFS measure incorporates a qualitative measure of reading pace alongside other aspects of fluency: expression and volume; phrasing; and smoothness. Interestingly, there is evidence from the case study interviews with teachers from the intervention schools, that PALS-UK may help pupils to read more quickly and accurately but may not lead to gains in prosody/reading with expression, due to the lack of an adult modelling of these aspects during PALS-UK lessons. Furthermore, over seven in ten teachers surveyed thought that pupils' oral reading fluency had improved; pupils within the case studies also shared this view. Taking the impact evaluation and IPE evidence together, we suggest that PALS-UK has a positive impact on the rate and accuracy of reading but does not appear to impact on expression/prosody.
- Pupils interviewed within the IPE generally felt that their overall reading ability had improved and put improvements down to the volume of reading they undertook.
- We find only weak evidence that improvements in reading motivation are important in bringing about improvements in attainment (through stimulating gains in both fluency and comprehension), while there is no evidence to support the thesis that reading self-efficacy has a role to play. Evidence from the IPE suggests that most teachers' perceived improvements in enjoyment of reading (55%) and self-monitoring of reading (just under 70%). Fewer than half of teachers perceived improvements in informal talk about reading among pupils (40%).
- Overall, the evidence from the impact evaluation does support the claim that PALS-UK will lead to an improvement in general reading attainment among pupils exposed to the intervention relative to pupils exposed to 'business as usual' practice. On average pupils in intervention schools made two months' additional progress in reading over that observed among control pupils. This finding accords with teachers' perceptions of reading, in that around six in ten agreed that reading had improved. There is some evidence that gains in comprehension may play a positive role in mediating the effect of PALS-UK, while the evidence that gains in fluency are important is more mixed.

Subgroups

We considered the effects of PALS-UK on pupils that qualified for FSM, EAL pupils, pupils with SEND, and low attaining readers. Each of these groups of pupils in intervention schools benefited from the programme relative to their counterparts in the control group and in line with the benefits in reading attainment seen across the intervention schools as a whole. There is weak, tentative evidence from the impact evaluation that pupils with SEND may have benefited disproportionately from PALS-UK when compared to non-SEND peers in intervention schools. Peer support approaches may be especially important for pupils with SEND, who otherwise 'might be left working without support over a prolonged period' (Toulia *et al.,* 2023, p. 225). Depending on a pupil's particular needs they may find it harder to stay on task when unsupported than some of their non-SEND peers (ONS, 2021). Given that PALS-UK provides pupils with SEND with one-to-one support throughout the lessons, this may confer a particular advantage to them. On the other hand, it has been suggested that for pupils not with SEND, peer support may lead to less substantial benefits, particularly if the peer tutoring is reciprocal in nature (i.e. if both peers within the pair adopt a tutoring role) (Toulia et al., 2023). This may be

due to the lower attaining reader within the pair being less able to tutor their partner effectively. In fact, this is something that was mentioned as an issue for some pupils within the IPE data.

It is worth noting, however that, in the case studies, teachers spoke of needing to make adaptations to PALS-UK for some pupils, for example, pairing a lower attaining pupil or a child new to English with a teaching assistant. Teachers also highlighted that there were some pupils with particular additional needs who did not access PALS-UK at all as it was not deemed suitable for them. It is important to note that some of the pupils will also have been excluded from the outcome analysis due to not being able to access the New PiRA. Tentatively, it therefore appears that some pupils with SEND were excluded from the intervention, while for others, teachers made successful adaptations, which allowed pupils with SEND to successfully engage with PALS-UK. It is not clear from the data available whether the adaptations made to PALS-UK for pupils with SEND actually enhanced the programme beyond what was being received by pupils not with SEND, or whether the adaptations simply allowed pupils with SEND to access essentially the same provision as was received by their peers. In the former case, the observed (albeit week) interaction effect for SEND would be interpreted as being due to the adaptations rather than the intervention itself; whereas in the latter case, the effect would suggest that there is something inherent to PALS-UK, which makes it particularly beneficial for some pupils with SEND.

Within the IPE survey data, teachers cited lower attaining readers most frequently as the group who they believe benefited the most from PALS-UK, but they did not mention pupils with SEND specifically. Similarly, within the interviews with case study teachers, teachers often spoke about less able readers but did not always distinguish between those with SEND and those without. It is important to note that it can be difficult to neatly separate out the subgroups within teachers' open-end responses to questions. For example, when a teacher suggests that PALS-UK was particularly beneficial for 'less able readers', some of the pupils that the teacher has in mind are also likely to have SEND, and/or be learning EAL, and/or be receiving FSM. Further research with a bigger sample is recommended to attempt to replicate the observed (weak) differential effect for pupils with SEND, including investigation of potential differential impacts of PALS-UK on pupils across different subcategories of SEND to provide a clearer picture of who does and does not benefit from the intervention. Further in-depth study of which adaptations teachers make for pupils with SEND and their impact in attainment is also warranted.

Unintended outcomes and consequences

The unintended pupil outcomes, which were suggested within the IPE data included an increase in informal 'book talk'. While the sessions themselves do not allow time for this, one case study school suggested that outside the lessons, pupils were talking about the books informally. The excitement around the books was also apparent in the pupil interviews. However, it was also noted by some teachers that one of the downsides of PALS-UK was that there was no space within the actual lessons for this social, sharing of books, as advocated within a 'community of reading' approach (e.g. Cremin, 2019). Another unexpected pupil outcome suggested by the case study data, was the development of metacognitive awareness around the reading process and an understanding of how the PALS-UK activities supported different aspects of their learning.

We also found that there were unintended outcomes in terms of teachers' practice. On the positive side, teachers talked about how PALS-UK had impacted their practice in other areas of the curriculum including the skills developed from 'retell' to focus the content of a paragraph, and the use of the ten-word 'paragraph shrinking' strategy to summarise knowledge in other subjects such as religious education and science. On the negative side, one teacher mentioned a small amount of time being lost from the timetable due to needing to move children around into their PALS-UK pairs.

Interpretation

This study represents the first successful attempt in the UK to replicate the positive findings from evaluations of PALS carried out in North America. The What Works Clearance House in the United States conducted a review of two North American studies of PALS, which reach their study quality thresholds (WWC, 2012). Evidence from Fuchs *et al.* (1997) showed an average effect size for reading fluency of 0.31 and 0.60 for reading comprehension, both considered 'substantively important'. Similarly, a study from Saenz *et al.* (2005) reported an effect size of 0.41 for fluency and 0.91 for comprehension. In the current study, there is likewise evidence that reading among pupils exposed to PALS-UK improved relative to the control group. For the primary reading attainment outcome an effect size of 0.12 (95% CI: 0.01 to 0.23) was obtained, while for the two oral reading fluency measures effect sizes of 0.14 (95% CI: 0.01 to 0.27) and 0.04 (95% CI: -0.13 to 0.20) were found. The effect size for reading comprehension was found to be 0.16 (95% CI: 0.02 to 0.31). Only one of the effect sizes had a 95% CI, which contained zero (the multidimensional fluency measure). Relative to North American studies, effect sizes reported here are somewhat smaller. It is worth pointing out that the

sample sizes for both the United States studies reported above were small, n = 40 and n = 20, respectively, and effect sizes did not reach statistical significance at the 95% confidence level. Large effect sizes obtained in small samples are known to be prone to exaggeration relative to the true effect sizes (see Button *et al.*, 2013 for a discussion).

This present study follows an unsuccessful attempt at replication by an evaluation team based at RAND Europe and the same delivery team as the current trial (Culora *et al.,* 2022). The lack of success in relation to the evaluation was caused by challenges related to the onset of the Covid-19 pandemic and the consequent disruption to schools. The evaluation was disrupted in three key ways: the overall attrition rate was high and the attrition rate differed across the control and treatment groups; delay of post-test measures, leading to a potentially diluted effect of PALS-UK on the primary reading attainment measure; and incomplete collection of the secondary reading comprehension and fluency measures. Within the current study we have been able to address these shortcomings, including improved implementation based on lessons learned by the delivery team in terms of successful elements of the previous attempts. Within the present study, we were able to investigate the impact of PALS-UK on all primary and secondary outcome measures, taken within a few weeks of the intervention ending. The current study is a rigorously conducted two-armed, cluster randomised controlled trial and we can have a fair degree of confidence in its findings. For the perspective of fidelity, this study represents a good test of PALS-UK against an active control group.

Broadly, our findings support the programme theory of change. Alongside gains in reading attainment, we found gains in reading comprehension and oral reading fluency (WIAT-III UK-T measure only). This aligns with the predictions of the theory of change that PALS-UK would support reading attainment by helping children to read both fluently and for meaning. These findings agree with research in the United States, which found that PALS had the potential to improve reading comprehension and reading fluency (WWC, 2012), and with evidence that other interventions involving peer reading, where gains in reading comprehension and reading fluency were also found (Algozzine *et al.,* 2009; Marr *et al.,* 2011).

The fact that children in intervention schools did not outperform children in control schools on the MDFS measure suggests that PALS-UK improves reading speed and accuracy but may not increase prosody/reading with expression. Alternatively, the lack of an effect on the MDFS fluency measure may be due to the measure being a qualitative, subjective assessment of fluency, and therefore, potentially less sensitive when compared to the WIAT-III UK-T. Teachers' concerns about PALS-UK not providing an adult role model and/or peer feedback for prosody/use of expression, could provide a plausible reason for why we see expected rate and accuracy improvements but not gains in expression/prosody. To our knowledge, this is the first study to test the impact of a peer reading intervention on a multidimensional measure of fluency, and there is a dearth of studies, which investigate effects of reading interventions on reading prosody more generally (Hudson *et al.,* 2020). Further research into the potential for such interventions to improve elements of fluency other than rate/accuracy (or not) are therefore warranted.

This study provides weak evidence for PALS-UK impacting on children's motivation for reading and finds no evidence that reading self-efficacy is affected by the programme. In this way we add to the research evidence, given that in the previous disrupted trial of PALS-UK, reading self-efficacy, but not motivation for reading was analysed. Interestingly, within the IPE just over half of the teachers (55%) agreed or strongly agreed that pupils' interest in reading had increased as a result of PALS-UK, and a high proportion of teachers agreed/strongly agreed (70%) that PALS-UK had improved pupils' ability to self-monitor their reading—an important aspect of self-efficacy. Within the case study data, both teachers and pupils generally felt that PALS-UK had had a beneficial impact on both motivation for reading and reading self-efficacy. The lack of an observed quantitative effect of PALS-UK on reading self-efficacy may be due to a potential misalignment between the particular self-efficacy constructs within the reading self-efficacy questionnaire and the actual reading self-efficacy constructs that were potentially impacted by the intervention. Self-efficacy is 'highly domain specific' (Carroll and Fox, 2017); given that reading ability is multi-faceted it is possible that some aspects of pupils' reading self-efficacy impacted by PALS-UK may not have been fully captured within the questionnaire. For example, while the questionnaire asks about pupil's belief in their ability to read without making lots of mistakes, it does not ask them about their ability to notice when they *are* making mistakes, something, which teachers reported to have increased. Given the complexities of measuring reading self-efficacy and the mixed results across the impact evaluation and IPE, further research into the impact of PALS-UK on reading self-efficacy is warranted.

Another aspect of the logic model, which was supported by the findings was the prediction that PALS-UK would promote peer support skills and peer relationships. While no direct measures of these constructs were taken, both teachers and peers perceived PALS-UK to have had a positive effect on both of these elements, with one case study teacher describing the social impact of PALS as being 'absolutely massive'. While teachers discussed the importance of (and sometimes difficulties with) ensuring that pupils were paired appropriately, overwhelmingly teachers felt that the peer-

coaching element of PALS-UK has benefits to pupils beyond their reading skills. The fact that some of the case study teachers talked about particular benefits of this dimension of the intervention for pupils exhibiting 'challenging' behaviour aligns with other studies, which found that PALS-UK reduced disruptive behaviour and increased the engagement of pupils exhibiting emotional or behavioural issues (Locke and Fuchs, 1995; Sutherland and Snyder, 2007).

The apparent success of PALS-UK as an intervention to accelerate reading attainment, aligns with evidence reported within the EEF Teaching and Learning Toolkit (EEF, 2021). The EEF Guidance Report on 'Improving Literacy in Key Stage 2' (Bilton and Duff, 2021) recommends a strong emphasis on developing pupils' fluency through guided instruction and repeated reading. This study extends this evidence base, demonstrating the efficacy/feasibility of fluency instruction delivered through a peer-based approach; however, the integrated impact evaluation/IPE findings suggest that adult modelling may also be required for children's multidimensional fluency (including use of expression) to improve. Alternatively, one teacher suggested that perhaps other elements of fluency could be integrated into PALS-UK in the form of prompts for pupils to feedback to each other on their use of expression, volume, etc. The importance of explicitly teaching for prosody as well as accuracy is also emphasised by Rasinski and others (e.g. Rasinski, 2012) as it forms an important part of the 'bridge from word recognition accuracy to text comprehension' (Rasinski, 2012, p. 519). This study's findings also aligns with the recommendation within the EEF Guidance Report (Bilton and Duff, 2021) to teach specific reading comprehension strategies 'that pupils can apply to monitor and overcome barriers to comprehension' (Bilton and Duff, 2021, p. 10), including prediction and summarising (retell)—two of the activities within PALS-UK, which most teachers felt were either quite useful or very useful. On the other hand, the IPE findings suggest that PALS-UK could perhaps be strengthened further if it incorporated some of the other reading comprehension strategies, which teachers noted are measured in the SATs test at the end of Key Stage 2 but are not explicitly addressed within PALS-UK (e.g. inference). Similarly, while vocabulary is not a strategy as such, it is a key competency, which skilled readers draw upon to make meaning from a text (Rasinski, 2012), which some teachers felt was missing from PALS-UK.

Teachers' reports of the high-quality books being a key strength of PALS-UK aligns with the EEF recommendation that texts should be carefully selected and evidence from other studies that a range and breadth of authors and genres is needed to provide pupils with rich reading experiences, support reading progression, and develop communities of readers (e.g. CLPE, 2023; Cremin *et al.,* 2014). It was noted within the case studies, however, that for the benefits of PALS-UK to be maximised, teachers needed to have a good knowledge of the books being used to facilitate monitoring of the children's progress in the moment.

The fact that there was some evidence to suggest that PALS-UK may have been especially beneficial for pupils with SEND aligns with the recommendation within the EEF guidance (Bilton and Duff, 2021) that highlight structured interventions may be particularly helpful for pupils who might struggle with literacy. A number of teachers in the IPE reported that PALS-UK benefits pupils as they know 'exactly what is coming'; although it was suggested that interest waned towards the end of the intervention, especially for some higher attaining readers. The study therefore, warrants further research into the optimal length of PALS-UK to maximise the potential benefits of a highly scaffolded approach, while maintaining the engagement of all pupils.

In summary, this study provides evidence that PALS-UK improves reading attainment, which may be mediated by gains in reading comprehension, oral reading fluency, and possibly also motivation to read. The qualitative evidence suggests that PALS-UK may have a beneficial impact on peer relationships and reading self-efficacy, although further research is needed to support this in the absence of evidence of a quantitative effect. This is the first successful replication of PALS in the UK. Further replication should be attempted through new samples of schools and pupils before decisions to move to an effectiveness trial and scale-up are taken. However, the results for the current study are promising. They suggest that PALS-UK is a cost-effective whole-class reading intervention, which is easy to deliver, affording it strong potential in terms of scalability. Within the current policy landscape, where the key emphasis lies on teachers using whole-class 'Quality First Teaching' (DfE, 2015) approaches to meet the needs of all pupils, and where staffing resources for group/individual interventions are limited, the peer-led nature of this intervention is a key strength. On the other hand, there were notable concerns among teachers that PALS-UK in its current form does not cover all the skills required to prepare pupils for the Key Stage 2 reading test. This suggests that it may be important to flag to schools choosing to use PALS-UK that it is not designed to cover all the aspects of reading that will be assessed in the national assessments, which are a key driver in relation to planning, teaching, and assessment within Key Stage 2 classes currently. Schools therefore, may want to ensure that they are covering the remaining reading content domains in the rest of their reading provision at other points in the week. Importantly, schools mentioned lots of benefits, which related to some of the reading content domains as well as some additional benefits, including developing positive identities in relation to reading (e.g. helping to promote a positive culture around reading/towards books) and developing pupils' independence/ability to work effectively with others. It is noted that one of the side benefits of PALS-UK reported by teachers, namely the

impact on peer relationships, may be particularly desirable in the current post-pandemic climate, where there are widespread concerns over a decline in young people's social and emotional development (e.g. Catton *et al.,* 2023).

Limitations and lessons learned

As is invariably the case with randomised trials conducted in 'field' settings, this trial faced challenges, and its findings are subject to qualification and caveat. The impact evaluation did suffer from a failure to retain schools and pupils in the sample. Of the 114 schools randomised, 103 were retained in the sample 'at analysis' (88% of schools assigned to the control group and 93% of schools assigned to the intervention group). Among pupils randomised, 75% were retained for analysis, 78% of intervention group pupils and 72% of control group pupils. Despite sample retention being lower in the control group, based on measured variables, this sample loss does not appear to have led to appreciable imbalances between intervention and control groups. Nonetheless, some bias in our analysis resulting from unmeasured factors cannot be ruled out.

In future, consideration needs to be given to additional steps towards retaining schools in experimental samples. This is particularly in light of the funding constraints faced by schools, higher levels of staff and pupil absences, and the general high level of strain on school resources. All these factors mitigate against initial school recruitment as well as school and pupil retention, and are particularly telling where studies, such as this one, rely heavily on school engagement with both baseline and endline assessment. The fact that the trial was conducted just after the COVID-19 pandemic, when schools were still adjusting to the 'new normal', may have been an additional barrier to recruitment and retention. Given the differences in retention across the control and intervention groups, further investigation is warranted into the factors that may lead to attrition in each condition, and that different approaches might be necessary to support retention of control versus intervention schools.

While the study was designed as an efficacy trial, and therefore the focus was a test of the intervention under more tightly specified conditions; and further, the study sought to achieve a test of the underlying theory of change embodied in PALS-UK; concern nonetheless naturally extends to questions around how far results can be generalised. The sample of schools recruited to this study was not selected at random from the population of primary schools; effectively schools selected themselves into the study. Furthermore, due to conditions placed on the funding of this trial by the DfE through its Accelerator Fund, schools had to be recruited from three English Regional School Commissioner regions of the North, East Midlands and the Humber, and West Midlands. However, the schools in our sample are similar to the population of all primary schools in term of levels of prior aggregate attainment at Key Stage 2; though in general, schools in this trial had slightly higher proportions of pupils qualifying for FSM than all schools, and schools allocated to control in this trial had on average marginally higher proportions of EAL pupils than the national average.

As highlighted in previous sections of this report, 'business as usual' in control schools around the teaching of reading implied a fairly active, although relatively stable, set of 'control conditions'. Although realistic in the sense that control school activities can be considered to represent reading instruction more broadly in schools, it is worth keeping in mind that the results we have described here represent a relative comparison with these active control conditions, and thus quite a high bar, albeit relevant one, against which to compare the effectiveness of PALS-UK.

There is often concern expressed over the extent to which test administration practices can lead to questions over the reliability of findings. At baseline, prior to randomisation, New PiRA assessments were collected from schools online. Some schools struggled with the technology required to complete assessments. Broadly, we were able to use nearly all responses received online, however, we do not know whether the quality of the data received may have been affected. For this reason and to help those schools with poorer technology, at endline New PiRA, we moved to paper and pencil self-completion. While this eliminated the issues associated with the use of technology (e.g. limited internet access in some schools, and pupils clicking through the pages too quickly), paper administration brought its own problems, for example, papers being returned without the pupil identification stickers on and pupils missing out pages accidentally. Furthermore, at endline, due to budgetary constraints, New PiRA assessments were overseen in schools by teachers. This means that endline assessments were not administered 'blind' to intervention/control group status. Where a lack of 'blinding' in test administration is felt to be important, it is crucial that budgets for data collection are adequate such that independent test administration can be achieved. Furthermore, consideration might be given to possibly consequences of mixed-mode administration of assessments and whether this would be at all feasible and valid in the context of a trial.

Turning attention next to the limitations and challenges faced in the IPE. The IPE consisted of baseline and endline surveys of teachers and headteachers. In total, the sample comprised some 174 classes (91 in the intervention group and 83 in the control group) and 114 schools. Thus, if we had achieved full response to our surveys, we might expect a sample of 174 teachers (one teacher per class) and 114 headteachers (one headteacher per school). We received 128

survey returns from intervention schools (teacher and headteacher combined) and 108 from control schools at baseline, and 98 and 78, respectively at follow-up. Very roughly a response rate of 86% (128 out of 91+57) was achieved in the intervention group and 77% (108 out of 83+57) in the control group at baseline; while at endline, the rates were 66% and 56%, respectively. These rates of response are fairly typical,[63] but the level of response and fact that survey response rates differed in the two arms of the trial raises questions as to the representativeness of these data, and thus the reliability of findings. Some consideration, therefore, needs to be given to how teachers and school leaders can be incentivised to take part in surveys and whether different approaches are needed for intervention schools versus control schools.

Finally, the case study data come from in-depth fieldwork conducted in six schools. It proved quite challenging to recruit case study schools, presumably due to the perceived additional burden of participating in case studies. As such, it is difficult to determine how far findings from case studies hold more broadly.

In terms of the implementation of PALS-UK, the evidence suggests that this was done consistent with the design intentions of the programme and to a high degree of fidelity. At the margin, there were some issues around teachers needing to make adaptations to PALS-UK and a sense that some pupils, possibly higher attaining children, tended to disengage towards the end. Overall, however, these are no major concerns.

Future research and publications

The evaluation and delivery team intend to publish the results of this trial in a high impact peer reviewed journal. Further joint journal articles and papers are anticipated. These outputs will seek to address questions around the relative role gains in comprehension and fluency might have in mediating the effects of PALS-UK, using formal statistical techniques of mediation analysis. Other future research will explore the various sub-elements of the MDFS measure, notably: expression and volume; phrasing; smoothness; and pace to examine whether effects on these subscales may have been masked in the aggregate. Analysis will also attempt to separate out aspects of the MDFS measure that capture prosody and fluency, to examine whether the results obtained through the WIAT-III UK-T measure of fluency can be replicated once prosodic elements are removed from the MDFS measure. Reading self-efficacy is difficult to measure but potentially important, one piece of future research could look at enhancing a measure of self-efficacy to better align its items with specific competencies we might expect PALS-UK to develop.

[63] For example the EEF evaluation of the School Partnership Programme had baseline survey responses from 80% of their treated schools, available at: SPP-Evaluation-report.pdf (d2tic4wvo1iusb.cloudfront.net).

References

Algozzine, B., Marr, M.B., Kavel, R. L. and Dugan, K.K. (2009) 'Using Peer Coaches to Build Oral Reading Fluency'. *Journal of Education for Students Placed at Risk,* 14: 3, 256–70.

Allen, R., Jerrim, J., Parameshwaran, M. and Thompson, D. (2018) 'Properties of Commercial Tests in the EEF Database'. EEF Research Paper, (001). London, UK: EEF.

Bilton, C. and Duff, A. (2021) 'Improving Literacy in Key Stage 2. Guidance Report. Second Edition.' London, UK: EEF. Available at: https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/literacy-ks2 (accessed 08 September 2024).

Boylan, M., Demack, S., Wolstenholme, C., Reidy, J. and Reaney, S. (2018) 'ScratchMaths: Evaluation Report and Executive Summary'. London, UK: EEF.

Braun, V. and Clarke, V. (2006) 'Using Thematic Analysis in Psychology'. *Qualitative Research in Psychology*, 3: 2, 77–101.

Breadmore, H.L., Vardy, E.J., Cunningham, A.J., Kwok, R.K.W. and Carroll, J.M. (2019) 'Literacy Development: Evidence Review'. London, UK: EEF. Available at: https://educationendowmentfoundation.org.uk/education-evidence/evidence-reviews/literacy-development (accessed08 September 2024).

Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinsonan, E.S.J., and Munafò, M.R. (2013) 'Power Failure: Why Small Sample Size Undermines the Reliability of Neuroscience'. *Nature Reviews,* 14: 365–76.

Carroll, J.M. and Fox, A.C. (2017) 'Reading Self-Efficacy Predicts Word Reading But Not Comprehension in Both Girls and Boys'. *Frontiers in Psychology,* 7: 2056.

Catton, S., Farquharson, C., Krutikova, S., McKendrick, A. and Sevilla, A. (2023) 'How Did Parents' Experiences in the Labour Market Shape Children's Social and Emotional Development During the Pandemic? London, UK: Institute for Fiscal Studies. Available at: https://ifs.org.uk/publications/how-did-parents-experiences-labour-market-shape-childrens-social-and-emotional (accessed 08 September 2024).

Centre for Literacy in Primary Education (CLPE). (2023) 'Power of Reading Impact Report'. London, UK: CLPE. Available at: https://clpe.org.uk/research/power-reading-impact-report-2023 (accessed 08 September 2024).

Clark, C. and Rumbold, K. (2006) 'Reading for Pleasure: A Research Overview'. London, UK: National Literacy Trust.

Cremin, T. (2019). 'Reading Communities: Why, What and How? *NATE Primary Matters*, Summer 2019'. Sheffield, UK: NATE.

Cremin, T., Mottram, M., Collins, F.M., Powell, S. and Safford, K. (2014) 'Building Communities of Engaged Readers: Reading for Pleasure'. Abingdon, UK: Routledge. https://doi.org/10.4324/9781315772585

Cullinanae, C. and Montacute, R. (2020) 'COVID-19 and Social Mobility. Impact Brief #1: School Closures'. London, UK: The Sutton Trust. Available at: https://www.suttontrust.com/wp-content/uploads/2021/01/School-Shutdown-Covid-19.pdf (accessed 08 September 2024).

Culora, A., Dimova, S., Ilie, S, Sutherland, A. and Gilder, L. (2022) 'Peer Assisted Learning Strategies – UK Evaluation Report'. London, UK: EEF.

Demack, S. (2019) 'Does the Classroom Level Matter in the Design of Educational Trials? A Theoretical & Empirical Review'. London, UK: EEF.

Department for Education (DfE). (2015) 'Special Educational Needs and Disability Code of Practice: 0 to 25 Years'. London, UK: DfE. Available at: https://www.gov.uk/government/publications/send-code-of-practice-0-to-25 (accessed 08 September 2024).

Department for Education (DfE). (2021). 'School pupils and their characteristics: 2020/21. Explore Education Statistics.' London, UK: DfE. Available at: https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics/2020-21 (accessed 01 July 2023).

Department for Education (DfE). (2023). 'The reading framework.' London, UK: DfE. Available at: https://assets.publishing.service.gov.uk/media/664f600c05e5fe28788fc437/The_reading_framework_.pdf (accessed 10 July 2024).

Department for Education (DfE). (2024) 'Key Stage 2 Attainment, Academic Year 2022/23'. Available at: https://explore-education-statistics.service.gov.uk/find-statistics/key-stage-2-attainment (accessed 08 September 2024).

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

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

Education Endowment Foundation (EEF). (2019) 'Cost Evaluation Guidance for EEF Evaluations'. London, UK: EEF. Available at: https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluation/evaluation-guidance-and-resources/evaluation-design(accessed 08 September 2024).

Education Endowment Foundation (EEF). (2021) 'The EEF Teaching and Learning Toolkit'. London, UK: EEF. Available at: https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/peer-tutoring/ (accessed 08 September 2024).

Education and Endowment Foundation (EEF). (2023) 'Attendance and Reading Key Barriers to Disadvantaged Pupils' Progress Say Three in Four Schools'. Available at: https://educationendowmentfoundation.org.uk/news/attendance-and-reading-key-barriers-to-disadvantaged-pupils-progress-say-three-in-four-schools (accessed 08 September 2024).

Fuchs, D., Fuchs, L.S., Mathes, P.G. and Simmons, D.C. (1997) 'Peer-Assisted Learning Strategies: Making Classrooms More Responsive to Diversity'. *American Educational Research Journal,* 34: 1, 174–206.

Gorard, S., Siddiqui, N., See, B.H., Smith, E. and White, P. (2017) 'Children's University: Evaluation Report and Executive Summary'. London, UK: EEF.

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: 52.

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R. and Kerr, K. (2016) 'Implementation and Process Evaluation (IPE) for Interventions in Education Settings: An Introductory Handbook'. London, UK: EEF.

Humphrey, N., Hennessey, A., Ashworth, E., Frearson, K., Black, L., Petersen, K., Wo, L., Panayiotou, M., Lendrum, A., Wigelsworth, M., Birchinall, L., Squires, G. and Pampaka, M. (2020) 'Good Behaviour Game. Evaluation Report and Executive Summary'. London, UK: EEF.

Jay, T., Willis, B., Thomas, P., Taylor, R., Moore, N., Burnett, C., Merchant, G. and Stevens, A. (2017) 'Dialogic Teaching: Evaluation Report and Executive Summary'. London, UK: EEF.

Klauda, S.L. and Guthrie, J.T. (2008) 'Relationships of Three Components of Reading Fluency to Reading Comprehension'. *Journal of Educational psychology,* 100: 2, 310.

Levin, H.M., McEwan, P.J., Belfield, C., Bowden, A.B. and Shand, R. (2017) 'Economic Evaluation in Education: Cost-Effectiveness and Benefit-Cost Analysis'. Thousand Oaks, CA: SAGE Publications, Inc.

Lloyd, C., Edovald, T., Kiss, Z., Morris, S., Skipp, A. and Ahmed, H. (2015) 'Paired Reading: Evaluation Report and Executive Summary'. London: EEF. Available at: https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/paired-reading (accessed 08 September 2024).

Locke, W.R. and Fuchs, L.S. (1995) 'Effects of Peer-Mediated Reading Instruction on the On-Task Behavior and Social Interaction of Children with Behavior Disorders'. *Journal of Emotional and Behavioral Disorders,* 3: 2, 92–9. https://journals.sagepub.com/doi/abs/10.1177/106342669500300204 (accessed 08 September 2024).

Lucas, M., Classick, R., Skipp, A. and Julius, J. (2023) '*Cost-of-Living: Impact on Schools'.* Slough, UK: National Foundation for Educational Research. Available at: https://www.nfer.ac.uk/media/lzjhu1ie/cost_of_living_impact_on_schools_staff.pdf (accessed 08 September 2024).

Marr, M.B., Algozzine, B., Nicholson, K. and Dugan, K.K. (2011) 'Building Oral Reading Fluency with Peer Coaching'. *Remedial and Special Education,* 32: 3, 256–64. https://doi.org/10.1177/0741932510362202

Morris, S.P., Edovald, T., Lloyd, C. and Kiss, Z. (2016) 'The Importance of Specifying and Studying Causal Mechanisms in School-Based Randomised Controlled Trials: Lessons from Two Studies of Cross-Age Peer Tutoring. *Educational Research and Evaluation,* 22: 7–8, 422–39.

Moser, G.P., Sudweeks, R.R., Morrison, T.G. and Wilcox, B. (2014) 'Reliability of Ratings of Children's Expressive Reading'. *Reading Psychology,* 35: 1, 58–79. doi:10.1080/02702711.2012.675417.

Office for National Statistics (ONS). (2021) 'Educational Experiences of Young People with Special Educational Needs and Disabilities in England: February to May 2022'. London, UK: ONS. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/educationandchildcare/bulletins/educationalexperiencesofyoungpeoplewithspecialeducationalneedsanddisabilitiesinengland/februarytomay2022 (accessed 08 September 2024).

Ofsted. (2023) 'Getting All Pupils Reading – Ofsted: Schools and Further Education & Skills (FES). Available at: https://educationinspection.blog.gov.uk/2023/08/30/getting-all-pupils-reading/ (accessed 08 September 2024).

O'Hare, L., Stark, P., Cockerill, M., Lloyd, K., McConnellogue, S., Gildea, A., Biggart, A., Connolly, P. and Bower, C. (2019) 'Reciprocal Reading: Evaluation Report'. London, UK: EEF.

Paige, D.D., Rasinski, T., Magpuri-Lavell, T. and Smith, G.S. (2014) 'Interpreting the Relationships Among Prosody, Automaticity, Accuracy, and Silent Reading Comprehension in Secondary Students'. *Journal of Literacy Research,* 46: 2, 123–56.

Peura, P., Aro, T., Viholainen, H., Räikkönen, E., Usher, E.L., Sorvo, R. and Aro, M. (2019) 'Reading Self-Efficacy and Reading Fluency Development Among Primary School Pupil: Does Specificity of Self-Efficacy Matter?' *Learning and Individual Differences,* 73, 67–78.

Pikulaki, J.J. and Chard, D.J. (2011) 'Fluency: Bridge Between Decoding and Reading Comprehension'. *The Reading Teacher,* 58: 6, 510–19.

Rasinski, T.V. (2003) 'The Fluent Reader: Oral Reading Strategies for Building Word Recognition, Fluency, and Comprehension'. New York, NY: Scholastic Inc.

Rasinski, T.V. (2004). 'Creating Fluent Readers'. *Educational Leadership,* 61: 6, 46–51.

Rasinski, T.V. (2012) 'Why Reading Fluency Should Be Hot!' *The Reading Teacher,* 65: 8, 516–22. https://doi.org/10.1002/TRTR.01077

Rasinski, T.V., Padak, N.D., McKeon, C.A., Wilfong, L.G., Friedauer, J.A. and Heim, P. (2005) 'Is Reading Fluency a Key for Successful High School Reading?' *Journal of Adolescent & Adult Literacy,* 49: 1, 22–7.

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

Rudd, P., Aguilera, A.B.V., Elliott, L. and Chambers, B. (2017) 'MathsFlip: Flipped Learning. Evaluation Report and Executive Summary'. London, UK: EEF.

Saenz, L., Fuchs, L.S. and Fuchs, D. (2005) 'Peer-Assisted Learning Strategies for English Language Learners'. *Exceptional Children*, 71: 3, 231–47.

Sutherland, K.S. and Snyder, A. (2007) 'Effects of Reciprocal Peer Tutoring and Self-Graphing on Reading Fluency and Classroom Behavior of Middle School Students with Emotional or Behavioral Disorders'. *Journal of Emotional and Behavioural Disorders,* 15: 2, 103–18.

Smith, G.S. and 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

Støle, H., Mangen, A. and Schwippert, K. (2020) 'Assessing Children's Reading Comprehension on Paper and Screen: A Mode-Effect Study'. *Computers & Education,* 151, 103861.

Topping, K., Millder, D., Thurston, A., McGavock, K. and Conlin, N. (2011) 'Peer Tutoring in Reading in Scotland: Thinking Big'. *Literacy*, 45: 1, 3–9.

Toulia, A., Strogilos, V. and Avramidis, E. (2023) 'Peer Tutoring as a Means to Inclusion: A Collaborative Action Research Project'. *Educational Action Research*, 31: 2, 213–29. https://doi.org/10.1080/09650792.2021.1911821

Vardy, E.J., Breadmore, H.L. and Carroll, J.M. (Under review). 'Measuring the Will and the Skill of Reading: Validation of the Self-Efficacy and Motivation to Read Scale'.

Weschler, D. (2018) 'Weschler Individual Achievement Test – Third UK Edition for Teachers: WIAT-III UK-T'. London, UK: Pearson.

WWC. (2012). 'Peer-Assisted Learning Strategies'. WWC Intervention Report. Washington DC: WWC. Available at: https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_pals_060512.pdf (accessed 08 September 2024).

Yildirim, K., Rasinski, T., Ates, S., Fitzgerald, S., Zimmerman, B. and Yildiz, M. (2014) 'The Relationship Between Reading Fluency and Vocabulary in Fifth Grade Turkish Students'. *Literacy Research and Instruction*, 53: 1, 72–89.

Appendix A: EEF cost rating

Table 25: Cost rating

Cost rating	Description
£ £ £ £ £	*Very low:* less than £80 per pupil per year.
£ £ £ £ £	*Low:* up to about £200 per pupil per year.
£ £ £ £ £	*Moderate:* up to about £700 per pupil per year.
£ £ £ £ £	*High:* up to £1,200 per pupil per year.
£ £ £ £ £	*Very high:* over £1,200 per pupil per year.

Appendix B: Security classification of trial findings

Rating	Criteria for rating			Initial score		Adjust		Final score
	Design	MDES	Attrition					
5	Randomised design	<= 0.2	0-10%					
4	Design for comparison that considers some type of selection on unobservable characteristics (e.g. RDD, Diff-in-Diffs, Matched Diff-in-Diffs)	0.21 - 0.29	11-20%					
3	Design for comparison that considers selection on all relevant observable confounders (e.g. Matching or Regression Analysis with variables descriptive of the selection mechanism)	0.30 - 0.39	21-30%	3		Adjustment for threats to internal validity **[0]**		3
2	Design for comparison that considers selection only on some relevant confounders	0.40 - 0.49	31-40%					
1	Design for comparison that does not consider selection on any relevant confounders	0.50 - 0.59	41-50%					
0	No comparator	>=0.6	>50%					

Threats to validity	Risk rating	Comments
Threat 1: Confounding	Low	Allocation was independently conducted. Imbalance at pre-test <0.05. Baseline scores included as covariates.
Threat 2: Concurrent Interventions	Low	One peer reviewer rated as moderate as there was indication that some intervention schools instigated other interventions at the same time as PALS, but the other peer reviewer felt there was no clear evidence of differential uptake between the groups and that concurrent interventions were comprehensively explored. On balance, this seems a low threat to the validity of the findings.
Threat 3: Experimental effects	Moderate	One peer reviewer rated as moderate, stating that there is perhaps some evidence that the intervention schools reduced time spent on other reading activities and control schools increased their focus on reading. The other reviewer rated as low with no comments.
Threat 4: Implementation fidelity	Low	Implementation and compliance is well defined, aligned to the logic model and is high in the intervention group. Intervention was delivered with fidelity.
Threat 5: Missing Data	Moderate	Pupil attrition was considerable, and there was imbalance between the groups, but missing data analysis yields similar results to complete case analysis. School attrition was lower.
Threat 6: Measurement of Outcomes	Low	Although issues were reported with digital administration of the tests at baseline, this was rectified by using paper tests at endline and no ceiling or floor effects were reported. Tests were marked blinded to allocation and the outcome measures used are reliable and valid.
Threat 7: Selective reporting	Low	Protocol and SAP cited, and trial registered. No evidence of selective outcome reporting. A very detailed report has been produced.

- **Initial padlock score:** 3 Padlocks – An intact RCT, sufficiently powered to 0.212 but there was significant attrition at the pupil level between randomisation and analysis (25%) which loses the trial 2 padlocks.
- **Reason for adjustment for threats to validity**: [0] - No adjustments as there is no evidence of serious threats to the internal validity of the study.
- **Final padlock score:** 3

Appendix C: Effect size estimation

Appendix Table D1: Effect size estimation for primary and secondary outcomes

			Intervention group		Control group		
Outcome	Unadjusted differences in means	Adjusted differences in means	n (missing)	Variance of outcome	n (missing)	Variance of outcome	Pooled variance
Endline PiRA score (primary)	0.97 (-0.42, 2.35)	1.02 (0.12, 1.92)	1,907 (403)	69.97	1,721 (378)	79.09	74.56
Oral reading fluency (rate) (secondary)	1.44 (0.19, 2.69)	1.16 (0.80, 2.25)	506 (1,804)	66.64	493 (1,606)	77.53	72.52
Oral reading fluency (multidimensional) (secondary)	0.19 (-0.38, 0.75)	0.10 (-0.38, 0.58)	512 (1,798)	8.12	497 (1,602)	8.44	8.29
Reading comprehension (secondary)	1.81 (-0.03, 3.64)	1.47 (0.16, 2.77)	467 (1,843)	69.22	455 (1,644)	92.29	81.37
Reading self-efficacy (secondary)	0.23 (-2.00, 2.46)	0.35 (-1.67, 2.39)	1,563 (747)	403.97	1,462 (637)	380.90	392.69
Motivation for reading (secondary)	0.84 (-0.56, 2.25)	0.89 (-0.33, 2.11)	1,561 (749)	165.32	1,458 (641)	168.90	167.13

Appendix Table D2: Effect size estimation for subgroup analysis

			Intervention group		Control group		
Subgroup (outcome: PiRA endline score)	Unadjusted differences in means	Adjusted differences in means	n (missing)	Variance of outcome	n (missing)	Variance of outcome	Pooled variance
FSM subgroup	-0.59 (-1.71, 0.54)	0.02 (-0.77, 0.82)	606 (158)	62.90	597 (143)	76.26	69.70
SEND subgroup	-0.04 (-1.38, 1.31)	0.97 (-0.02, 1.96)	279 (111)	76.25	273 (80)	80.65	78.67
Lowest quartile baseline PiRA subgroup	0.21 (-0.86, 1.27)	0.22 (-0.65, 1.08)	454 (101)	44.84	418 (70)	48.78	47.36
Lower half PiRA and EAL subgroup	0.97 (-0.82, 2.76)	0.26 (-1.00, 1,51)	207 (28)	53.28	235 (31)	68.86	62.02

Further Appendices

Please see separate technical report with further appendices:

Appendix D: Recruitment documents

Appendix E: Research instruments

Appendix F: Further details about usual practice

Appendix G: Randomisation and Primary Analysis Code

Appendix H: Statistical outputs

This document is available for download at https://educationendowmentfoundation.org.uk

Education Endowment Foundation

The Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP

https://educationendowmentfoundation.org.uk
@EducEndowFoundn
Facebook.com/EducEndowFoundn