Education
Endowment
Foundation

Paired Reading

Evaluation report and Executive summary

July 2015

Independent evaluators:

Cheryl Lloyd, Triin Edovald, Zsolt Kiss, Stephen Morris, Amy Skipp and Hashim Ahmed (NatCen Social Research)

NatCen
Social Research

The Education Endowment Foundation (EEF)

The Education Endowment Foundation (EEF) is an independent grant-making charity dedicated to breaking the link between family income and educational achievement, ensuring that children from all backgrounds can fulfil their potential and make the most of their talents.

The EEF aims to raise the attainment of children facing disadvantage by:

- Identifying promising educational innovations that address the needs of disadvantaged children in primary and secondary schools in England;

- Evaluating these innovations to extend and secure the evidence on what works and can be made to work at scale;

- Encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust, as lead charity in partnership with Impetus Trust (now part of Impetus-The Private Equity Foundation) and received a founding £125m grant from the Department for Education.

Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.

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

Robbie Coleman
Research and Communications Manager
Education Endowment Foundation
9th Floor, Millbank Tower
21-24 Millbank
SW1P 4QP

p: 020 7802 1679
e: robbie.coleman@eefoundation.org.uk
w: www.educationendowmentfoundation.org.uk

About the evaluator

The project was independently evaluated by a team from NatCen Social Research. The impact evaluation was led by Cheryl Lloyd and Triin Edovald who were assisted by Zsolt Kiss and Stephen Morris. The process evaluation was designed and overseen by Amy Skipp who was assisted by Hashim Ahmed.

A team from the University of Bristol, led by Paul Clark managed the evaluation project through set up and randomisation for the trial.

NatCen would like to thank Emily Tanner and Gemma Lewis for their support and contribution to the project, as well as all the schools who gave up their time to participate in our interviews.

Contact details:

Emily Tanner
Head of Children, Families & Work
NatCen Social Research
35 Northampton Square
London

Tel: 0207 549 8580
Email: Emily.Tanner@natcen.ac.uk

Contents

Executive summary

Paired Reading is a peer tutoring programme in secondary schools which trains teachers to support and encourage the regular tutoring of Year 7 pupils (aged 11-12 years) by Year 9 pupils (aged 13-14 years). The Paired Reading programme aims to improve pupils' general literacy in addition to speaking and listening skills. This is achieved by pupils working together to follow the Paired Reading steps to choose the material to read, and discuss this, together with the older pupil (tutor) supporting the reading, correcting errors and praising the younger pupil (tutee) throughout.

The 16-week programme is intended to take place during normal school hours in timetabled sessions, for 20 minutes each week. Teachers in participating schools received training from the delivery team, a detailed programme manual and extensive digital resources.

The impact of Paired Reading on 2,736 pupils in 120 classes in ten participating schools (1,370 in Year 7 and 1,366 in Year 9) was tested using a cluster randomised controlled trial design with 58 classes randomly allocated to receive the programme and 62 classes allocated to the control condition. Schools from the North Tyneside local authority (LA) and in neighbouring LAs of South Tyneside and Sunderland took part in the trial over the 2013/14 academic year.

Key Conclusions
1. This evaluation does not provide any evidence that the Paired Reading programme had an impact on overall reading ability, sentence completion and passage comprehension of participating pupils.
2. There was no evidence of the Paired Reading programme having an effect on overall reading ability, sentence completion and passage comprehension of FSM pupils.
3. There was some variation in the intervention group schools in terms of the programme set-up and delivery. There was also a varying level of support provided to pupils within the intervention by the teachers involved, based mainly on the reading ability of the pupils. However, these appear to be natural variations between the settings of the schools involved and are unlikely to have affected the dosage of the intervention for the pupils involved.

What impact did it have?

There is no evidence from this evaluation that the Paired Reading programme had any positive impact on overall reading ability of participating pupils, measured using the New Group Reading Test (NGRT) scores (see key findings for the primary outcome, reading ability, below).

The process evaluation suggested that there were some differences amongst participating schools with regards to the set-up, delivery and implementation of the programme. There was also a varying level of support provided to pupils within the intervention by the teachers involved, based mainly on the reading ability of the pupils.

Security rating

> *Security rating awarded as part of the EEF peer review process*

Findings from this evaluation have moderate to high security. The study was set up as a randomised controlled trial, which aimed to compare the progress of Year 7 and 9 pupils who received the programme to similar pupils who did not. The trial was classified as an effectiveness trial, meaning that it sought to test the intervention in realistic conditions in a large number of settings.

The study was a large and well-conducted study. All classes allocated to intervention or control arms completed the study. As the randomisation took place at class level this does introduce the possibility of an exchange of information between teachers and pupils in different trial groups. The process evaluation did not indicate information was exchanged.

At the end of the intervention period all pupils were asked to complete the New Group Reading Test, as a standardised measure of general reading ability. As the testing was administered by the schools the padlock rating is reduced to 4.

How much does it cost?

The cost of the Paired Reading programme is estimated to be £10.50 per pupil. This estimate includes training and support costs and materials required for schools to be able to deliver the programme. These estimates are based on a class of 25 pupils (and five classes per school).

Group	Effect size	Estimated months' Progress	Security rating	Cost
Intervention vs. control (Year 7)	-0.02	-1[1]	🔒🔒🔒🔒🔓	**£**
Intervention vs. control (Year 9)	-0.06	-1	🔒🔒🔒🔒🔓	**£**
Free school meal pupils (Year 7)	-0.04	-1[2]	–	–
Free school meal pupils (Year 9)	-0.06	-1	–	–

---

[1] *Since this report was published, the conversion from effect size into months of additional progress has been slightly revised. If this result was reported using the new conversion, it would be reported as 0 months of additional progress rather than -1. See **here** for more details.*

[2] *See footnote 1.*

Introduction

Intervention

Paired Reading is a Peer Tutoring programme in secondary schools which trains teachers to support and encourage regular tutoring of Year 7 pupils (aged 11-12 years) by Year 9 pupils (aged 13-14 years). The Paired Reading programme aims to improve pupils' general literacy in addition to speaking and listening skills. This is achieved by pupils working together to follow the Paired Reading steps to choose the material to read, and discuss this together with the older pupil (tutor) supporting the reading, correcting errors and praising the younger pupil (tutee) throughout.

In Paired Reading, pupils in Year 7 are paired with pupils in Year 9 using recent reading assessment scores and spend 20 minutes each week reading together, over 16 weeks. Pupils are trained and supported by their teachers to follow clear procedures for each session. Pupils select reading materials (e.g. books, magazines) above the independent reading ability of the tutee, then spend a short period reading together before the tutee begins to read aloud on their own. The tutor then focuses on error correction. Each time the tutee makes a mistake the tutor will wait four or five seconds before intervening with the correct word. In addition this Paired Reading process includes a meta-cognitive element developed through the use of questioning and feedback by the pair.

The Paired Reading programme is used in conjunction with the school's English curriculum. It is not intended to replace other English/literacy teaching but rather is incorporated into lessons to support pupil's reading development. Teachers in participating schools received training from the delivery team, a detailed programme manual and extensive digital resources (Thurston & Cockerill, 2013).

Schools from North Tyneside local authority (LA) in North East England and in neighbouring LAs of South Tyneside and Sunderland took part in the trial.

Background evidence

Peer tutoring is a student-led instructional strategy used to support improved academic achievement and social-emotional outcomes of pupils. Research suggests that pupils involved in peer tutoring show higher academic achievement, improved relationships with peers, improved personal and social development and increased motivation. The evidence for peer tutoring tends to be overwhelmingly positive. For example, a meta-analytic review summarising the findings from 65 independent evaluations of school tutoring programmes by Cohen, Kulik & Kulik (1982) showed that these programmes have positive effects on the academic performance and attitudes of those who receive tutoring but also on tutors. A more recent review by Rohrbeck and colleagues (2003) looking at comparison design studies evaluating peer-assisted learning interventions with elementary school students showed that such interventions increased pupils' academic achievement. The review also indicated that interventions that used interdependent reward contingencies, ipsative evaluation procedures, and provided students with more autonomy showed more positive effects.

A recent meta-analytic review by the Washington State Institute for Public Policy (WSIPP) (Pennucci & Lemon, 2014) explored the effects of both cross-age and same-age/class wide peer tutoring.[3] The

---

[3] In the two evaluations included in the meta-analysis of cross-age peer tutoring, the average cross-age peer tutoring programme provides 30 hours tutoring time and 7.5 hours of training time per class. In the evaluations included in the meta-analysis of same-age peer tutoring, pupils from the same classrooms provide academic assistance to struggling peers. Same-age tutoring assistance occurs through one-on-one interactions or in small groups, and in some instances, students alternate between the role of tutor and tutee. The same-age peer tutoring programmes included in the WSIPP's meta-analysis provide, on average, 30 hours of peer tutoring time each year and about five hours of training time for teachers and students to learn programme procedures.

results of respective meta-analytic reviews indicated that both types of programmes had a positive effect on participant test scores. However, the strength of the evidence was found to be stronger for same-age and class wide peer tutoring when compared cross-age peer tutoring in the WSIPP inventory of evidence- and research-based practices. The meta-analytic review on cross-age peer tutoring included only two relatively old studies (1980 and 1983) that included limited information on the demographic characteristics of the pupils (for further details on the criteria applied to the included studies please see Pennucci & Lemon, 2014[4]).

Poor literacy outcomes continue to affect many young people in England, including at the transition between Primary and Secondary school and this is most acute for pupils living in areas of multiple disadvantage. Gaps in attainment widen between pupils known to be eligible for Free School Meals (FSM) and their peers through Key Stage 2, and persist during Key Stage 3 at secondary school (DfE, 2014a). Whilst reading tends to show smaller differences in progress by FSM status, this is still of concern to policy makers and practitioners (DfE, 2014a).

The EEF and Sutton Trust Pupil Premium Toolkit (Higgins, Katsipataki, Coleman et al., 2014)) included a review of the effectiveness of different teaching and learning strategies and concluded that based on current evidence peer tutoring was a potentially effective approach to improve attainment for both the tutors and tutees – who have to think carefully about how to support the tutees. The evidence of impact is relatively high, typically equating to about a GCSE grade, and there is some evidence that children from disadvantaged backgrounds and low attaining pupils make the biggest gains.

A large-scale efficacy trial in Fife (Tymms, et al., 2011) developed the peer tutoring approach in primary school maths and reading using the Paired Reading technique for the latter. The study showed that cross-age peer tutoring had a consistent positive impact on attainment in both math and reading for both the younger and older cohorts (and produced better results when compared to same-age tutoring). The study had limitations including researchers not being blind to treatment allocation, loss of schools and pupils during the project, schools changing their intervention group and a differential loss to follow up among the older cohort. Despite limitations, this study suggested that peer tutoring has promise when scaled up.

Recently, a research team from Durham University modified the existing 'Fife' Cross-Age Peer Tutoring model for maths. The impact of this cross-age peer tutoring pedagogy called Durham Shared Maths was tested in 82 primary schools in four English local authorities to assess how effective it is at raising the attainment of disadvantaged pupils when implemented on a large scale and delivered by non-specialists in peer tutoring. This effectiveness trial, however, did not provide evidence of an impact on attainment in maths as the primary outcome (Lloyd, et al., *in review*). The results of this trial are expected to be published in May 2015.

This EEF effectiveness trial of Paired Reading builds on the Fife experience by delivering the programme to secondary school pupils in the North East of England to assess how effective it is at raising the literacy attainment of disadvantaged pupils, shortly after the transition to secondary education.

Evaluation objectives

The aim of the impact evaluation was to measure the impact of the Paired Reading programme on the reading ability attainment of participating pupils:

- in Year 7 (the tutee)

---

[4] For combined meta-analytic results on class-wide peer tutoring, peer-assisted learning strategies and reciprocal peer tutoring by WSIPP please visit the following link: http://www.wsipp.wa.gov/BenefitCost/Program/107

- in Year 9 (the tutor)
- who are eligible for FSM in both tutee and tutor year groups
- who are in the special educational needs (SEN) category in both tutee and tutor year groups
- who have English as an additional language (EAL) in both tutee and tutor year groups
- by looking separately at girls and boys in both tutee and tutor year groups

The aims of the process evaluation were to:

- Understand, at a strategic level, what attracted schools to sign up to the programme.
- Understand the process of setting up the programme in the school- including pairing the pupils and finding space in the timetable.
- Explore how the programme worked in practice and any implementation, fidelity or sustainability issues.
- Find out how teachers and senior leaders in schools perceived the programme to be making (or not making) an impact – whether on the pupils or on the school at a wider level.
- Examine how the teachers who delivered the programme in the schools felt the programme could be improved if it were to be implemented more widely.
- Look into issues around sustainability of the programme.

Project team

The programme was implemented by a delivery team led by Maria Cockerill of North Tyneside Council, in partnership with Allen Thurston and Nicole Craig of Queen's University, Belfast, with design collaboration from Christine Merrell of Durham University.

The independent evaluation was set up by a team from the University of Bristol, led by Paul Clark who managed the project through set up and randomisation for the trial. However, Paul Clark left the University of Bristol and the EEF re-tendered for the evaluation work. For the remaining stages of the project a team from NatCen Social Research were the independent evaluators. The impact evaluation was led by Cheryl Lloyd who was assisted by Triin Edovald, Zsolt Kiss and Stephen Morris. The process evaluation was designed and overseen by Amy Skipp, assisted by Hashim Ahmed.

Ethical review

NatCen Social Research obtained ethical approval from its own ethics board for the evaluation, comprising the process evaluation and analysis of test results. This approval included the processes for the research team communicating with and carrying out interviews with school staff.

This trial has been registered at **www.controlled-trials.com** with the trial number ISRCTN03012788.

Methodology

Trial design

In order to assess the impact of the Paired Reading programme on the reading ability of participating children, a cluster randomised controlled trial with assignment at class level was carried out in ten participating secondary schools. Classes at these schools were assigned at random into two trial groups – intervention and control (see Randomisation section below for more details). Pupils in intervention classes participated in Paired Reading while pupils in control classes received 'business as usual' (BAU). Thus, pupils in the control group continued to operate as usual. Schools participated in the project from January to June 2014.

The design study is a pragmatic randomised controlled trial. Namely, the schools participating in the study were the ones whose head teachers had volunteered to be part of the trial and that included classes where the logistics of the standard school day allowed the implementation of peer tutoring.

Eligibility

Ten secondary schools in North East England were recruited to take part in the trial, selecting schools where the proportion of pupils eligible for free school meals (FSM) was above the national average. Further details on recruitment are presented in the section on Participants below. All Year 7 and Year 9 classes in recruited schools were eligible to participate in the study. However, due to the nature of the Paired Reading programme, an equal number of Year 7 (tutee) and Year 9 (tutor) classes of similar size were required in each participating school. Therefore, some of the classes in Year 7 or Year 9 group were excluded from the trial. For further details see Randomisation section below.

The head teachers at participating schools signed a Memorandum of Understanding (MoU) with the delivery and evaluation teams, agreeing that they consented on behalf of their school to randomisation, pupil testing and running the Paired Reading programme. Following this agreement, each school sent a letter to the parents of all pupils in Year 7 and 9 which was provided by the delivery team. This letter explained that the school was participating in the Paired Reading programme, including details of the intervention, pre- and post-testing and asked parents to inform the school if they did not wish their child to take part (see Appendix A for the parent Opt out letter).

Intervention

The Paired Reading programme is a cross-age peer tutoring programme in secondary schools which involves older Year 9 pupils tutoring younger Year 7 pupils. The participating children work together to follow the Paired Reading steps to choose the material to read which can be a book, magazine or any other written document, and discuss this together with the older pupil (tutor) supporting the reading, correcting errors and praising the tutee throughout. Paired Reading has been developed from a number of earlier peer tutoring projects, originating from Keith Topping and colleagues' work at Dundee University.

The Paired Reading programme aims to improve pupils' general literacy in addition to speaking and listening skills. This is achieved by the tutor and tutee working together to read text aloud, with the older pupil correcting errors and the two pupils discussing the content of the text to deepen their understanding of the issues raised and critically thinking about these which aims to encourage reading. The two pupils are required to listen to each other, with the older pupil praising the tutee as appropriate throughout which develops speaking and listening skills (Thurston & Cockerill, 2013).

The programme follows the steps:

1. The younger pupil, who is acting as the tutee, chooses a book which is above their independent reading ability.
2. The younger and older pupils spend a short time reading the book aloud together.
3. The older pupil then stops reading aloud while the younger pupil continues reading on their own.
4. The older pupil, who is acting as the tutor, will then focus on error correction.
5. When the tutee struggles to or cannot pronounce a word, the tutor waits for five seconds before intervening with the correct pronunciation.
6. Both the tutee and tutor begin reading together again.
7. In addition to correcting pronunciation, the tutor and tutee will question one another, before, during and after the reading. The questions could be comprehension based and focused on the book that the tutee has chosen to read.
8. The tutor and tutee provide ongoing assessment through feedback using verbal feedback recorded in a structured log book.

This provides a formal system of ongoing peer assessment through feedback.

The Paired Reading delivery team were responsible for providing teacher training for both Year 7 and Year 9 teachers and on-going support throughout the programme. (Further details on both teacher and pupil training can be found below.)

Prior to the start of the programme, teachers in participating schools received training from the delivery team and a detailed programme manual. The Peer Tutoring in Secondary Schools manual (Thurston & Cockerill, 2013) provided background information about the programme, how this links to the National Curriculum and how it should be delivered to pupils in their school. (Please see Process evaluation – Role of the teacher for further details.).

As set out in the manual, the Paired Reading programme is intended to take place during normal school hours in timetabled sessions. Pupils in Year 7 are paired with pupils in Year 9 using recent reading assessment scores and spend 20 minutes each week reading together, over 16 weeks. Both tutors and tutees are trained and supported by their teachers to follow clear procedures for each session.

Teacher training

Prior to the start of the programme, teachers from participating schools received a day of training, covering the rationale and background to the intervention, as well as the practical aspects of delivering the programme on a day to day basis. The training covered the processes for training pupils, pairing pupils, selecting books for reading, supported reading, error correction, questioning and praise. Attendees were provided with a manual advising them on techniques and the approach. A further ½ day, was provided four weeks into the programme for teachers to further improve their practice and to consolidate teacher school teams for the duration of the programme. Further materials were provided during the second training event in respect of enhancing communication skills between peer tutors and tutee pupils. In most cases Year 7 and Year 9 teachers attended the training. However, if another subject teacher was responsible for delivering the Paired Reading programme in the school, this teacher attended instead. Following the training, participating schools were provided with a resource pack containing the training slides, other background information about the intervention and other practical materials required, such as log books and question mats. In addition, teachers were provided with access to a website containing all online resources, including training films and teacher and student training PowerPoint resources for sessions 1-16.

Contact time

It was recommended that schools set aside 30-minute periods for the programme with 20 minutes of this dedicated to quality reading time. The other 10 minutes were planned to provide time for pupils to

move between classrooms and getting them settled for reading. Schools were advised to conduct the intervention during normal school hours, replacing another lesson in the pupils' timetable. One school was unable to do so due to timetable constraints and, instead, brought pupils into school early in order to conduct the intervention, replacing the normal registration period.

The intervention was designed to be used at any time during the school day (not specifically during English lesson classes). 52 classes delivered the intervention during normal school hours and 6 classes delivered half of every session before the school day and the other half at the start of the day - consecutively. Of the 52 classes, 20 received the intervention during English class time.

Pupil training

The Paired Reading manual recommends that schools train both tutors and tutees together to ensure that both received the same information about the Paired Reading programme. The programme manual advises that the first four sessions should be used as follows to train pupils and ensure all pairs were confident with how the programme works:

1. The first session for explaining the background and rationale as well as any starter activities with the pairs working together.
2. The second session for practising praise and error correction.
3. The third session to introduce questioning.
4. The fourth session to correct any areas of weakness and consolidate learning to date.

Pupil training took place either with Year 9s and Year 7s together or separately. PowerPoint resources were provided for the 16 sessions, with separate distinct PowerPoint slideshows for sessions 1 to 4 to incorporate the training element for pupils and a further generic PowerPoint for sessions 5-16. There was some variation across schools in terms of how the training took place and in terms of the training length; for example, some schools used only the first lesson to train the pupils, while others provided a staged introduction to the programme. Some schools only trained Year 9 pupils and there was an incident where one school did not carry out any advance pupil training (for further details see Process Evaluation - Training the pupils section below).

Pairing classes and pupils

Year 7 and Year 9 classes in the intervention group were paired up based on reading ability, using the pre-intervention test scores [Standard Age Scores from the digital New Group Reading Test (NGRT)].[5] The evaluation team at NatCen were responsible for recommending which pairs of classes from each Year group should work together for the Paired Reading programme. This was done by ranking classes within each school according to their average test scores, each class in Year 7 was then matched with the same ranking class in Year 9. These matches were then checked to ensure that the demographic characteristics of pupils were similar in terms of gender, FSM, EAL and ethnicity.

Teachers were responsible for pairing pupils from Years 7 and 9 within paired classes assigned to the intervention group. Pupils were ranked in the order of reading levels based on their pre-intervention test scores (see Outcomes section for more details about the tests). Pupils in each pair of Year 7 and 9 classes were then paired with one another, with the top pupil in Year 7 paired with the top pupil in Year

---

[5] The preferred option of pairing classes and individual pupils is based on a recent reading test to enable the best match between Year 9 tutors and Year 7 tutees in respect of reading age. For the purposes of this project the pre-intervention test score was used, however, a specific assessment is not required for the Paired Reading programme so another appropriate, recent measure of reading ability would normally be used.

9, the second pupil in Year 7 paired with the second pupil in Year 9, and so on. Other than reading levels, schools were advised to consider other factors when pairing pupils, such as gender, maturity, working habits, cultural background, personality and attendance.

If there was an uneven number of pupils in Year 7 and 9, then some pupils were allocated into 'trios'. The trios could either be with two tutors and a tutee or two tutees and a tutor. Poor attendance was recommended as a factor for inclusion into a trio, to ensure there was at least an opportunity for pupils to read in pairs during every session (even in the absence of a member of a trio).

Training the teachers

The training for participating staff was delivered at a local council run education provider centre. It was delivered over two separate sessions, totalling one and a half days for each participating teacher. Session 1 lasted a full day and Session 2 was a follow up two to three weeks into the programme delivery, which lasted half a day. The teachers who would be delivering the intervention, from all participating schools, received the training but not all at the same time; rather they attended separate sessions in order to minimise the number of teachers away from the school at once. Teachers commented that receiving the training along with teachers from other schools was a positive aspect as it helped to widen their professional understanding and offered the opportunity for them to learn from others.

'It is always good to meet other teachers and learn from them and their experiences'

(Teacher)

The format of the training was that teachers:

- Were given an introduction to the programme, including the rationale behind the intervention
- Went through how the programme works and what is expected of the teachers and pupils
- Were put in pairs and practiced the programme as if they were pupils receiving the intervention.

All participants were given a support pack that contained resources and materials explaining how to deliver the programme. This included a copy of the presentation slides and the materials needed for the programme; such as question cards, question mats and log books. In addition, teachers were given access to all resources, including films and detailed PowerPoint presentations for sessions 1-16 of the programme. There was ongoing technology support available from the delivery team to ensure all online resources were accessed by teachers, and where teachers required support the information was sent via email to anyone who requested it.

Teachers spoke very positively about the training, in particular its more practical aspects. The practice activities were felt to provide a deeper understanding of the programme and gave teachers a clearer idea of what it would involve when implemented in their classrooms.

'That (the practice activity) was helpful because it made it a lot easier to explain it to our kids'

(Teacher)

It was also felt that the training was complimented well by the support pack. The materials provided were felt to cover the practical issues teachers were likely to encounter as well as setting out what the sessions should involve and answering any outstanding questions they had about the process.

'Once I received my pack with all the information, I was pretty happy that I knew what to do with it all'

(Teacher)

Teachers appreciated the 'hands-on' format of the training as they felt it helped them get experience of the approach required, before getting their pupils to carry it out. This helped them to be better prepared

to deliver the programme once it began. They practiced the programme step by step, including selecting a book, reading together, correcting, questioning and praising.

It wasn't just sitting and watching someone read from a PowerPoint…there was a chance to have a go at some of the things that were part of the programme'

(Teacher)

However, there were suggestions for how the training could be improved further:

- Some of the resources contained grammatical errors, which did not create a good impression.
- Some teachers experienced difficulties with electronic materials and downloading these from a dedicated website. Due to these difficulties some of the schools ended up creating their own resources, rather than rely on the ones provided. However these problems were not reported elsewhere and were by no means the prevailing experience.

'The whole thing could have been sent out by email. The whole scheme has been really time consuming'

(Teacher)

- The training could have been reduced into a day, or even half a day, or delivered as departmental time within their schools. There was concern about the amount of time the programme delivery took up overall and reducing the training length would have helped to alleviate this.

Selecting reading text

Schools were instructed to allow the tutee to select their reading material. Teachers were encouraged to help pairs try different genres of text and ensure the books were at the right level of reading difficulty. Pupils tended to select a book and continue reading the book until it was finished, at which point they would select a new one.

The 'Five Finger Test' illustrated below (source: Thurston & Cockerill, 2013) was recommended to select a book. Ideally, the book would be above the independent reading ability of the tutee but below that of the tutor to facilitate the tutor being able to help the tutee as far as possible.

Five Finger Test

Choosing 'Just Right' books
Use the Five Finger Test

- Tutor or tutee: put 5 digits on the page
- Tutee: read the words under each digit
- Repeat the process over four pages

Add up the total number of mistakes:

- 1 or less = book is too easy
- 2 – 3 = book is 'Just Right'
- 4 or more = book is too hard

Paired Reading process

Once books have been selected the approach then encourages pupils to follow the steps set out in detail below (Thurston & Cockerill, 2013):

1. **Supported reading**

The pairs begin to read together, with the tutor helping the tutee by modulating the speed of their reading to be just behind the tutee. After a short while, the tutor stops reading with the tutee continuing to read alone until a mistake was made. At that point, the pair would begin reading together again for another short while.

2. **Error correction**

When a tutee makes an error in pronunciation, the tutor waits for 4 or 5 seconds, providing the tutee with the opportunity to self-correct. Should this self-correction not occur, the tutor then says aloud the correct pronunciation of the word, which the tutee repeats. The tutor then gives praise and the pair starts reading together again.

3. **Questioning**

Questioning is considered to be one of the programme's key aspects as it meant both tutors and tutees needed to think about and understand the book they were reading. The questioning involves both tutors and tutees asking each other questions about the book before, during and after the reading. Question mats with suggested questions are provided to help with this, although pupils were also encouraged to think of their own questions, particularly as the intervention progressed. The questions generally covered:

Before the reading

- The reasons for choosing the book
- What will happen in the book
- Whether they think they will like the book

During the reading

- Who the favourite characters are
- What the book is about
- What is likely to happen next

After the reading

- Whether they enjoyed the story and why
- Whether they would recommend it to others
- Whether the story ended as expected

The question mats provide questions of varying difficulty so that higher ability pupils can ask more difficult questions while the lower ability pupils ask questions that are easier for them to understand. There were three levels of question mats provided:

Basic

Standard

Advanced (1 and 2)

Teachers were instructed to use their own judgement to decide which mat to provide each pair with at the start of the paired reading project. They would then, whilst observing the pairs, make a decision about when to move them onto a more (or indeed less) advanced question mat as the programme progressed.

4. Praise

Tutors are encouraged to praise the tutees, for example when the tutee reads a difficult word or sentence, when the tutee pronounces a corrected word and when the tutee has read a section on their own. Tutors are provided with a praise card with words such as 'good reading' and 'brilliant expression' written on them and are encouraged to use these, along with their own praise words.

5. Log books

Both tutors and tutees are provided with log books in which to write after each session. The log allows pupils to note how far they had read into the book with space for both tutees and tutors to write how the tutee's reading has improved from the previous session.

Control classes

The control group classes followed 'business as usual' during the trial and did not receive any Paired Reading materials or training. 'Business as usual' included interventions such as Literacy Clubs, additional reading support groups, withdrawal groups, specific 'Literacy' lessons and dedicated reading time (please see Other literacy and reading interventions in the Process evaluation section for further details on 'business as usual').

Outcomes

Primary and secondary outcomes

The New Group Reading Test (NGRT), which was developed by GL Assessment and the National Foundation for Educational Research (NFER) was used to measure the outcomes of interest. The primary outcome of this trial was reading ability that was measured using the Overall Reading Scale of the NGRT that combines two NGRT subscales: the Sentence Completion Scale and the Passage Comprehension Scale. These two subscales of NGRT were used to measure the two secondary outcomes: sentence completion and passage comprehension.

The NGRT was administered by computer, using a computer adaptive version which assesses pupil performance during the test and adapts the questions asked to be appropriate for the pupil's reading level.

Pupil responses from each of these modules were used to calculate scale scores for sentence completion, passage comprehension and overall reading ability. Furthermore, Standard Age Scores (SAS) score was calculated for reading ability.[6] The SAS is based on the pupil's raw test score which has been adjusted for age and placed on a scale that makes a comparison with a nationally representative sample of students of the same age across the UK. The average score is 100, with pupils scoring between 90 and 100 being considered 'average', those achieving scores of over 100 being 'above average' and those scoring under 90 performing 'below average'.

For further details on specific scales and scores used in the analysis please see the Analysis section below.

Administration of the NGRT

All Year 7 and Year 9 pupils in participating schools undertook the pre-test in September 2013 and the follow-up test in June 2014. The NGRT was completed online and administered to pupils in groups or as a class, in exam conditions.

The tests were administered by schools with some support from the Paired Reading delivery team in respect of technical set-up and not assistance with test completion. Teachers and the Paired Reading delivery team were aware of the pupils' group allocation. As the computer administered version of the NGRT is an adaptive assessment, tailored to each individual pupil according to their abilities, the assessment marking is automated, and thus blinded to pupils treatment allocation status.

Prior to test administration, the staff member responsible for Information Technology (IT) at each school uploaded pupil Unique Pupil Numbers (UPN) and background information to the NGRT GL Assessment website. Then, shortly after testing was complete the delivery team were able to securely access test scores and pupil background information. Once all testing was complete the data were provided to the evaluation team for matching to the National Pupil Database (NPD) and analysis.

Sample size

The sample size calculations that were used to inform the size of the trial were conducted by Bristol University before NatCen were involved in the project. NatCen have been unable to adequately ascertain the parameters and the design characteristics that were used in the calculation. So for the sake of clarity we have conducted a post-hoc power calculation for the number of schools and pupils actually recruited.

Following the collection and processing of trial data, minimum detectable effect sizes (MDESs) for the primary outcome were calculated more precisely, using actual values for the ICC obtained from an unadjusted analysis[7] (see below). The ICC for Year 7 data were generally higher than those obtained from Year 9. The proportion of variance explained at Level 2 – the class – subsequent to including covariates in the adjusted analysis was very high at 0.90, and at Level 1 - the pupil - around 0.60. MDESs at randomisation and as analysed are the same for both Year 7 and Year 9, 13 per cent of one standard deviation, and 11 per cent respectively.

---

[6] http://www.gl-assessment.co.uk/sites/gl/files/images/Files/NGRTsub-scalepaper.pdf

[7] The unadjusted primary analysis includes all pupils who had post-test outcome data.

Table 1: MDESs for Year 7 and Year 9 at randomisation and at analysis

| | As randomised | | As analysed | |
	Year 7 Classes N=60	Year 9 Classes N=60	Year 7 Classes N=60	Year 9 Classes N=60
Probability level	0.05	0.05	0.05	0.05
ICC	0.18	0.06	0.18	0.06
Average cluster size	23	23	22	21
R-squared (Level 2 variance explained)	0.90	0.90	0.90	0.90
Power	80	80	80	80
Effect size	0.13	0.11	0.13	0.11

Randomisation

All Year 7 and Year 9 classes in participating schools were eligible for randomisation. However, as indicated above, an equal number of Year 7 (tutee) and Year 9 (tutor) classes of similar size were required in each participating school to deliver the Paired Reading programme. In order to achieve an equal or an approximately equal number of pupils (within classes) in both year groups included in the trial, all classes were stratified by a year group (Year 7 and Year 9), and listed alphabetically (e.g. 7a, 7b, 9a, 9b) or numerically (e.g. 7/1, 7/2, 9/1, 9/2). The year group with the smaller number of classes in each school set the limit of classes in both year groups to be included in randomisation. Any additional classes in either Year 7 or Year 9 group were excluded before randomisation.

In one of the participating schools pupils were separated by academic ability into English classes. More specifically, classes were streamed in 5 classes in Year 7 and 4 classes in Year 9. In both year groups, one class was half the size of the other classes and composed of the lowest ability pupils. Year 7 had four classes of similar size (1, 2, 3, 4) with classes 2, 3, 4 being of similar ability and class 5 with smaller numbers in the lowest ability set. Year 9 had three classes of similar size (1, 2, 3), with classes 2 and 3 being of similar ability, and class 4 with a smaller number of pupils showing lowest ability. In this school, although the classes in each year group were first listed as described above, Year 7 English class 4 was excluded pre-randomisation to ensure that the four possible streamed classes in each year group were of similar size. This was part of the trial design to ensure a similar number of pupils would be paired between classes as tutors or tutees.

Randomisation was conducted as follows. Once an even number of Year 7 and Year 9 classes in each school was achieved, two groups were created, group A and group B, by going through the list of Year 7 and Year 9 classes in each school and alternately assigning a class first to group A and then group B. The classes in group A and group B were then randomly assigned to either intervention or control group. A member of the Bristol University evaluation team carried out the randomisation procedure by tossing a coin. When the randomisation was conducted, 'heads' was identified as intervention and 'tails' as control. As the result was 'tails', the first class available on the list was allocated to the control group. From then all subsequent available classes were allocated alternately to control group or intervention group. The evaluator carrying out the randomisation procedure declared the outcome over the phone to the delivery team.

See section Pairing classes and pupils above for more details on how classes and pupils were paired post-randomisation.

Analysis

The analysis was conducted in STATA version 13 (Stata Corporation, College Station, Texas, USA). Impacts were estimated on the basis of intention to treat, whereby all classes and pupils were analysed

according to the study arm to which they were initially assigned, regardless of whether they went on to participate in the intervention.

An overview of the study outcomes, the NGRT scales used to measure the outcomes and various scores used in the analysis is provided in Table 2.

Table 2: An overview of primary and secondary outcomes and outcome indicators used in the analysis

Outcome	Measure	Scale	Score used in unadjusted descriptive analysis	Score used in unadjusted regression analysis	Score used in adjusted regression analysis
Reading ability (primary)	NGRT	The Overall Reading Scale	Scale score*	Age-standardised score**	Age-standardised score**
Sentence completion (secondary)	NGRT	The Sentence Completion Scale	Scale score*	Scale score*	Scale score*
Passage comprehension (secondary)	NGRT	The Passage Comprehension Scale	Scale score*	Scale score*	Scale score*
* The scale score is the equivalent of a raw score for an adaptive test **** Age-standardised score has the mean standardised score of around 100 with a standard deviation of around 15.**					

As indicated above in Table 2, descriptive analyses (for both the primary and secondary outcomes) are based on scale scores for pupils for whom post-test data were available.

Hierarchical linear regression modelling was used to compute the effect sizes on both the primary and secondary outcomes. In models designed to test the primary outcome the *age standardised* scores were used which provide an indication of whether pupils are performing at the average level for their age, below or higher than average.[8] The scale scores (as the only available scores) were used for the secondary outcomes (i.e. sentence completion and passage comprehension). The modelling strategy used involved fitting a multi-level linear regression model with random intercepts. Pupils represented Level 1 and classes were indicative of Level 2. Fixed effects at Level 2 to account for the potential variation at school level were also used. Schools were modelled as fixed effects because the sample of schools was chosen pragmatically and was not selected on the basis of random sampling. The following covariates were included in the adjusted analysis:

- Class level: A dummy variable indicating whether the class was an intervention class and a set of dummy variables indicating schools.
- Pupil level: Pre-test scores, eligibility for FSM, EAL, ethnic group, gender and month of birth. These covariates that are possibly predictive of the outcome under study were included in the analysis to adjust for possible imbalances between intervention and control pupils.

---

[8] Age standardised scores are based on the pupils' raw score which has been adjusted for age and placed on a scale that makes a comparison with a nationally representative sample of pupils of the same age across the UK. The average score is 100.

Separate pre-specified regression models were estimated for Year 7 and Year 9 pupils and for those who qualified for FSM. To confirm the findings with regards to FSM an interaction test was also implemented.[9]

Following a similar approach, separate pre-specified regression models were also implemented for the two genders, boys versus girls (separately for Year 7 and Year 9 pupils). The analyses followed a similar approach to the one described above.

Finally, even though specified in the protocol, we unable to carry out separate analyses for pupils with EAL due to very low sample sizes.[10] Furthermore, the current report does not include SEN subgroup analysis on outcomes of interest. However, the project team will be carrying out further analysis to assess the impact of the programme on SEN pupils.

Effect sizes and their respective 95 per cent confidence intervals were calculated following the procedure set out in Tymms (2004):

$$ES = \frac{\beta}{\sigma}$$

Where β represents the adjusted difference in outcomes between intervention and control groups obtained from the adjusted multi-level model and σ the square root of the pupil level variance obtained from fitting an unadjusted multi-level model. The unadjusted model contains a constant, an intervention dummy variable and class level fixed effect dummies, but no further covariates. The Standard Error of the Effect Size was computed according to Shagen and Elliot (2004).

Finally, summary descriptive statistics are produced along with estimates of ICCs for each regression model estimated.

Process evaluation methodology

The purpose of the process evaluation was to explore and understand the experiences of schools and teachers in implementing the Paired Reading programme in their school. The aims were to:

- Understand, at a strategic level, what attracted schools to take part in the programme.
- Understand the process of joining and setting up the programme in the school including training, pairing the pupils and timetabling.
- Explore how the programme worked in practice and any implementation, fidelity or sustainability issues.
- Find out in which ways teachers and senior leaders perceived the programme to be making an impact at pupil and school level.
- Explore how the teachers felt the programme could be developed if it were to be implemented more widely.
- Look into issues around sustainability of the programme

---

[9] The interaction models were implemented using a random intercept and random slope hierarchical regression model (multi-level model). We specified a random slope for the individual level variable indicating if the student was eligible for free school meals. This means that we allowed the effect (slope) of free school meal eligibility on the post-test to vary across classes. We further specified that any difference in the variance of these slopes (i.e. in the effect receiving free school meals has on the outcome) is due to having been in a class which was in the treatment group versus the control. This latter specification was included through a cross-level interaction between receiving free school meals and the allocation into treatment or control classes. The results of this interaction indicate if being in a treatment versus control class has any impact on the effect FSM has on the outcome.

[10] There are only 31 EAL pupils in Year 7 and 46 EAL pupils in Year 9 (as randomised). As analysed, the numbers drop to 28 in Year 7 and to 45 in Year 9.

A qualitative approach was used to meet these objectives and involved in-depth interviews with teachers and senior leaders from participating schools.

In addition to staff interviews, the original study design included an online teacher survey to complement the information provided by the project team on participating schools. The results of the survey were intended to be used to select a representative sample of experiences to follow up in depth. However, it was not feasible to carry out the online survey as schools were reluctant to introduce further research burden in addition to already existing data collection as part of the trial.

Process evaluation was set out to carry out interviews with two staff members from six schools including a mixture of both Year 7 and Year 9 teachers and senior leaders or operations managers. As indicated below, there were 8 interviews in total achieved as part of the process evaluation across 8 schools including 3 senior leaders, 3 Year 7 teachers and 2 Year 9 teachers.

Sampling and recruitment

The table below provides a breakdown of the achieved sample in relation to the primary sampling criteria. The aim was to speak to senior leaders as well as teachers. The table below provides a breakdown of the achieved sample by the type of interviewee. All ten secondary schools that participated in the programme were approached to take part in the process evaluation as they had previously signed a MoU with North Tyneside Council, in which they agreed to participate in research.

Table 3: Achieved sample for the process evaluation, based on number of pupils, FSM eligibility and Special Education Needs

Interview focus	Interviewee	Number of pupils in Year 7	Number of pupils in Year 9	FSM	School Action	School Action Plus	Statement
Strategic	Senior leader	117	107	27%	10%	9%	2%
	Senior leader	225	184	15%	17%	9%	2%
	Senior leader	260	222	9%	8%	7%	2%
Year 7 pupils	Teacher	166	200	20%	12%		10%
	Teacher	105	113	17%	18%	10%	2%
	Teacher	115	158	29%	11%	4%	1%
Year 9 pupils	Teacher	132	132	16%	6%	8%	2%
	Teacher	168	149	19%	12%	6%	3%

Each school was contacted and provided with details of the process evaluation. They were invited to nominate one member of teaching staff taking part in the programme to be interviewed. An online booking system was set up to allow participating staff to arrange a suitable interview time.

Individual depth interviews were conducted with five teachers and three senior leaders in eight of the schools. Interviews took place over the telephone and lasted approximately half an hour each.

The eight achieved interviews covered schools with a wide range of characteristics. As the above table shows, schools included in the process evaluation varied in terms of their size, number of pupils eligible for FSM and with special educational needs (School Action, School Action Plus, Statement). This variation in key demographics helped give a good understanding of how the programme worked in a range of different school settings.

The interviewees were also varied. Among the senior leaders, interviews were conducted with 2 deputy head teachers as well as a teaching and learning co-ordinator. All of the teaching staff interviewed were responsible for the day to day running of the programme,

Ideally, interviews would have been conducted with all of the schools involved in the programme. However, this was not possible due to the unavailability of teachers; many were too busy to take part in an interview towards the end of the summer term despite flexibility in times and dates offered to them.

Carrying out the interviews

The interviews all took place by phone. They were based around a topic guide to ensure systematic coverage of key issues, but were also intended to be flexible and interactive, allowing issues of relevance for individual respondents to be covered through detailed follow-up questioning. Follow-up questioning of interviewees was led by their responses, ensuring questioning was responsive and all issues were explored. Copies of the topic guides have been attached to this report in the appendices (see Appendix B).

Two topic guides were developed: one for use with senior leaders, the other for interviews with teaching staff.

The interviews with senior leaders were more focused on the school-level implementation of the programme:

- Joining the peer tutoring programme - including any strategic considerations.
- Running the programme - including cost and resources required to implement the programme
- Response to the programme from the staff.
- Implications of the intervention - including any perceived impacts and whether it would be sustained or rolled out further in the future.

The teacher topic guide was focused more on the day to day delivery of the programme and covered:

- Other literacy interventions in the school, including those running before or at the same time as the peer tutoring programme.
- Training provided to teachers, its coverage and suitability.
- Informing and training the pupils.
- The process of pairing the pupils and setting up the programme.
- The Paired Reading sessions - including the format, any barriers to successful implementation, progress monitoring and fidelity.
- Perceived impact - on the pupils as well as on the wider school.
- Proposed development, any necessary changes or recommendations.

Analysis

All interviews were digitally recorded and stored securely. An analysis framework was set up and the research team populated cells in the framework for each respondent. Following this charting, the framework was analysed thematically.

Costs

Please see page 43 for further details on the costs.

Impact evaluation

Timeline

Table 4: Project timeline

Date	Activity
March 2013	Recruitment of schools into the study
September 2013	RCT begins with 10 schools in North East of England
September 2013	Reading test 1 with Year 7 and Year 9 classes at participating schools
September 2013	Randomisation of classes into intervention and control groups at participating schools
September 2013	Pairing of Year 7 and Year 9 intervention classes at participating schools
November 2013	First teacher training session held
January 2014	Programme delivery begins with Year 7 and Year 9 intervention classes at participating schools
January 2014	Remaining classes at participating schools continue operating as usual
January 2014	Second teacher training session held
May 2014	Staff interviews with senior leaders and teachers at participating schools
June 2014	Programme delivery ends at participating schools
June 2014	Reading test 2 with all Year 7 and Year 9 classes
September 2014	Reading test scores received by NatCen

Participants

School recruitment

Schools were recruited directly by the delivery team and were contacted by email followed by in-person meetings with the head teacher or senior leader. Given that the randomisation was done at class level, to recruit 120 classes the delivery team approached 15 secondary schools with the proportion of pupils eligible for free school meals (FSM) above the national average which were invited to participate in the trial. Of the total number of schools, ten were from within the North Tyneside local authority while five schools were based in the neighbouring authorities: three schools in Newcastle, one school in Gateshead, one school in South Tyneside and one school in Sunderland. All schools that were

approached met the inclusion criterion (FSM eligibility in the pupil population), however, four schools declined to participate. One school which had initially agreed to be a part of the research withdrew before the pre-test (and before randomisation) due to an imminent Ofsted inspection.

Schools within North Tyneside were invited to participate first according to the EEF grant, followed by school in the neighbouring authorities with above average proportion of pupils eligible for FSM.[11]

Class recruitment

Within each school that was included in the trial, all Year 7 and Year 9 classes were eligible to participate in the study and thus for randomisation. However, when there were an uneven number of classes between year groups, the extra classes (either Year 7 or Year 9) were eliminated. (See Randomisation section for more details).

All schools sent opt-out letters to parents of all children in Year 7 and Year 9 (see Appendix A). The letter explained to parents that their child, as part of their class, will be randomly selected to either receive the intervention programme and assessments, or assessments only. Parents were asked to inform their child's teacher if they did not wish their child participating in the study. No parent opted that his/her child did not take part in the study.

Participant flow

Figure 1 shows the flow of participants in relevant arms through the study. There were a total of 2,736 pupils in the study at the point of randomisation (1,370 and 1,366 in Year 7 and Year 9 respectively).

At the onset of the trial, 58 classes in 10 schools (containing 1,313 participating pupils) were randomly allocated to receive the intervention and 62 classes in 10 schools (containing 1,423 participating pupils) to the control condition. There were 652 Year 7 pupils and 661 Year 9 pupils in the intervention condition and 718 Year 7 pupils and 705 Year 9 pupils in the control condition.

After randomisation and pre-testing, no classes dropped out of the study. Thus, there was no attrition at the level of randomisation. However, a number of pupils did not complete the post-test. The number of pupils in the intervention arm that were lost to follow-up was 66 pupils (25 in Year 7 and 41 in Year 9). The loss to follow-up in the control arm was 85 pupils (36 in Year 7 and 49 in Year 9). Overall, the proportion of pupils allocated to intervention and to the control group lost to follow-up was approximately 5 per cent and 6 per cent respectively.

In addition, we eliminated from the analysis those who did not have a valid Overall Reading score. This overall score is computed based on the results of the Sentence Completion and the Passage Comprehension scales (the secondary outcomes in this study). In the first stage of the assessment all pupils fill in the Sentence Completion Scale. In the second stage (based on their ability) pupils are either assigned to do the Passage Comprehension Scale or the Phonics Scale. However, the Phonics Scale does not generate scalable results and its results cannot be included in the computation of the scale score for the Overall Reading Scale. This is why the pupils who completed the Phonics Scale do not have valid results on The Overall Reading Scale and need to be eliminated. We identified 8 such students in the intervention groups (2 in Year 7 and 6 in Year 9) and 2 in the control groups (1 in Year 7 and 1 in Year 9).

The final number of pupils who had valid post-test data and were thus identified as eligible for unadjusted primary analysis was 1,239 (94 per cent of those allocated) in the intervention arm (625 who

---

[11] The 8 schools in North Tyneside also meet the above average FSM criterion.

started in Year 7 and 614 in Year 9) and 1,336 (94 per cent of those allocated) in the control arm (681 who started in Year 7 and 655 in Year 9).

However, it is important to note that further participants were excluded from adjusted primary and secondary analysis due to missing pupil level covariates capturing relevant socio-demographics (e.g. whether the pupil qualified for free school meals, EAL, ethnic group, sex and month of birth) in the data received from the National Pupil Database (NPD). Furthermore, pupils who did not complete the Passage Comprehension pre-test were also excluded. In the intervention arm, 2 Year 7 pupils and 3 Year 9 pupils were excluded from adjusted analysis due to not having socio-demographic or valid pre-test data. As for the control arm, 4 Year 7 pupils and 1 Year 9 pupil were excluded from adjusted analysis due to not having socio-demographic data or valid pre-test data.

Figure 1: CONSORT diagram
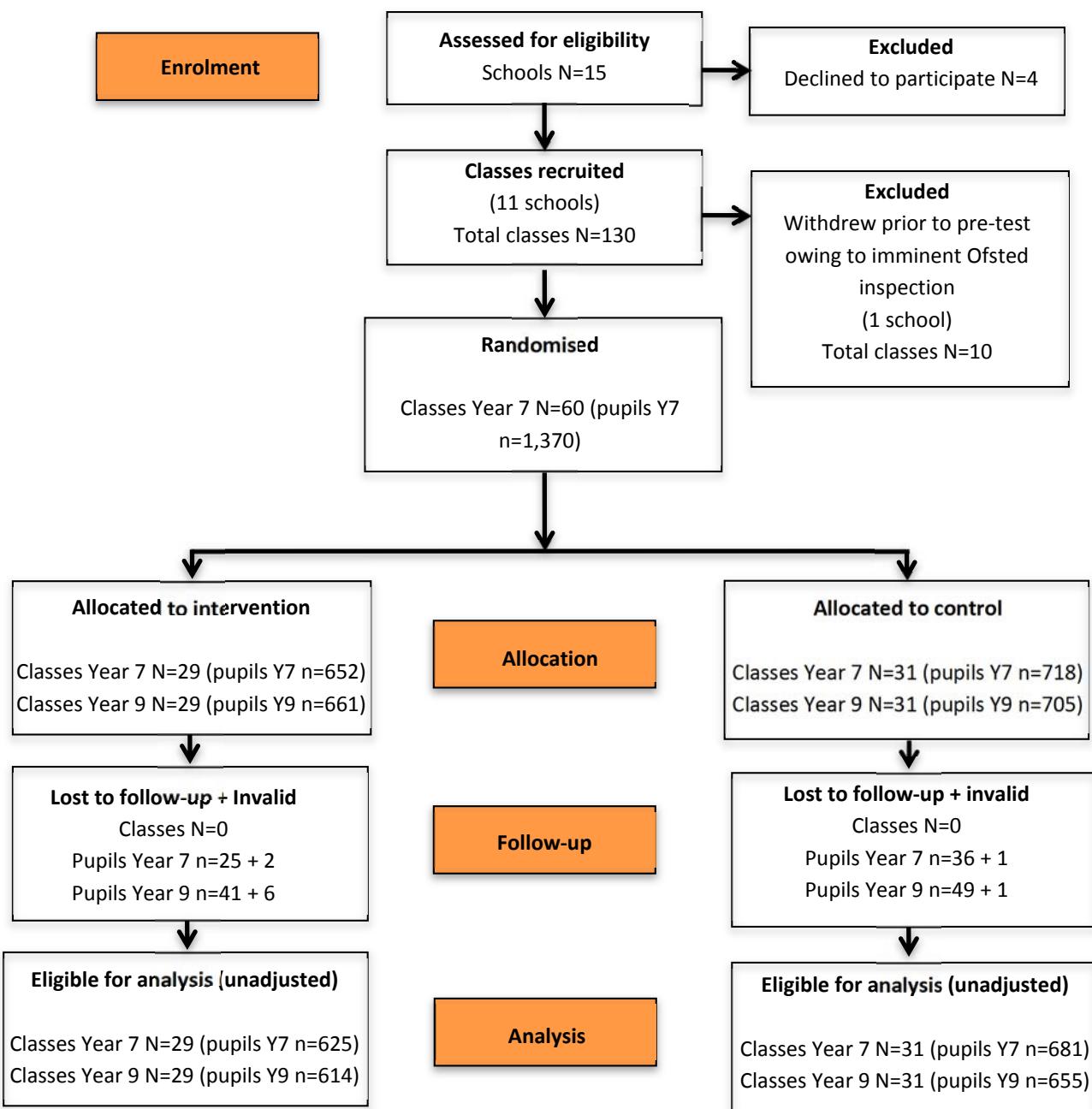
Note: A further 5 pupils in the intervention arm and 5 in the control arm were excluded from adjusted analyses due to missing pupil level covariates capturing relevant socio-demographics.

School characteristics

In this section we look at the characteristics of the ten schools that were included in the trial. As illustrated in Table 5, most of the schools had Foundation status, with two schools being Academies and another two having Community School status. In terms of size, the schools included in the trial varied from smaller (maximum capacity of 504) to large schools (maximum capacity of 1,650). As indicated earlier, all schools were above the national average in terms of the proportion of pupils eligible for FSM. Finally, the number of pupils in each participating school who were included in the intervention group versus the control group is shown in Table 6.

Table 5: School level characteristics

	Type of School	Highest age	Capacity	Number of pupils	Number of boys	% FSM[12]
School 1	Community	18	1526	1360	681	20.3
School 2	Foundation	18	1022	780	420	23
School 3	Academy	16	900	No info	No info	No info
School 4	Foundation	18	1284	1015	512	14.7
School 5	Foundation	16	885	785	410	17.7
School 6	Foundation	18	1010	995	512	15.4
School 7	Foundation	16	504	460	239	35.3
School 8	Foundation	18	951	575	294	18.5
School 9	Community	16	1100	690	384	38.1
School 10	Academy	18	1650	1665	810	10.4

Table 6: Year 7 and Year 9 pupils in participating schools

	Year 7 pupils				Year 9 pupils			
	Year 7 total	FSM eligible	Year 7 Control	Year 7 Treatment	Year 9 total	FSM eligible	Year 9 Control	Year 9 Treatment
School 1	166	33	83	83	200	47	103	97
School 2	117	26	60	57	107	28	56	51
School 3	132	17	70	62	110	25	55	55
School 4	198	27	109	89	184	29	92	92
School 5	144	33	74	70	149	24	74	75
School 6	104	23	59	45	113	17	68	45
School 7	64	23	29	35	75	32	37	38
School 8	70	13	47	23	77	14	50	27
School 9	115	45	43	72	129	43	53	76
School 10	260	30	144	116	222	25	117	105

Class characteristics

As discussed earlier, the randomisation into the intervention and control groups was carried out at class level. Table 7 illustrates the relevant differences between classes in terms of their composition. The average number of pupils per class (at randomisation stage) in the intervention arm is slightly lower for both years compared to the control arm. In Year 7, there are, on average, 22.48 pupils per class in the intervention group and there are 23.16 pupils in a class placed in the control group. In Year 9, the difference between the sizes of classes in the two arms is smaller: an intervention class has an average number of 22.79 pupils while a control class is slightly larger at 22.74 pupils.

---

[12] The pupil is either eligible for free school meals or has been eligible for free school meals at some point in the last 6 years or has been looked after continuously for 6 months during the year + aged 5-15 (2013/2014).

Table 7: Class level characteristics

	Year 7		Year 9	
All pupils	**Intervention class (n=29)**	**Control class (n=31)**	**Intervention class (n=29)**	**Control class (n=31)**
Number of pupils in trial	652	718	661	705
Mean (SD) pupils per class	22.48 (4.80)	23.16 (5.01)	22.79 (4.60)	22.74 (4.44)
Median (Min; Max) pupils per class	23 (8; 30)	24 (13; 30)	24 (10; 29)	24 (10; 30)
FSM pupils only	**Intervention class (n=27)[13]**	**Control class (n=31)**	**Intervention class (n=29)**	**Control class (n=31)**
Number of pupils in trial	127	143	144	140
Mean (SD) FSM pupils per class	4.70 (2.57)	4.61 (2.45)	4.97 (2.77)	4.52 (2.51)
Median (Min, Max) FSM pupils per class	4 (1; 11)	4 (1; 10)	5 (1; 12)	4 (1; 12)

Pupil characteristics

The study involved the participation of both Year 7 and 9 pupils. Table 8 presents a summary of Year 7 pupils' characteristics in intervention and control groups as randomised and as analysed (i.e. those eligible for unadjusted analysis based only on the post-test scores for the primary -and secondary[14]- outcome). There were 652 pupils in intervention classes and 718 pupils in control classes at randomisation. Restricting analysis to just the pupils with post-test scores (as analysed) reduces the number. As such, there are 625 pupils in the intervention classes and 681 in the control classes.

The baseline characteristics of pupils in intervention and control classes on the primary outcome were similar at randomisation and again at analysis, suggesting that the loss of pupils between randomisation and post-test analysis did not introduce bias on observable variables into the sample. There are slight differences in the distributions of scores at Key Stage 1 (at both randomisation and primary analysis) and there is no data for Year 7 on most Key Stage 2 English and Writing scores.

---

[13] Two intervention classes in Year 7 are lost when looking only at FSM pupils as in these classes there are no pupils who qualify for FSM.
[14] Given that the secondary outcomes are the components of the primary outcomes all students who have a score on the post-test of the primary outcome will automatically also have a score for the secondary outcomes.

Table 8: Pupil characteristics Year 7

	All pupils as randomised			All pupils as analysed (with post-test scores)		
	Intervention	**Control**	**Difference[15]/ (Effect size)**	**Intervention**	**Control**	**Difference / (Effect size)**
Baseline score (reading ability)[16]	99.31	99.56	-0.25 / (0.017)	99.26	99.82	-0.56 / (0.039)
Baseline score (passage comprehension)	324.20	324.38	-0.18 / (0.003)	324.00	325.60	-1.60 / (0.027)
Baseline score (sentence completion)	328.90	331.40	-2.50 / (0.054)	328.74	331.91	-3.17 / (0.069)
KS1 Reading – L1	15%	13%	2%	14%	13%	1%
KS1 Reading – L2a	24%	26%	-2%	24%	25%	-1%
KS1 Reading – L2b	22%	24%	-2%	22%	24%	-2%
KS1 Reading – L2c	15%	16%	-1%	16%	15%	1%
KS1 Reading – L3	21%	19%	2%	20%	19%	1%
KS1 Reading – Missing	3%	3%	0%	3%	3%	0%
KS2 Reading – Expected Level	88%	91%	-3%	88%	91%	-3%
KS2 Reading – above expected	43%	45%	-2%	42%	45%	-3%
Female	51%	50%	1%	51%	51%	0%
EAL	3%	2%	1%	2%	2%	0%
Native English speaker	97%	98%	-1%	97%	98%	-1%
FSM	20%	20%	0%	19%	19%	0%
Ever FSM	41%	40%	1%	40%	40%	0%
SEN	20%	21%	-1%	20%	20%	0%
White	95%	96%	-1%	96%	96%	0%
Asian	1%	1%	0%	1%	1%	0%
Black	1%	0%	1%	1%	0%	1%
Chinese	0%	0%	0%	0%	0%	0%
Mixed	1%	0%	1%	0%	1%	-1%
Other	2%	1%	1%	2%	1%	1%
Number of pupils (all)	652	718	-	625	681	-

---

[15] The difference is computed by subtracting the score in the control group from the score in the intervention groups.
[16] We eliminate those who do not have a valid Overall Reading Score due to not having taken the Passage Comprehension test. There are 2 such pupils in Year 7

Note. L = Level. There were no data available on the following characteristics: KS2 English – L2, KS2, English – L3, KS2 English – L4, KS2 English – L5, KS2 English – Missing, KS2 English – Low attainment, KS2 English – Expected level, KS2 English – High attainment, KS2 Writing – Expected level.

The results of a similar exploration of Year 9 pupils' characteristics are provided in Table 9. The intervention schools had 661 pupils and control schools had 705 pupils at randomisation, with these numbers falling to 614 and 655 respectively in the sample for unadjusted analysis. Slightly larger differences in average Key Stage 1 scores are observed for Year 9 pupils compared to Year 7, both at randomisation and at analysis. In addition, in Year 9 classes there seem to be slightly more FSM pupils in the intervention classes compared to the control classes both at randomisation and at analysis. The adjusted multivariate analysis does control for FSM.

Table 9: Pupil characteristics Year 9

	All pupils as randomised			All pupils as analysed (with post-test scores)		
	Intervention	**Control**	**Difference / (Effect size)**	**Intervention**	**Control**	**Difference / (Effect size)**
Baseline score (reading ability)[17]	98.76	99.15	-0.52 / (0.035)	99.00	100.00	-1.14 / (0.078)
Baseline score (passage comprehension)	348.63	349.06	-0.43 / (0.007)	349.78	352.25	-2.47 / (0.043)
Baseline score (sentence completion)	353.59	356.45	-2.86 / (0.058)	354.12	358.71	-4.59 / (0.093)
KS1 Reading – L1	13%	12%	1%	13%	12%	1%
KS1 Reading – L2a	24%	24%	0%	25%	25%	0%
KS1 Reading – L2b	23%	20%	3%	24%	19%	5%
KS1 Reading – L2c	14%	14%	0%	14%	14%	0%
KS1 Reading – L3	20%	24%	-4%	20%	25%	-5%
KS1 Reading – Missing	6%	6%	0%	5%	5%	0%
KS2 English – L2	1%	1%	0%	1%	1%	0%
KS2 English – L3	13%	12%	1%	12%	12%	0%
KS2 English – L4	57%	52%	5%	58%	52%	6%
KS2 English – L5	26%	30%	-4%	26%	31%	-5%
KS2 English – Missing	79%	89%	-10%	72%	93%	-21%
KS2 English – Low Attainment	2%	4%	-2%	2%	3%	-1%
KS2 English – Expected level	84%	83%	-1%	85%	85%	0%
KS2 English – High attainment	26%	30%	-4%	27%	32%	-5%
KS2 Reading – Expected Level	86%	85%	-1%	86%	86%	0%
KS2 Reading – Above expected	42%	47%	-5%	43%	49%	-6%
KS2 Writing – Expected level	77%	77%	0%	78%	79%	-1%
Female	48%	48%	0%	47%	48%	-1%
EAL	4%	3%	1%	4%	3%	1%
Native English speaker	96%	96%	0%	96%	96%	0%

---

[17] We eliminate those who do not have a valid Overall Reading Score due to not having taken the Passage Comprehension test. There are 3 such pupils in Year 9.

FSM	22%	20%	2%	21%	18%	3%
Ever FSM	41%	38%	3%	40%	36%	4%
SEN	22%	22%	0%	21%	20%	1%
White	94%	95%	-1%	94%	95%	-1%
Asian	1%	1%	0%	1%	2%	-1%
Black	1%	1%	0%	1%	1%	0%
Chinese	0%	0%	0%	0%	0%	0%
Mixed	1%	1%	0%	1%	1%	0%
Other	2%	2%	0%	2%	1%	1%
Number of pupils (all)	661	705	-	614	655	-

Outcomes and analysis

Primary outcome: The overall reading score

The analysis and results presented below first focus on the primary outcome. The two secondary outcomes will be discussed in subsequent sections. All outcomes analysed were pre-specified in the protocol. All sub-group analyses were pre-specified in the protocol as a separate FSM analysis is a requirement of all EEF evaluations.

Summary statistics of the primary outcome

Initial analysis is based only on the post-scores for the primary outcome (the NGRT Overall Reading Scale[18]), without taking into account pre-test scores or pupil characteristics and without assessing whether the differences we observe in the scores are statistically significant[19]. Such tests are performed in the subsequent sections. As the summary statistics displayed in Table 10 indicate, the mean post-test score on the primary outcome for Year 7 pupils in the intervention group (338.27) is very close (although lower) to that in the control group (341.69). As analysed, the sample of Year 7 students consists of just over 1,300 pupils across 60 classes. Post-test scores were missing for around 4 per cent of analysed pupils in intervention classes and 5 per cent of pupils in control classes.

Table 10: Unadjusted average overall reading scores Year 7 pupils (primary outcome)

	Intervention	Control	Total
Sample as analysed classes	n= 29	n= 31	n= 60
Sample as analysed pupils	n= 625	n= 681	n= 1306
Primary outcome	**Overall Reading Scale**		
Mean (SD)	338.27 (53.43)	341.69 (52.26)	340.06 (52.83)
Median (Min, Max)	340 (144; 468)	340 (164; 496)	340 (144; 469)
Missing (% of pupils randomised)	4.14%	5.15%	4.67%

The 'as analysed' summary statistics describing the mean and the median levels of the post-test for the two groups for Year 9 are displayed in Table 11. Similarly to Year 7 pupils, the difference in

---

[18] We excluded from the 'as analyzed' group those pupils who had a valid score on the Sentence Completion Scale but did not complete the Passage Comprehension Scale. There was 1 such pupil in the Year 7 control group and 2 pupils in the Year 7 intervention group. In Year 9 we excluded 1 pupil in the control group and 6 pupils in the intervention group.

[19] For this specific analysis the scale score was used as the age-standardised score would have a fixed mean and standard deviation and therefore the results would not be informative.

average post-test scores between intervention (360.33) and control (365.87) pupils is small. Approximately 7 per cent of post-test scores are missing for pupils, as analysed, in both the intervention schools and control schools.

Table 11: Unadjusted average overall reading scores Year 9 pupils (primary outcome)

	Intervention	Control	Total
Sample as analysed classes	n= 29	n= 31	n= 60
Sample as analysed pupils	n= 614	n= 655	n= 1269
Primary outcome	Overall Reading Scale		
Mean (SD)	360.33 (57.03)	365.87 (58.17)	363.19 (57.66)
Median (Min, Max)	368 (144; 512)	372 (168; 512)	372 (144; 512)
Missing (% of pupils randomised)	7.11%	7.09%	7.10%

Impact of Paired Reading on the overall reading score

The analysis of the primary outcome was undertaken using a multi-level regression model with random intercepts, separately for Year 7 and 9 pupils. For each year group two analyses are presented, one based on an unadjusted model and the second one based on adjusted analysis. The adjusted analysis includes the baseline measure (pre-test) and EEF designated variables as covariates. These covariates are students' month of birth, sex, eligibility for FSM and second level dummy variables representing schools.

Table 12 reports estimates for Year 7 pupils. The difference between the effect of the allocation on the post-test score in both the adjusted and unadjusted analyses was very small and did not reach conventional levels of statistical significance. This indicates that there is no evidence of any impact on the reading competency of the tutees (Year 7 pupils) who received tutoring on reading from Year 9 pupils as part of the Paired Reading programme The adjusted analysis was conducted on a final sample of 1,300 Year 7 pupils across 60 classes in all 10 schools. The adjusted analysis reveals a difference of -0.28 (95% CI: -1.58 to 1.02) in the age standardised overall reading score at post-test between the intervention and the control group. The negative score indicates that the reading score was lower in the intervention group. This is equivalent to an effect size of -0.02 (95% CI: -0.15 to 0.11).

Table 12: Analysis of the overall reading score - Year 7 pupils (primary outcome)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.02 (-0.15 to 0.11)
Regression coefficient (95% CI)	-0.75 (-4.35 to 2.85)	-0.28 (-1.58 to 1.02)
P-value	0.683	0.672
ICC (SE)	0.176 (0.034)	0.037 (0.015)
Variance class level (SE)	40.74 (9.44)	2.88 (1.18)

Variance pupil level (SE)	190.50 (7.65)	75.32 (3.02)
Total sample size (classes)	1,305 (60)	1,300 (60)

Similar analysis was performed for Year 9 pupils. Table 13 displays the results. Similarly to Year 7, the difference between the allocated groups on the post-test score in both the unadjusted and adjusted analyses was very small and did not reach conventional levels of statistical significance. Therefore, there was no evidence of impact on reading competence for Year 9 pupils. This means that assisting younger pupils in improving their reading does not make a visible difference in their level of reading competence. The adjusted analysis was conducted on a final sample of 1,265 Year 9 pupils across 60 classes in all 10 schools. The adjusted analysis reveals a difference of -0.911.05 (95% CI: -2.08 to 0.25) in the age standardised overall reading score at post-test between the intervention and the control group. The negative score indicates that the reading score was lower in the intervention group. This is equivalent to an effect size of -0.06 (95% CI: -0.14 to 0.02).

Table 13: Analysis of the overall reading score - Year 9 pupils (primary outcome)

	Unadjusted analysis	**Adjusted analysis**
Effect Size (CI)	-	-0.06 (-0.14 to 0.02)
Regression coefficient (95% CI)	-1.60 (-4.16 to 1.00)	-0.91 (-2.08 to 0.25)
P-value	0.231	0.125
ICC (SE)	0.059 (0.018)	0.018 (0.012)
Variance class level (SE)	14.44 (4.71)	1.41 (0.93)
Variance pupil level (SE)	230.81 (9.40)	78.58 (3.20)
Total sample size (pupils)	1,266 (60)	1,265 (60)

Impact of Paired Reading on the overall reading score: FSM subgroups

In ensuring that the analysis conforms to EEF standards, a set of additional analyses were carried out. These subgroup analyses were performed to examine whether there was evidence that the Paired Reading programme had an impact on the overall reading score (primary outcome) depending on pupil eligibility for FSM. This analysis was carried out for both Year 7 and 9 pupils separately and was implemented using two distinct methodological approaches. In the first approach, both the unadjusted and adjusted multilevel regression models were re-estimated including only pupils eligible for FSM. In the second approach interaction tests were implemented to assess whether the eligibility for FSM has an impact on the effect the intervention (versus the control) has on the primary outcome.

Amongst the 246 Year 7 pupils eligible for FSM there is no evidence of impact on the primary outcome by the allocation groups; results reveal a difference between allocated groups equivalent to an effect size of -0.04 but this did not reach conventional levels of statistical significance (95% CI:-0.21 to 0.13).

The equivalent result for the Year 9 sample, which comprised 246 pupils eligible for FSM, was an effect size of -0.06and again did not reach conventional levels of statistical significance (95% CI: -0.14 to 0.02). The negative results indicate that the reading score was lower in the intervention group. Full results from these analyses are presented in Appendix D (see Tables D1a and D4a for Year 7 and Year 9 sample respectively).

The interaction analysis described as our second approach is also presented in Appendix D (Tables D1b and D4b). For both Year 7 and Year 9 the results displayed for the interaction models is fully consistent with the analysis discussed above and shows that the effect of allocation does not vary according to whether pupils are FSM eligible or not.

Impact of Paired Reading on the overall reading score: Gender subgroups

Along with performing subgroup analyses for pupils eligible for FSM we also carried out analysis by gender. To accomplish this we re-ran the analysis discussed above for each gender separately. These analyses were carried out to determine whether there was evidence that the Paired Reading programme had a different impact on the overall reading ability of boys versus girls.

For Year 7 boys (635 pupils) there is no evidence the Paired Reading programme had an impact. The results reveal a difference between groups equivalent to an effect size of -0.05 but this did not reach conventional levels of statistical significance (95% CI:-0.17 to 0.07). The equivalent result for the Year 7 girls (665 pupils) was similar and also showed no impact. The effect size was equal to 0.02 and was not statistically significant (95% CI: -0.08 to 0.12). The negative result indicates that the reading score was lower in the intervention group while the positive score shows that reading score was higher in the intervention group. Full results from these analyses are presented in Appendix E (see Tables E1a and E1b for Year 7 boys and girls respectively). These results were also confirmed by the second analytical strategy: the interaction analysis. The results presented in Table E1c (Appendix E) indicate that gender does not have an impact on the effect the allocation (i.e. treatment) has on reading ability.

Looking at Year 9 boys (662 pupils), the results are consistent with those for Year 7 boys: the treatment does not seem to have had an impact on their overall reading ability. The effect size is -0.01 and is not statistically significant (95% CI: -0.11 to 0.09). The same is not the case for Year 9 girls (603 pupils). The results show that for girls in Year 9 the treatment did have an effect on their overall reading ability. The effect is, however, negative, meaning that having acted as a tutor in the Paired Reading programme actually decreased their overall reading ability. The effect size is equivalent to -0.12 and is statistically significant (95% CI: -0.22 to -0.02). The second analytical strategy comes close to confirming that, it is likely that the treatment had a negative effect on girls. The regression coefficient corresponding to the interaction term fails to attain statistical significance at the 0.05 level but it is significant at the 0.1 level (the actual value is 0.075).

Secondary outcomes

The two secondary outcomes of interest are the two components of the Overall Reading Scale. These are the Passage Comprehension Scale and the Sentence Completion Scale. The paragraphs below present the results for both these outcomes. We first assess passage comprehension and subsequently consider sentence completion.

Summary statistics of secondary outcome: Passage comprehension

Similarly to the initial analysis of primary outcome, we present summary statistics for all students with values on the post-test, without taking into account pre-test scores or pupil characteristics. The summary statistics displayed in Table 14 indicate that the mean post-test score on the passage comprehension for Year 7 pupils in the intervention group (337.15) is similar to that in the control group (339.15). As mentioned before, the sample of Year 7 students consists of just over 1,300 pupils across

60 classes. Post-test scores were missing for around 4 per cent of analysed pupils in intervention classes and 5 per cent of pupils in control classes.

Table 14: Unadjusted average passage comprehension scores Year 7 pupils (secondary outcome)

	Intervention	Control	Total
Sample as analysed classes	n= 29	n= 31	n= 60
Sample as analysed pupils	n= 625	n= 681	n= 1306
Secondary outcome	Passage Comprehension Scale		
Mean (SD)	337.15 (60.62)	339.15 (60.68)	338.19 (60.64)
Median (Min, Max)	340 (148; 472)	344 (148; 520)	340 (148; 520)
Missing (% of pupils randomised)	4.14%	5.15%	4.67%

The results presented above are similar for Year 9 students (Table 15). In essence there does not seem to be any differences between the intervention group and the control group when it comes to the mean value of the post-test. As the figures show, the difference in average post-test scores between intervention (354.72) and control (361.94) pupils is very small.

Table 15: Unadjusted average scores Year 9 pupils – passage comprehension (secondary outcome)

	Intervention	Control	Total
Sample as analysed classes	n= 29	n= 31	n= 60
Sample as analysed pupils	n= 614	n= 655	n= 1269
Secondary outcome	Passage Comprehension Scale		
Mean (SD)	354.72 (65.35)	361.94 (67.53)	358.45 (66.55)
Median (Min, Max)	364 (92; 520)	372 (88; 520)	372 (88; 520)
Missing (% of pupils randomised)	7.11%	6.96%	7.03%

Impact of Paired Reading on the passage comprehension

The analysis of the secondary outcomes was undertaken in an identical way to that of the primary outcome: a multi-level regression model with random intercepts was used separately for Year 7 and 9 pupils. Similarly to the above, the unadjusted and adjusted models are presented.

The findings for Year 7 pupils are shown in Table 16. Similarly to the analysis of the primary outcome, the difference between the effect of the intervention on the post-test score in both the adjusted and unadjusted analyses was very small and did not reach conventional levels of statistical significance. This indicates that there is no evidence of any impact on the passage comprehension of Year 7 pupils. The adjusted analysis was conducted on a final sample of 1,300 Year 7 pupils across 60 classes in all 10 schools. The adjusted analysis reveals a difference of -0.48 (95% CI: -6.07 to 5.17) in the passage

comprehension score at post-test between the intervention and control groups. The negative score indicates that the reading score was lower in the intervention group. This is equivalent to an effect size of -0.01 (95% CI: -0.12 to 0.10).

The results are similar for Year 9 students (Table 17) and indicate no significant difference between the intervention group and the control. The adjusted analysis was conducted on a final sample of 1,265 Year 9 pupils across 60 classes in all 10 schools. The adjusted analysis reveals a difference of -5.22 (95% CI: -11.95 to 1.16) in the passage comprehension score at post-test between the intervention and control groups. The negative score indicates that the reading score was lower in the intervention group. This is equivalent to an effect size of -0.09 (95% CI: -0.20 to 0.02).

Table 16: Analysis of passage comprehension - Year 7 pupils (secondary outcome)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.01 (-0.12 to 0.10)
Regression coefficient (95% CI)	-1.60 (-15.22 to 12.02)	-0.48 (-6.07 to 5.17)
P-value	0.818	0.876
ICC (SE)	0.164 (0.033)	0.036 (0.015)
Variance class level (SE)	572.81 (135.12)	53.60 (22.38)
Variance pupil level (SE)	2929.99 (117.59)	1425.66 (57.27)
Total sample size (pupils)	1,306 (60)	1,300 (60)

Table 17: Analysis of passage comprehension - Year 9 pupils (secondary outcome)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.09 (-0.20 to 0.02)
Regression coefficient (95% CI)	-6.98 (-17.32 to 3.36)	-5.22 (-11.59 to 1.16)
P-value	0.186	0.109
ICC (SE)	0.058 (0.018)	0.035 (0.014)
Variance class level (SE)	230.94 (74.96)	66.92 (27.93)
Variance pupil level (SE)	3724.15 (151.40)	1825.36 (74.28)

Total sample size (pupils)	1,269 (60)	1,265 (60)

Impact of Paired Reading on the passage comprehension: FSM subgroups

Using the same approaches presented in the subgroup (FSM) analysis of the primary outcome, we assessed the impact of FSM eligibility on the effect the intervention has on the relevant secondary outcome.

Amongst the 246 Year 7 pupils eligible for FSM there was no evidence of impact on the primary outcome by the allocation groups; results reveal a difference between allocated groups equivalent to an effect size of -0.02 without it reaching conventional levels of statistical significance (95% CI:-0.20 to 0.16). The results for Year 9 pupils (246 pupils) indicate an effect size of -0.03 without reaching conventional levels of statistical significance (95% CI: -0.22 to 0.16). The negative scores indicate that the reading score was lower in the intervention group. As before, full results from these analyses are presented in Appendix D (see Tables D7a and D10a for Year 7 and Year 9 sample respectively).

Similar results are obtained by the implementation of the second approach: the interaction analysis. (Tables D7b and D10b in Appendix D). For both Year 7 and Year 9 the results are consistent with the analysis discussed above and show that the effect of allocation does not vary according to whether pupils are eligible for FSM.

Impact of Paired Reading on the passage comprehension: Gender subgroups

The same approach (to the one used in the analysis of the primary outcome) was used to assess the separate impact the treatment might have had on boys' and girls' level of passage comprehension.

The analysis of Year 7 boys (635 pupils) showed no evidence of impact. The effect size is equivalent to -0.03 and is not statistically significant (95% CI: -0.16 to 0.10). The same can be said about Year 7 girls (665 pupils): available evidence does not allow us to conclude that the treatment had an effect. The size of the effect we discover is 0.02 without being statistically significant (95% CI: -0.10 . to 0.14). The positive score indicates that the reading score was higher in the intervention group. The full results are displayed in Appendix E (Tables E9a and E9b). The second analytical procedure (the interaction model) supports these findings. The results are included in Table E9c in the Appendix E.

For Year 9 pupils the results resonate with those identified in the analysis of the primary outcome. The results show that for Year 9 boys (662 pupils) the treatment had no effect on their level of passage comprehension. The effect size is equal to 0.01 and does not attain statistical significance (95% CI: -0.12 to 0.14). However, there is evidence that the treatment had a statistically significant effect on the level of Year 9 girls' (603 pupils) passage comprehension (effect size: -0.20; 95% CI: -0.21 to -0.19). The results indicate that for a Year 9 girl the inclusion in the Paired Reading programme as a tutor had a negative effect decreasing the level of passage comprehension. This finding is also confirmed by the interaction analysis displayed in Table E13c in Appendix E.

Summary statistics of secondary outcome: Sentence completion

Finally, the summary analysis of the second secondary outcome, sentence completion, produces results which are consistent with the findings obtained for the overall reading score (primary outcome) and the passage comprehension (secondary outcome). The summary statistics indicate that the mean post-test score on the sentence completion for Year 7 pupils (Table 18) in the intervention group (340.66) is similar to that in the control group (345.41). The sample size and the proportion of pupils missing on the post-test are similar to the summary statistics presented for the primary outcome.

Table 18: Unadjusted average sentence completion scores Year 7 pupils (secondary outcome)

	Intervention	Control	Total
Sample as analysed classes	n= 29	n= 31	n= 60
Sample as analysed pupils	n= 627	n= 682	n= 1309
Secondary outcome	Sentence Completion Scale		
Mean (SD)	340.66 (49.42)	345.41 (50.11)	343.14 (49.82)
Median (Min, Max)	344 (72; 464)	344 (36; 504)	344 (36; 504)
Missing (% of pupils randomised)	3.83%	5.00%	4.45%

Similar results were found for Year 9 pupils (Table 19). The summary statistics indicate no differences between the intervention group (367.32) and the control group (371.98) when it comes to the mean value of the post-test.

Table 19: Unadjusted average sentence completion scores Year 9 pupils (secondary outcome)

	Intervention	Control	Total
Sample as analysed classes	n= 29	n= 31	n= 60
Sample as analysed pupils	n= 620	n= 656	n= 1276
Secondary outcome	Sentence Completion Scale		
Mean (SD)	367.32 (59.35)	371.98 (55.10)	369.71 (57.23)
Median (Min, Max)	368 (64; 536)	368 (0; 536)	368 (0; 536)
Missing (% of pupils randomised)	6.20%	6.82%	6.52%

Impact of Paired reading on sentence completion

A similar multi-level analysis approach was implemented for the second secondary outcome (sentence completion) as well. Consistent with the previous analysis, once again we find no impact of the treatment on the post-test score of this secondary outcome. The adjusted analysis for Year 7 pupils (presented in Table 20) was conducted on a final sample of 1,304 Year 7 pupils across 60 classes in all 10 schools. The adjusted analysis reveals a difference of -1.60 (95% CI: -6.01 to 2.80) in the Sentence Completion score at post-test between the intervention and control groups. The negative score indicates that the reading score was lower in the intervention group. This is equivalent to an effect size of -0.04 (95% CI: -0.14 to 0.06).

Table 21 displays the results for Year 9 students, indicating again, the lack of effects detectable at the conventional levels of statistical significance. The adjusted analysis was conducted on a final sample of 1,275 Year 9 pupils across 60 classes in all 10 schools. The adjusted analysis reveals a difference of -0.43 (95% CI: -4.67 to 3.80) in the Sentence Completion score at post-test between the intervention and control groups. The negative score indicates that the reading score was lower in the intervention group. This is equivalent to an effect size of -0.01 (95% CI: -0.09 to 0.08).

Table 20: Analysis of sentence completion - Year 7 pupils (secondary outcome)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.04 (-0.14 to 0.06)
Regression coefficient (95% CI)	-4.29 (-15.02 to 6.45)	-1.60 (-6.01 to 2.80)
P-value	0.434	0.476
ICC (SE)	0.142 (0.032)	0.017 (0.013)
Variance class level (SE)	344.49 (86.76)	19.78 (15.21)
Variance pupil level (SE)	2068.95 (83.07)	1169.18 (47.12)
Total sample size (pupils)	1,309 (60)	1,304 (60)

Table 21: Analysis of sentence completion- Year 9 pupils (secondary outcome)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.01 (-0.09 to 0.08)
Regression coefficient (95% CI)	-4.06 (-11.68 to 3.57)	-0.43 (-4.67 to 3.80)
P-value	0.297	0.841
ICC (SE)	0.024 (0.013)	0.000 (0.000)
Variance class level (SE)	77.10 (42.10)	0.00 (0.00)
Variance pupil level (SE)	3079.07 (124.97)	1447.60 (57.33)
Total sample size (pupils)	1,276 (60)	1,275 (60)

Impact of Paired Reading on the sentence completion: FSM subgroups

The analysis of the FSM subgroup for the sentence completion (secondary outcome) was carried out in a similar fashion to the analysis described in the FSM analysis of the primary outcome. As such, amongst the 248 Year 7 pupils eligible for FSM there is no evidence of impact on the primary outcome by the allocation groups; results reveal a difference between allocated groups equivalent to an effect size of -0.15 without it reaching conventional levels of statistical significance (95% CI:-0.39 to 0.09). The result for Year 9 pupils (249 pupils), indicate an effect size of -0.13, which did not reach

conventional levels of statistical significance (95% CI: -0.31 to 0.05). The negative scores indicate that the reading score was lower in the intervention group. As before, full results from these analyses are presented in Appendix D (see Tables D13a and D16a for Year 7 and Year 9 sample respectively).

Finally, the results obtained using the interaction models (Tables D13b and D16b in Appendix D) suggest that, for both Year 7 and Year 9, being eligible for FSM does not have an impact on what effect the intervention has on this secondary outcome.

Impact of Paired Reading on the sentence completion: Gender subgroups

Similarly to previous analysis described above, separate analyses were run for each gender to test the effect of the programme on sentence completion (secondary outcome). The analysis was implemented in a similar fashion to the previous sections.

The results indicate that for Year 7 boys (639 pupils) the treatment seemed not to have an effect on the level of sentence completion. The difference between the groups that we are able to highlight is equivalent to an effect size of -0.03 and is not statistically significant (95% CI:-0.18 to 0.12). The result for Year 7 girls (665 pupils), also show no effect. The effect size is equal to -0.03 and is not statistically significant (95% CI: -0,15 to 0.12). The results (including the interaction analysis) are presented in Appendix E, Tables E17a, E17b and E17c.

For Year 9 boys (669 pupils) the analysis, once more, indicate the lack of any effects. The effect size is equal to 0.04 and does not attain statistical significance (95% CI: -0.07 to 0.15). Looking at Year 9 girls (606 pupils) we also find no effects. It appears that when it comes to sentence completion the treatment does not have an effect on Year 9 girls who act as tutors. The effect size is equal to 0.06 but does not attain statistical significance (95% CI: -0.18 to 0.06). The interaction model supports this finding. The results are included in Appendix E, Tables E21a, E21b and21c.

Cost

The costs connected with the delivery of the Paired Reading programme have been provided by the delivery team. Several assumptions have been made when considering the cost of the programme to schools. These include:

- Delivery to a medium sized school with 5 class entry per year group.
- All Year 7 and 9 pupils would take part with an average of 25 pupils per class, resulting in 250 participating pupils per school.
- The programme involves 30 minute sessions once a week and lasts 16 weeks.

Training and support costs. Professional development and additional support for staff is recommended, particularly in the early stages of setting up a programme. It is estimated that this would cost a total of £2,000 per school.

Materials required for schools would be a copy of the manual for each teacher involved and resources for pupils to use (including question and praise cards). The costs for these would be around £625 per school (or £5 per pupil pair, £2.50 per child).

There was no additional workload intended for teachers to deliver the programme so no extra costs were associated with this.

Based on the assumptions listed above the cost of the programme delivered over 16 weeks, is **£2,625** per school. This translates into a unit cost of **£10.50** per pupil per academic year.

Process evaluation

The process evaluation helps us to address three important questions:

1. Were there issues in setting-up the Paired Reading programme?

   This question helped us to understand the process of setting up the programme in the school; including the strategic decision to join, the pairing of pupils and finding space in the timetable. It also helped us to examine how the teachers who delivered the programme in the schools felt these aspects could be improved if it were to be implemented more widely.

2. Was there any variation in fidelity to the running of the Paired Reading programme?

   This question helped us to explore how the programme and 'Paired Reading' classes worked in practice and any implementation, fidelity or sustainability issues that were faced by the schools involved. It also helped us to look into issues around the sustainability of the programme.

3. Were there any perceived benefits reported by staff to pupils' reading ability, or any wider impacts of the programme?

   This question helped us to find out how teachers and senior leaders in schools perceived the programme to be making (or not making) an impact; whether on the pupils or on the school at a wider level.

Implementation

This section covers issues around preparing to deliver the programme in schools, including separate sections about:

- Why schools took part in the Paired Reading programme
- Literacy interventions already in place
- Teacher training
- Training pupils and Introducing the programme to them
- Pairing of pupils
- Timetabling the programme sessions

Participating in the Paired Reading Programme

Schools were initially approached by North Tyneside Council to take part in the Paired Reading programme. The decision to participate was guided by either pupil-based or school-based factors.

Pupil based factors:

- Reading levels. The programme was seen as an opportunity to improve reading levels of both year groups. Senior leaders hoped that Year 7 pupils' reading would improve by being mentored by an older pupil while the Year 9 pupils would also improve due to the process of mentoring and correcting a younger pupil.
- Confidence. Senior leaders hoped that the programme would be a pleasurable process that would build the confidence of pupils. They hoped it would help to build relationships

between pupils across different age groups and promote a friendly community atmosphere around the school.

School based factors:

- Staff development. The programme was seen as an opportunity for the teaching staff to develop professionally. Senior leaders wanted the teaching staff to lead on a project and develop their own leadership potential.
- Research participation. Schools were interested in taking part in research and trialling a promising approach which had previously been shown to have positive impact on student performance.

"We're a teaching school. Part of our drive is looking at things that actually work. What is behind teaching and learning is really important to me"

(Senior leader)

These were also some apprehensions about taking part:

- Administration and extra work - There was a perceived risk that the programme may increase the administrative burden for schools.
- Timetabling - Some schools felt they would have difficulty fitting the programme into their timetable.

Teachers and senior leaders were largely positive about taking part in the programme. It was seen as an opportunity to do something different, though for some the idea of the peer mentoring programme was so radically new and unprecedented, they could not imagine how it was going to look, or be run, in practice.

Other literacy and reading interventions

It is important to consider that Paired Reading was not the only literacy intervention that was taking place in the participating schools. Schools already used a variety of different techniques and interventions with a view to improving literacy with their pupils.

The existing interventions in the schools ('business as usual') included:

- Literacy Clubs: These tended to occur weekly and were available to all pupils. The pupils would come to the club and read a book together with the others, taking turns reading passages from the book. They would then, as a group, discuss the book with a member of staff.
- Additional reading support groups: These were run by either staff within the school or volunteers from the community and involved one-to-one literacy support to pupils who had been identified as requiring it. They would take place outside of school hours.
- Withdrawal groups: These were run with pupils who were identified as requiring additional support with their reading. The pupils would be withdrawn from a standard timetabled class (usually a foreign language class) and given intensive literacy support by support assistants at the school.
- Specific 'Literacy' lessons: Schools also included 'Literacy' as a standard lesson in the timetable where the class would read excerpts from novels and cover other skills such as spelling, punctuation and paragraphing.
- Dedicated reading time: Reading was considered to be a very important part of the syllabus with some schools dedicating 15 minutes of every English lesson purely to reading, with an 'accelerated reader' hour on top. Schools that followed this were having at least two hours of reading per week with their pupils. The schools involved also gave significant importance to

literacy as part of their syllabus. This suggests that the schools involved had a particular interest in literacy, which may have influenced their decision to join the pilot.

Introducing the programme to the pupils

Teachers were responsible for introducing the Paired Reading programme to their pupils. They described presenting the programme in a way which would appeal to their pupils to encourage enthusiasm from the outset.

Some of the schools chose to inform their pupils by organising a special assembly (for all Year 7s or Year 9s separately, or both together) whereas others introduced it to individual classes. All schools focused on the individual benefits the programme would potentially bring pupils and particularly their reading ability.

Some schools felt that this programme tied in well with the Citizenship syllabus, where the pupils were taught personal skills, communication and being a responsible citizen. The school linked the Paired Reading programme with the Citizenship syllabus and explained to the Year 9 pupils that it was part of their role as citizens in the community.

'They were helping others and benefitting themselves in the process'

(Teacher)

Teachers described a mixed reaction from the pupils once they had been informed of the programme, although generally they were positive towards it.

'I thought they were really good. The Year 9s liked the responsibility they had been given over the Year 7s and the Year 7s enjoyed having that reading buddy to talk to and do a bit of work with'

(Teacher)

There were some difficulties especially with Year 9 pupils not understanding how they would benefit from the programme. Running the programme outside of normal timetabled hours or replacing private reading time also caused resentment towards the programme in some cases.

Training the pupils

Schools were advised to train both tutors and tutees together to ensure that they all received the same message and that the idea of 'we are all learning together' was immediately conveyed. It was recommended that the training time be specified well in advance, so pupils could look forward to the experience. As well as a training session, schools were also advised to use the first four sessions of the intervention as a gradual introduction.

In practice, some schools used the resources provided during the teacher training, in order to train the pupils (generally feeling that the resources, whilst aimed at adults, were appropriate for Year 7 or Year 9 pupils to use). However, while the pupils could 'understand' the resources, there was a view that pupils did not find the materials particularly exciting. Other schools therefore created their own resources. Many teachers also used the online videos provided by the delivery team in their training.

On the whole, schools did not report any major problems with training the pupils. It was suggested that the delivery team could produce materials that were more engaging for the younger age group which would help to deliver more consistent messages to the participants and excite them about the programme.

Pupil training took place either with Year 9s and Year 7s together, or the two age groups separately. The delivery team suggested that year groups of pupils are gradually introduced to the programme

together, step by step but there was further variation in terms of the way schools trained pupils which are outlined below.

Gradual introduction. The training took place as the intervention started (as prescribed) but its length varied from school to school. Some schools used just the first lesson to train the pupils, while others provided a staged introduction to the programme, with the first few lessons dedicated to training and practice activities. While this may have reduced the dose of the intervention for the pupils, teachers felt it was necessary for pupils to fully get to grips with the programme and what was required of them before embarking on it.

'The first lesson we got them reading together, the second lesson we introduced questioning and the third week we did a trial and the programme actually started in the fourth week'

(Teacher)

Only training Year 9 pupils. Schools spent the first two lessons or so going through the PowerPoint presentation with Year 9 pupils, explaining how the programme should be run, the rationale behind it, the structure and what their approach should be. They were told how they should support the Year 7s, give verbal feedback and use the question cards. Schools that took this approach then needed a lesson to practice before embarking on the programme fully. The Year 7s did not receive any additional training.

No advance pupil training. One school took the approach of not training the pupils at all. In this case, schools spent a little bit of time in the first lesson going through the five finger test and selecting the books. However, after this, pupils went straight into the programme and were considered *'ready to go'*. Though this was not by any means the prevailing approach taken by schools, it was felt in this case that the pupils did not require an extended training or preparation period.

Pairing pupils

Prior to starting the programme all pupils in Year 7 and Year 9 took a literacy test to establish their reading ability. For the purposes of the project this assessment was also the pre-intervention test to avoid additional testing. Year 7 and Year 9 classes were paired based on these test results by the evaluation team.

Teachers were responsible for pairing the pupils individually. The delivery team suggested that reading test results should be the primary factor considered when pairing pupils. They also suggested that teacher knowledge (such as the maturity and behaviour of the pupil) and outside factors including attendance should be taken into account. In practice, teachers used a mix of test results, teacher knowledge and other outside factors when pairing the pupils individually:

- Test results only: Pupils with the closest reading ages were paired together. No other factors were taken into account.
- Test results and teacher knowledge: Test results and the personalities of the pupils and whether they were likely to get along with each other were considered.
- Test results, teacher knowledge and other outside factors: Factors based on individual (non-academic) circumstances which would make it inappropriate to match pupils together such as a dispute between the families of the pupils, were considered along with test results and other teacher knowledge.

When schools had an odd number of pupils in either Year 7 or Year 9 they the delivery team suggested to put some in trios rather than pairs. This was also the case when one of a pair was absent. A trio could be with two Year 7s with a Year 9 pupil or two Year 9s with a Year 7 pupil. In practice, teachers used this method and in one case, when a Year 9 pupil was absent a teaching assistant would take their role.

Timetabling the programme

There was no guidance on which particular lessons, within the school day, the intervention should replace. Schools were found to have timetabled the programme in the following ways:

- Replaced an existing class - Schools took this approach out of choice or necessity:

  - Some had an appropriate class that could be replaced, such as an English lesson or Citizenship lesson, which they felt worked well with the programme.
  - Others replaced a modern language and ICT class. It was felt that the reading levels of the pupils were more important than progress in these subjects and they did not have sufficient registration time or timetable flexibility to make alternative changes.

'By splitting the lesson in half it created a very staggered lesson. If we're doing Shakespeare, you go through the scene, you're about to start the work and then it's paired reading, so it will have to be pushed back to the next lesson'

(Teacher)

- Replaced registration or tutorial time - This was an easier task for schools that had a specific time set aside in the mornings for this purpose. If schools had 20 minutes or more tutorial time available they fitted the programme into this period.
- Ran the programme outside of school hours - Where not possible to replace an existing class or run the programme during registration or tutorial time the only remaining option was seen to be running the programme outside of the normal timetable hours. One school (6 classes) brought participating pupils in early one morning a week and held the programme before other lessons began for 15 minutes pre registrations and 15 minutes during registration. This relied on the goodwill of the teaching staff to do so.

Staff interviews indicated that timetabling was felt to be easier for schools that were aware of the programme requirements before they set the timetable for the period covered. It was not always easy to replace an existing class in a set timetable. Some schools reported disruption when the programme replaced an existing class, or part of a class. There was also concern about pupils missing other lessons. When the intervention group was doing Paired Reading for half a lesson once a week, these pupils missed a lot of other lesson time over the 16 weeks of the programme. It also required teachers to plan their lessons in a different way to allow for certain pupils missing sections of teaching.

This inconsistent timetabling of the programme amongst schools meant that the length of sessions, and the amount of time each pupil spent reading, varied between schools. Some schools were able to fit in a 40 minute Paired Reading session while others had to add a 20-minute slot onto the school day. This meant that pupils from different schools were receiving varying amounts of Paired Reading.

Wider programme constraints

The main issue that was raised by members of the senior leadership teams concerned the levels of administration the programme entailed. This related mainly to the evaluation of the programme, rather than the programme itself. While senior leaders did understand that the programme was being trialled with a view to being extended elsewhere, it appeared there was a misunderstanding about which aspects of the administrative tasks related to the programme itself and which aspects related to the evaluation of the programme.

Senior leaders at schools listed the following as some of the administrative tasks their staff have had to complete during the programme:

- Case studies
- Logs/diaries

- Questionnaires
- Reading tests for all pupils

There were occasions when the amount of administrative tasks had an effect on the school. For example, one senior leader was forced to give the teachers involved with the programme a free period to allow them to catch up with and get through the paperwork. The senior leader felt that, without this free period, there simply would have been too much for them to do. There was a view that the many different strands of data collection and administration that needed to be coordinated were putting a strain on the staff involved.

'Now and again I felt there was a lot and the staff were under pressure. At times I felt there was possibly too much being put on us'

(Senior leader)

The administrative aspect of the programme and its evaluation led to the following constraints being raised:

- Programme length - There was a view that the programme went on too long and this resulted in an erosion of the enthusiasm that the staff and pupils initially felt.

'By the end, I think everybody was getting a little bit sick of it, the students were getting bored and for the staff the novelty wore off, it was just a little bit too long'

(Teacher)

- Busy time of year - Teachers commented that the programme was held at a very busy time of the year. The second and third terms are when schools are preparing Year 9 pupils for their GCSEs, which will begin in the next academic year. There was a view that the Paired Reading programme added to teachers already large workload and made it difficult for them to fully engage and monitor the progress of the pupils formally.
- No instruction for assessment - Teachers reported receiving no instruction on how to assess the pupils as the programme progressed, either formally or informally. This was particularly an issue where schools had half of their form group with another teacher for their programme sessions and so were not aware what progress was being made.

Fidelity

This section looks at issues around the delivery of the programme, including levels of fidelity and any barriers that schools faced with delivery.

- Book selection
- Lesson content and format
- Role of the teacher
- Programme issues

Book selection

Participating schools took two different approaches to select books to read:

- The Year 7 pupils would select a book from among the school's own resources. The pupils would then use the 'Five Finger Test' (see Intervention section above) to ensure the book was neither going to be too easy nor too difficult for them to read.
- The teacher would pick the books for the pupils by suggesting two or three books to each pupil and then allowing the pupil to choose which one they would like to read.

The reasons schools chose to select books for pupils, instead of relying on the 'Five Finger Test' were around not feeling that this technique helped to find a suitable book.

'Occasionally they would do the five finger test and still find a book that is too hard'

(Teacher)

It was also apparent that teachers' understanding of how to do the 'Five Finger Test' test varied.

'Any reader could pick up a book, put their five fingers across and find 'and', 'or', 'because', 'but', 'other' and think 'Oh I can read this book'. So at secondary level there needs to be a far more rigorous way. There need to be recommended novels. It doesn't work at all'

(Teacher)

The programme resources did not state that pupils had to choose a book from the school. On the contrary, they suggested pupils had a choice of material, including for example, football match programmes. However, it was not always possible for pupils to bring their own books or reading materials from home. Teachers commented that many of the pupils from deprived backgrounds would not have access to books at home. For many of them, school would be the only time that they had access to books and the opportunity to read them.

The programme guidance was for the younger pupil to select the book. However, some schools tried to allow the Year 9 pupils to choose the book. The idea was that this would help the older pupil to become more engaged as they would be more interested in the reading material. More often than not, however, the book chosen was not suitable for the Year 7 pupil to read.

Teachers felt that the best solution for the selection of books was for Year 7 pupils to choose the book from a recommended reading list based on reading age and covering a variety of genres.

Lesson content and format

Overall, teachers felt that the Paired Reading sessions worked reasonably well. This was despite apprehensions about having a large number of pupils reading out loud in the same room.

'At first I thought it would be far too much noise to have all those pupils reading but it was actually fine'
(Teacher)

Teachers felt the programme took a little time for the pupils to get used to and for them to use the prescribed techniques.

'It worked for the most part really well but it took a long time to engrain. For some groups we felt like we were just getting it and the programme came to an end'

(Teacher)

Whilst teachers found that some pupils were very good at following the steps, having a reminder in front of them certainly helped. Teachers felt that without a step-by-step guide pupils (especially those with lower ability levels) would have struggled to remember what to do. Teachers were also breaking the sessions into several shorter slots of reading, to give lower ability students more frequent breaks in the task.

Correcting pronunciation. Tutors were good at correcting the pronunciation of words. The Year 9 pupils were correcting the pronunciation of words but were not explaining the meaning to the tutee, nor providing an example of its use in another sentence. Teachers felt this was a missed opportunity to develop both pupils' literacy.

Some pupils found the correcting process more difficult. There were examples of the tutors waiting until the tutee said they were stuck before correcting it. These pupils needed extra encouragement to read together with the Year 7 pupil and follow the steps correctly.

Questioning. There were several sets of question cards that pupils could use depending on ability level. As a result of this, higher ability pupils were asked more difficult questions than lower ability pupils. Examples of the suggested questions include:

- Which of the characters is most like you, and why?
- How would you cope if you were transported into the story?
- Describe what might happen. Can you justify why you think this?

Schools took two different approaches to this aspect of the programme:

- There was a set time in every class when pupils would be instructed to stop reading. The Year 9 pupils would then spend 5 minutes or so asking questions to the Year 7 pupils.
- There was no set time with Year 9 pupils being instructed to ask questions at the end of a chapter or at the end of an appropriate page.

The success of the questioning aspect of the lesson was felt to depend on two factors:

- Lesson length - The longer the lesson that the programme replaced, the more successful this aspect was. Schools with shorter sessions carried out less questioning and sometimes included questioning at the start of a session about what pupils had read in their previous lesson. This was seen by teachers to make this aspect of the programme more difficult for the tutees, as the story would not have been as fresh in their mind as if they had just been reading the book.

'The writing part was absolutely fine, they were articulating it but we only had 20 minutes so that was one of the parts that was cut short because we needed more reading time. With an extra five minutes that would have been a lot more effective'

(Teacher)

- Question cards - Without these, teachers felt that most pupils would have struggled to think of or ask appropriate questions. The majority of pupils only used the question cards and did not use their own questions. The success of this aspect of the lesson was therefore felt by teachers to depend on how well the tutor used the question cards and asked the appropriate question at the appropriate time. Higher ability pupils were more likely to move away from the question cards and think of their own questions to ask, which the lower ability pupils were unable to do.

Praising. The Year 9 pupils had 'praise cards'. These were used to praise the Year 7 pupils when they read well or pronounced a difficult word correctly. Teachers at participating schools felt that the Year 9 girls were better than the boys at this, as the boys tended to struggle to go beyond saying 'well done'. The boys in Year 9 needed more encouragement from the teacher in order to praise effectively.

Log books. Pupils and teachers had to fill in a log book at the end of each lesson. These recorded the book and amount of reading done, what pupils had read well and what they needed to work on for the next lesson. Examples included:

'Work on your pace', 'Try to de-code words better', 'Pronounce your Ts'.

Teachers reported that the students with lower literacy abilities struggled with this part of the programme. They tended to write the same thing in their log books every week.

'They didn't have the literacy levels to vary what they wrote after each lesson'

Teachers felt there was a need for more differentiation between higher and lower ability pupils. This could be achieved by providing less scaffolding to the higher ability pupils but more for the lower ability pupils to better support their learning. The schools overcame this by the teacher or teaching assistant providing extra help to pupils who needed it.

Senior leaders and teachers also felt that some of the pupils did not understand the need to fill in log books at the end of each session. This was something that required behaviour management and persuasion from the teacher to ensure it was done properly.

Pupil pairings and the use of trios. Overall, teachers felt trios, when they had to be used, worked well, however they identified some problems which were not present with pairs:

- When two Year 7s were together with one Year 9, it slowed down the amount of reading covered, as the Year 7s were each taking turns to read their book.
- The Year 7 pupils tended to get bored when they were not reading. This resulted in behaviour issues with pupils talking to or distracting others.
- It was difficult for one Year 9 pupil to focus on the reading of both Year 7s. Teachers felt this resulted in less correction and the quality of the 'buddying' being lowered.
- It was also difficult when there were two Year 9s with a Year 7. Not only was it intimidating for the Year 7 but one of the older tutors would invariably become more dominant and the other would become bored and distract others.

'They had to sit squashed on either side of this poor Year 7 and some of our year 9s are huge and the Year 7s are tiny so it was a little intimidating for them'

(Teacher)

Role of the teacher

The teacher had an important role in delivering the programme. This included demonstrating the programme approach and monitoring pupil progress. The monitoring of pupil progress occurred by observing and listening to them as the programme progressed but also by reading the pupils' log books. Teachers were instructed to intervene during their monitoring of pupils but only when a pair was having difficulty to the point of being unable to move forward with the process. However, there were times when teachers needed to take a more pro-active role with certain pairs. This may have been due to personality clashes (such as over-dominant partners or gender issues) or poor communication skills. Furthermore, teachers were also instructed to intervene if pupils had selected books that were either easy or too difficult. The manual suggested that teachers take turns at observing pairs with each observation lasting approximately six minutes.

One of the issues teachers faced when monitoring the progress of the pupils was that of a lack of time. This was a particular problem for those schools which had shorter Paired Reading lessons.

As for programme related issues, teachers felt there were some issues which limited how closely the schools were able to adhere to the prescribed approach.

Teachers additionally identified a number of perceived barriers to pupils' tutoring or being tutored:

- Cross-age tutoring - There were some concerns about the social aspect of the programme, with pupils of different ages finding it difficult to work together. This was caused by either or both of the year groups feeling shy or uncomfortable with the other. Some Year 7s had not previously interacted with Year 9 pupils in any way before the start of the programme.

'Our students aren't on vertical tutor systems. It is very horizontal and the Year 9s and Year 7s very rarely mix together'

(Senior leader)

For many pupils, the first Paired Reading lesson was the first time they had interacted or worked with a pupil from a different year group. It was felt schools may need a little more time for the groups to get to know each other so that the social element was less of a constraint in the running of the programme.

- Personality clashes - Some of the pupils simply did not get along very well, or were disruptive when paired together. Teachers found that this affected the entire class and so they tried to pre-empt or deal with issues as soon as possible.

'If it had been students who had volunteered or with Year 9s and Year 11s, it may have been better'

(Teacher)

- Adhering to the approach

The programme was felt to contain a lot of information for the Year 9 pupils to take in, particularly when they first started. There was a view that immaturity was preventing some Year 9s from engaging with the programme. It was commented that they were not mature enough to understand that they were in a role of leadership and responsibility. To counteract this, some of the schools held separate training sessions for the Year 9 pupils. However, it was felt that a more formalised training process for the Year 9 pupils would help them to understand their role more clearly.

- Being paired with lower ability Year 7s

When lower ability Year 9s were paired with lower ability Year 7s teachers felt neither improved their reading ability nor their confidence. It was often the case that the lower ability Year 7 had a higher reading age than the Year 9 and this resulted in a considerable loss of confidence and morale for the Year 9 pupil. Lower ability pairings which were struggling with their reading, also tended to have behavioural issues during the class.

'From that they thought they're rubbish and they're crap. Why should they bother? Why should they have to do it if they're not going to get any better?'

(Teacher)

Perceived impacts

Teachers and senior leaders identified a number of areas where they felt the programme had an impact on the school and the pupils, both positively and negatively.

Literacy and transferable skills

- **Literacy.** Although the impact evaluation has found no significant impact on the reading levels of participating pupils, teacher's commented on the ease at which some of the Year 7 pupils were reading at the end of the programme. Some of the pupils also finished the books they were reading previously and were moving onto, what the teachers described as, *'harder books'*. Not all senior leaders we spoke to were convinced that the programme had, or would, improve literacy levels among the pupils due to the lack of directed teaching.

'Two kids reading together is fine but you can't get a Year 9 kid to teach a Year 7 kid how to decode a word'

(Senior leader)

- **Transferable skills.** Teachers and senior leaders spoke of the transferable techniques that had been gained from this programme and used in the broader curriculum. For example, the questioning structure which teachers believed made the pupils think more about their reading.

Non-cognitive skills

- **Improved relationships.** The programme was felt to improve relationships between pupils with the pairs having a positive social experience from peer tutoring. In some cases, teachers felt that pupils who had never spoken to someone two years above them had gained in confidence as a result of their interactions on the programme.
- **Confidence of pupils.** Teachers spoke of a visible improvement in the confidence of some pupils, in both years, who took part in the programme. The prevailing view was that, whilst at the beginning some of the Year 7s may have felt a little intimidated, by the end the fluency with which many of the pupils were reading, questioning and praising, had improved.
- **Loss of confidence for lower ability Year 9s.** While some pupils in Year 9 gained in confidence, teachers felt that especially for the lower ability Year 9 pupils, the programme had the opposite effect. It was felt they were embarrassed about having a reading ability the same as or lower than a Year 7 pupil and so disengaged from learning.

Wider school impacts

- **Up-skilling teachers.** Senior leaders felt that the programme had also brought the added benefit of up-skilling teachers in the literacy element of teaching. This fell into the wider school improvement plans for schools.

'You look at literacy for school and people think it sits with the English department but we've had a big push on that. Everybody is a teacher of literacy'

(Senior leader)

- **Profile of teachers.** Senior leaders felt that the programme had provided a clear focus to the teachers and those who ran the programme now have a higher profile within the school.

Pupil attainment

- **Students falling behind.** Teachers felt that participating pupils fell behind their peers in subjects that the Paired reading sessions replaced. In total, they were taken out for at least 30 minutes a week over 16 weeks and there was a worry that this may result in them doing worse in their assessments than the control group pupils.
- **Impact of existing programmes.** Some senior leaders felt that their school's own existing literacy programmes were better than the Paired Reading programme.

Formative findings

There were mixed reports on whether participating schools would continue with the Paired Reading programme. In general schools wanted to see whether it had an impact on the reading ability of students before they committed to continuing with the approach.

Schools were concerned about the implementation of the programme including timetabling the lessons. There were concerns that missing other subject lessons was having an adverse effect on attainment. This would clearly influence schools in whether they removed pupils from other lessons, or held sessions during tutorial time, or added sessions on to the school day.

However, the concept of peer tutoring was very well received and schools were more likely to continue with this practice. It was the fully prescribed approach of the Paired Reading programme which schools found challenging to follow when implementing the approach routinely in their school.

Members of the school senior leadership teams and teachers generally felt that the concept of the Paired Reading programme was sound. However, for the programme to be successful, the following aspects should be considered:

- Programme length: Some teachers and senior leaders felt that the programme needed to be shortened to around a term's length, rather than 16 weeks. It was argued that, for the pupils, Paired Reading felt like a regular lesson but with 16 weeks of repeating the same thing. Some pupils therefore lost focus and enthusiasm. This was further compounded by the rigidity of the method, with little scope for variation.
- Timetable constraints: The programme would easily slide into a school that has a significant time for registration in the morning. However, schools that do not have a 'collapsed' lesson struggled to fit it in and would need to see evidence of impact before committing to a change in the way they set up the timetable.

The programme was perceived to have the potential to improve the reading levels of pupils at a minimum cost to the school. Teachers commented that the gold standard in terms of a child's ability to succeed in life is their ability to read. Teachers and senior leaders identified a number of changes which they felt the programme should adopt if it were to be rolled out further in future:

- Motivation for the Year 9 pupils to take part in this programme. One teacher suggested that they be given a certificate at the end of the programme to commend their participation.
- Year 9 pupils need some additional support in the form of mentoring training to help them understand and develop their tutoring role and reduce the burden on them as they embark on the programme.
- There was a suggestion to use 'colour coding' for the question cards to help pupils differentiate their difficulty more easily.
- There needs to be more scope to develop 'understanding' of words, rather than simply how to pronounce it. Teachers commented that the best way to learn how to pronounce a word is to know what it means and how it can be used in different circumstances. This element was seen as being necessary to include in a literacy programme for this age pupils.

Control group activity

While pupils in intervention classes were taking part in the Paired Reading programme, the pupils in remaining Years 7 and 9 classes continued their 'business as usual'. This meant that they continued in their normal timetabled lessons and normal classes. The pupils in control group classes may have also been receiving additional literacy support, such as those discussed in section 'Other literacy and reading interventions' section of the process evaluation chapter. Some of the schools were also running other targeted literacy interventions for pupils, which intervention and/or control group children could have been receiving. These data were not collected.

Conclusion

Key Conclusions
1. This evaluation does not provide any evidence that the Paired Reading programme had an impact on overall reading ability, sentence completion and passage comprehension of participating pupils.
2. There was no evidence of the Paired Reading programme having an effect on overall reading ability, sentence completion and passage comprehension of FSM pupils.
3. There was some variation in the intervention group schools in terms of the programme set-up and delivery. There was also a varying level of support provided to pupils within the intervention by the teachers involved, based mainly on the reading ability of the pupils. However, these appear to be natural variations between the settings of the schools involved and are unlikely to have affected the dosage of the intervention for the pupils involved.

This evaluation has sought to provide new evidence on the effectiveness of an intervention that uses cross-age peer tutoring method to improve the reading ability of pupils in Year 7 and Year 9. The results of analysis suggest that the Paired Reading programme had no positive effect on the reading ability of pupils participating in the programme and a negative effect on Year 9 girls' outcomes.

Limitations

Considering that all pupils in intervention cohorts were intended to receive the Paired Reading programme, the pupil recruitment was successful. There was no attrition at the level of randomisation (i.e. classes). As for pupils, 95 per cent of those allocated to the intervention group and 94 per cent of those allocated to the control group were eligible for unadjusted primary analysis yielding a pupil level attrition rate of 5 and 6 per cent respectively. Moreover, the sample characteristics in terms of observable variables were similar for the randomised and the analysed sample and any differences in intervention and control group composition are likely to be due to chance.

It has to be acknowledged that the key limitations of this cluster randomised trial design is that randomisation at class rather than school level increases the possibility of 'contamination' (exchange of information among pupils and teachers in different trial groups) and increases likelihood of collaboration among teachers within each school, and thus the chances of treatment diffusion (i.e. some aspects of the experimental stimulus such as the intervention are passed on from the intervention group to the control group).

This trial was undertaken within a narrow geographical area of the UK. More specifically, the trial was run in three LA areas in North East England – North Tyneside, South Tyneside and Sunderland. The Paired Reading programme was delivered by teachers in a regular school context which makes the trial results more applicable to a real-world setting. The proportion of FSM pupils in trial schools ranged from approximately 10 per cent to 35 per cent. On average 21 per cent of pupils in trial schools were eligible for FSM which is higher than the national average in state funded secondary schools (15 per cent, DfE, 2014b). As for pupils with SEN, the trials schools had a higher percentage of pupils in this group 21 per cent), compared to 18 per cent across all pupils in state funded secondary schools (DfE, 2014c). Furthermore, the trial schools had much lower percentage of pupils with EAL (less than 4 per cent) compared to the national average (14 per cent, DfE, 2014b). Overall, the sample characteristics suggests that the schools participating in the study were slightly more disadvantaged than the average secondary school in England (as intended) while having less EAL pupils when compared to the national average. It also has to be acknowledged that the pragmatic approach taken to the trial design is likely to mean that the treatment effect is estimated on more engaged schools (and classes within schools) where the logistics of the standard school day could have affected how the classes were paired to allow easier implementation of peer tutoring.

It has to be noted that as part of the process evaluation, a short online teacher survey was originally planned to gather further information from teachers about training that they received, support they have access to, perceptions of the programme, implementation issues and fidelity to the programme. The survey results were to be used to assist the selection of a sample selection for the process evaluation. However, a decision was made not to proceed with the survey as schools felt strongly overburdened by data collection demands as part of this project and further data collection was deemed to be unfeasible. (It had already led to two schools dropping out of the process evaluation at a decision-making stage.)

Furthermore, there were no observations of the intervention itself so researchers were unable to see a 'Paired Reading' lesson in practice. The result of this was that researchers had to rely on teachers to comment on issues of fidelity or how the programme worked in practice. It is therefore not possible to verify the observations and comments of teachers to be an accurate reflection. As such, conclusions around the fidelity of implementation should be seen as indicative rather than conclusive.

Finally, even though specified in the protocol, we unable to carry out separate analyses for pupils with EAL due to very low sample sizes. Furthermore, the current report does not include SEN subgroup analysis on outcomes of interest. However, the project team will be carrying out further analysis to assess the impact of the programme on SEN pupils.

Interpretation

The results of this cluster randomised controlled trial do not provide evidence that the Paired Reading programme had an impact on pupils' reading ability. Even more, the findings seem to indicate that the programme was in fact detrimental to Year 9 girls (who acted as tutors) levels of reading ability (primary outcome) and passage comprehension (secondary outcome that forms a part of the primary outcome). This finding is inconsistent with existing research (e.g. Tymms, 2011) that shows overwhelmingly positive effects of peer tutoring on raising attainment in school-aged children across a number of subjects.

As indicated above, there is a risk of contamination that trials of educational interventions and treatment diffusion – the process by which teacher and pupils in control classes either gain access to the intervention being evaluated, or receive a similar intervention by way of being compensated for their allocation to control conditions. It is possible that contamination may have reduced the magnitude of effect estimates and therefore also increased the chance that estimates are not statistically significant.

In addition to the risk of treatment diffusion, the process evaluation found that there had been some variation amongst participating schools with regards to the set-up, delivery and implementation of the programme. For example, it was recommended to run the pupil training in four sessions. However, schools used a range of approaches to train the pupils involved, from not giving any advance training at all to using the first few sessions to gradually introduce them to the process. Similarly, the length of the sessions varied from school to school, with some able to use the recommended half an hour, while others were forced to condense the lesson into a shorter time period. There was also a varying level of support provided to pupils within the intervention by the teachers involved, based mainly on the reading ability of the pupils. However, these appear to be natural variations between the settings of the schools involved and are unlikely to have affected the dosage of the intervention for the pupils involved.

The schools involved in the programme also provided a range of literacy initiatives, in addition to the Paired Reading programme. As mentioned previously, these included literacy clubs, support groups and withdrawal groups in addition to literacy support and focus within normal timetabled lessons. This suggests that the schools involved already placed an emphasis on reading and therefore would have been fairly engaged in literacy issues. It is likely that such schools would have given the Paired Reading programme their best attempt to make it a success, but that due to the existing emphasis on reading in

these school it might be difficult for a programme such as Paired Reading to generate marginal improvements in outcomes that might be considered meaningful and thus worthy of investment.

Furthermore, it has to be acknowledged that 'business as usual' in participating schools could be accessed by both the intervention group and the control group which makes it somewhat challenging to interpret the results. In some of the cases (20 classes) it was the English classes that were replaced with Paired Reading, whereas in the majority of cases (32 classes) it replaced classes other than English. Therefore, the findings of this study reflect a mixture of providing Paired Reading instead of English classes in some of the cases and providing Paired Reading in addition to English classes in other cases.

Future research and publications

As indicated there is a body of literature showing the positive effects of peer tutoring on pupil outcomes. In the light of the current findings one area for future research is to explore the core components of peer tutoring interventions that make them work. Furthermore, additional research on the level of technical support required to deliver the programme with fidelity would benefit the programme.

References

Cohen, P. A., Kulik, J. A., & Kulik, C-L. C. (1982). Educational outcomes of peer tutoring: A meta-analysis of findings. *American Educational Research Journal, 19* (2), 237-248.

Department for Education (2014a). *How do Pupils Progress During Key Stages 2 and 3?* Research Report DFE-RR096 London: DfE. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/182413/DFE-RR096.pdf

Department for Education (2014b). Schools, pupils and their characteristics: January 2014. Statistical First Release 15/2014. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/335176/2014_SPC_SFR_Text_v101.pdf

Department for Education (2014c). Special educational needs in England: January 2014. Statistical First Release 26/2014. Available at: https://www.gov.uk/government/statistics/special-educational-needs-in-england-january-2014

Higgins, S., Kokotsaki, D. & Coe, R. (2011). Toolkit of strategies to improve learning. Summary for schools spending the pupil premium. Available at: http://www.bristol.gov.uk/sites/default/files/documents/children_and_young_people/child_health_and_welfare/children_in_care/What%20works.%20Sutton%20Trust.pdf

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology, 95* (2), 240–257.

Thurston, A., & Cockerill, M. (2013). Peer tutoring in secondary schools. Belfast, UK: Queen's University Belfast.

Tymms, P. (2004). Effect sizes in multilevel models. In I. Shagen & K. Elliot (Eds.), *But what does it mean? The use of effect sizes in educational research* (pp. 55-66). Slough, Berkshire: National Foundation for Educational Research.

Tymms, P., Merrell, C., Thurston, A., Andor, J., Topping, K., & Miller, D. (2011). Improving attainment across a whole district: School reform through peer tutoring in a randomised control trial. *School Effectiveness and School Improvement, 22* (3), 265-289

Appendix A: Parent opt out letter

Peer Tutoring in Secondary Schools

Helping to raise standards in Literacy

We are delighted to introduce the Peer Tutoring in Secondary Schools programme – an initiative that aims to raise standards in reading through peer tutoring. This project is a partnership between North Tyneside Council, Queen's University, Belfast and Durham University. The project is funded through a grant from the *Education Endowment Foundation.* All qualifying schools in North Tyneside and some schools in neighbouring Authorities are being invited to participate in this project.

Excellence for All

The Peer Tutoring in Secondary Schools programme will help to raise reading levels by improving learning. It will give teachers new techniques and links well with other developments in schools such as "Literacy across the Curriculum" and "Assessment for Learning". The project will run for three years.

Peer Tutoring Programme

During the project your child will undertake peer tutoring in reading. Students in Year 9 will tutor students in Year 7, for thirty minutes per session over sixteen weeks in total. During peer tutoring the teacher organises materials, manages the class, and monitors the learning that is taking place, to ensure quality. Both tutors (Y9) and tutees (Y7) will gain. Such approaches to learning have been proven to raise attainment in schools.

Why are we doing it?

- It should help raise levels in reading.
- Children should be more motivated to learn and more confident in school.
- It should enhance interpersonal skills.

Peer Tutoring in Secondary Schools programme

Paired reading using peer tutoring is a kind of supported reading. It follows a tried and tested method which has been refined over many years and was recently used in a very successful project in Scotland. Emphasis is placed on reading for understanding and enjoyment. It encourages pupils to talk about the text or book and a rapport is built up between the students. Importantly, increases in reading ability bring benefits to all areas of the curriculum.

Your child as part of their class will be randomly selected to either receive the intervention programme and assessments, or assessments only. All students will be asked to take pre and post programme student assessments. Student test responses and any other pupil data will be treated with the strictest confidence.

Students will complete the New Group Reading Test which is a computerised assessment, and responses will be collected online by the school and accessed by the project research team and the external evaluator. Named data will be matched with the National Pupil Database and shared with project research

team, external evaluator, EEF, and the UK Data Archive for research purposes. We will not use your child's name or the name of the school in any report arising from the research.

We expect that your child as part of their class will enjoy doing the tests and being part of the programme. Your child may withdraw from the tests at any time. If you prefer for your child NOT to take part, please inform their teacher.

Further information Details

If you would like to learn more about the Peer Tutoring in Secondary Schools programme, or have any questions about any aspect of the initiative, please contact: Maria Cockerill (Project Lead), Peer Tutoring in Secondary Schools programme, Early Years and School Improvement Service, Schools, Learning and Skills, North Tyneside Council, The Langdale Centre, Langdale Gardens, Howdon, Wallsend, NE28 0HG. Email: **maria.cockerill@northtyneside.gov.uk**

Appendix B: Topic guides

Topic Guide for senior leaders

Aim of the telephone interview:

The aim of the interview with teachers is to provide a broad overview of the programme at a strategic level in the school, as well as its perceived impact and implications for the future

The topic guide:

This guide sets out a number of necessary contextual and factual topics and questions that will be covered during interviews. The guide does not contain follow-up probes and questions like `why', `when', and `how', etc. as participants' contributions will be explored in this way, as far as is feasible, during the 30 minute interview. Researchers will use prompts and probes in order to understand how and why views, behaviours and experiences have arisen.

The interview will last for **approximately 30 minutes.**

---

1. **Introductions**

   - Introduce yourself and NatCen Social Research

   - Introduce the study:
     - Evaluation of the Paired Reading Programme
     - Commissioned by the Education Endowment Foundation
   - Overall project aims:
     - To understand school experiences of delivering the programme
     - To understand the impact (both soft and hard outcomes) that the programme is having on students
   - Interview is one of 3 interviews with head teachers in different schools
   - Digital recording – check OK, and reassure re: confidentiality
   - Data kept securely in accordance with Data Protection Act
   - How we'll report findings – anonymity of teachers and schools
   - Any questions/concerns?

Background and Context

2. **Background**

Aim: To gather background information including detail of the context in which the school operates, the profile of their pupils and the teacher in question

   - Profile of school:
     - Reading and overall literacy levels (particularly year 7)
     - Turnover of pupils at the school
     - Special Education Needs (SEN)

o English as an additional language (EAL)

o Disadvantaged children (based on FSM take-up)

The programme itself

3. Joining the Peer Tutoring Programme

Aim: To provide understanding of how the school joined the programme, the initial expectations and the process

- First hearing about the programme:

    o When was this?
    o Who did you hear about the programme from?
    o What was the initial understanding?

- Deciding to join:

    o What did you hope would come out of it for the school?
    o Any apprehensions?
- Process of joining:
    o Liaising with North Tyneside Council
    o Receiving information and guidance
    o Any improvements to this process

4. The Peer Tutoring Programme

Aim: To provide an understanding of the set up and running of the programme from the perspective of the leadership team and the school as a whole

- Process of setting up the programme
    o Informing staff
    o Informing pupils
    o Informing parents
    o Response to the programme from all the above
- Costs to school
    o Training teachers
    o Training pupils
    o Materials
    o Other running costs?

The future

5. Impact and implications

Aim: To understand the wider perceived impact and implications of the programme for the school as a whole as well as the potential for rolling the programme out in the future

- What changed at school as result of programme?
    o Positive changes, e.g. literacy outcomes, improved community feel at school, any other?

- o   Any negative changes?
- Would implement in the future?
  - o   If so, would do anything differently?
  - o   If not, why not?

Would recommend to other schools?

For Year 7 teachers

Aim of the telephone interview:

The aim of the interview with teachers is to provide a broad overview of the way that tutors for the programme are trained, how the paired reading lessons run in practice, what barriers there are to successful implementation and what broader impacts the programme is having on the students, the department and the school.

The topic guide:

This guide sets out a number of necessary contextual and factual topics and questions that will be covered during interviews. The guide does not contain follow-up probes and questions like `why', `when', and `how', etc. as participants' contributions will be explored in this way, as far as is feasible, during the 45 minute interview. Researchers will use prompts and probes in order to understand how and why views, behaviours and experiences have arisen.

The interview will last for **approximately 45 minutes.**

---

1. **Introductions**

   - Introduce yourself and NatCen Social Research

   - Introduce the study:
     - Independent Evaluation of the Paired Reading Programme
     - Commissioned by the Education Endowment Foundation
   - Overall project aims:
     - To understand school experiences of delivering the programme
     - To understand the impact (both soft and hard outcomes) that the programme is having on students
   - Interview is one of 9 interviews with teachers in several different schools
   - Digital recording – check OK, and reassure re: confidentiality
   - Data kept securely in accordance with Data Protection Act
   - How we'll report findings – anonymity of teachers and schools
   - Any questions/concerns?

Background and Context

2. Background

Aim: To gather background information including detail of the context in which the school operates, the profile of their pupils and the teacher in question

   - Profile of school:
     - Reading and overall literacy levels (particularly year 7)
     - Turnover of pupils at the school
     - Special Education Needs (SEN)

     - English as an additional language (EAL)

- o Disadvantaged children (based on FSM take-up)
- Profile of teacher:
  - o When qualified
  - o Time as teacher
  - o Additional responsibilities (e.g. after school clubs etc)

3. Literacy and reading pre the intervention

Aim: To get contextual information about literacy and reading in the school, for years 7, outside of the Peer Tutoring Programme

- Literacy pre Peer Tutoring Programme
  - o Literacy as part of normal school syllabus for year 7
  - o What interventions in place before/regular format of literacy support
  - o What interventions in place now (for pupils not on programme)
  - o Nature of these interventions- brief information (e.g. how pupils selected, what they go through, who delivers)

The Peer Tutoring Programme

4. Training for teachers

Aim: To provide understanding of how teachers are trained and any preparation that is done by the teacher in advance of the programme starting

- Training for the teaching staff:

  - o Nature of this
  - o How effective this was
  - o How prepared they felt to implement the programme

- Materials:

  - o Which materials used
  - o How easy these are to use
  - o Suggested improvements to these

Anything they would have liked to have seen added to the training?

5. Training for pupils

Aim: To provide an understanding of how pupils are trained to be tutees (Yr 7) and any preparation that is done in advance of the programme starting

- What were Yr7 tutees told about the programme prior to starting
- Provided any training of any kind? If so… (if not, was any prep needed?)
- Format of the training sessions for tutees:

  - o How well these worked
  - o What guidance was provided to teachers for the training
  - o Fidelity to the guidance
  - o Whether tutors and tutees were trained together and how well this worked

- Materials: (if none, any needed?)
    - Which materials used
    - How easy these are to use
    - Whether targeted at the right age and abilities
    - Suggested improvements to these

- Tutees initial attitudes to the programme

6. Prior to the lesson

Aim: To understand the process of matching pupils together and selecting pupils to take part

- *Matching pupils*
    - *Process of matching pupils*
    - *Use of test results to do this*
    - *Use of teacher knowledge to do this*
    - *Suggested better ways of matching pupils*

Did this work right? Any changes needed? Any suggestions for doing it better?

7. The lessons

Aim: To provide understanding of how the lessons run in practice, how consistent they are and what barriers there are to successful implementation.

- Format of the lesson:
    - What did it replace? How easy to timetable and resource?

    - Impact of absences
    - Use of trios or other set-ups
- Specific aspects of the lesson:
    - Selection of books
    - Selection of passages within books
    - Correction
    - Other specific aspects
    - What worked well
    - What didn't work well

- Differences within the lesson between different subgroups (E.g. high ability, low ability, middle ability, FSM):

    - How this is managed within the lesson
    - Differences in fidelity to the programme- was it stuck to by all or varied? (if varied, how?)

- Overseeing tutoring:

    - How inconsistencies are dealt with
    - How consistency could be improved
    - What type of help/guidance needed – for tutors and tutees
- Support:
    - Nature of support (from SLT, other teachers etc)
    - Nature of support needed

- Barriers to being tutored as perceived by teacher:

  - How these could be tackled

8. Monitoring pupil progress

Aim: To establish how pupil progress is monitered both within the lessons (by tutors) and more broadly by the department

- Tutees progress monitoring:

  - Use of pupil record-keeping of tutees
  - How did tutees find this process

- Teachers monitoring pupil progress:

  - Use of formal assessment
  - Use of informal assessment
  - How involved tutors are in informal assessment

Impact of the programme and the future

9. Impacts

Aim: To provide an understanding of any soft or hard skills gained from the programme, for tutors and tutees. To gather information about the positive and negative impacts on the maths department and the school more broadly

- Impacts on tutees:

  - Reading ability
  - Ability to verbalise thinking
  - Impact of immediate feedback
  - Confidence
  - Social impact
  - Enjoyment of reading

- Impacts on English department and school:

  - Any impacts on English department
  - Staffing
  - Resources
  - Time spent planning
  - Transferring techniques to broader curriculum

10. Views of programme and recommendations

Aim: To gather views on the programme as a whole, whether they would recommend it and how it can be improved.

- Does the programme meet its aims:

- How/ Why not?

- Any additional support they would have liked:

    - At what stage?
    - In what form?

- Any improvements:

    - Materials
    - Support structures
    - Programme itself
    - Anything else

- Recommend use to other schools:

    - Why/ why not?

11. Sustainability

- What would they change if the programme was being rolled out nationally?
- What would be the biggest challenges other schools might face?Will the school continue with peer tutoring after this project? Why / why not?

For Year 9 teachers

Aim of the telephone interview:

The aim of the interview with teachers is to provide a broad overview of the way that tutors for the programme are trained, how the paired reading lessons run in practice, what barriers there are to successful implementation and what broader impacts the programme is having on the students, the department and the school.

The topic guide:

This guide sets out a number of necessary contextual and factual topics and questions that will be covered during interviews. The guide does not contain follow-up probes and questions like `why', `when', and `how', etc. as participants' contributions will be explored in this way, as far as is feasible, during the 45 minute interview. Researchers will use prompts and probes in order to understand how and why views, behaviours and experiences have arisen.

The interview will last for **approximately 45 minutes.**

---

1. **Introductions**

    - Introduce yourself and NatCen Social Research

    - Introduce the study:
        - Independent Evaluation of the Paired Reading Programme
        - Commissioned by the Education Endowment Foundation
    - Overall project aims:
        - To understand school experiences of delivering the programme
        - To understand the impact (both soft and hard outcomes) that the programme is having on students
    - Interview is one of 6 interviews with teachers in several different schools
    - Digital recording – check OK, and reassure re: confidentiality
    - Data kept securely in accordance with Data Protection Act
    - How we'll report findings – anonymity of teachers and schools
    - Any questions/concerns?

    Background and Context

    2. Background

    Aim: To gather background information including detail of the context in which the school operates, the profile of their pupils and the teacher in question

    - Profile of school:
        - Reading and overall literacy levels (particularly year 9)
        - Turnover of pupils at the school
        - Special Education Needs (SEN)

      o   English as an additional language (EAL)

      o   Disadvantaged children (based on FSM take-up)
- Profile of teacher:
  - When qualified
  - Time as teacher
  - Additional responsibilities (e.g. after school clubs etc)

3. Literacy and reading pre the intervention

Aim: To get contextual information about literacy and reading in the school, for years 9, outside of the Peer Tutoring Programme

- Literacy pre Peer Tutoring Programme
  - What interventions in place before / regular format of literacy support
  - What interventions in place now (for pupils not on programme)
  - Nature of these interventions- brief information (e.g. how pupils selected, what they go through, who delivers)
  - Literacy as part of normal school syllabus for year 9

The Peer Tutoring Programme

4. Training for teachers

Aim: To provide understanding of how teachers are trained and any preparation that is done by the teacher in advance of the programme starting

- Training for the teaching staff:

  - Nature of this
  - How effective this was
  - How prepared they felt to implement the programme

- Materials:

  - Which materials used
  - How easy these are to use
  - Suggested improvements to these

- Anything they would have liked to have seen added to the training?

5. Training for pupils

Aim: To provide an understanding of how pupils are trained to be tutors (Yr 9) and any preparation that is done in advance of the programme starting

- What were Yr9 tutors told about the programme prior to starting
- Provided any training of any kind? If so…(if not, was any prep needed?)
- Format of the training sessions for tutors:

  - How well these worked
  - What guidance was provided to teachers for the training

- o Fidelity to the guidance
- o Whether tutors and tutees were trained together and how well this worked
- Materials: (if none, any needed?)
  - o Which materials used
  - o How easy these are to use
  - o Whether targeted at the right age and abilities
  - o Suggested improvements to these

- Tutors initial attitudes to the programme

6. Prior to the lesson

Aim: To understand the process of matching pupils together and selecting pupils to take part

- Matching pupils

  - o Process of matching pupils
  - o Use of test results to do this
  - o Use of teacher knowledge to do this
  - o Suggested better ways of matching pupils

- Did this work right? Any changes needed? Any suggestions for doing it better?

7. The lesson

Aim: To provide understanding of how the lessons run in practice, how consistent they are and what barriers there are to successful implementation.

- Format of the lesson:
  - o What did it replace? How easy to timetable and resource?

  - o Impact of absences
  - o Use of trios or other set-ups
- Specific aspects of the lesson:
  - o Selection of books
  - o Selection of passages within books
  - o Correction
  - o Other specific aspects
  - o What worked well
  - o What didn't work well

- Differences within the lesson between different subgroups (E.g. high ability, low ability, middle ability, FSM):

  - o How this is managed within the lesson
  - o Differences in fidelity to the programme- was it stuck to by all or varied? (if varied, how?)

- Overseeing tutoring:

  - o How inconsistencies are dealt with

- o How consistency could be improved
- o What type of help / guidance needed- for tutors and tutees
- Support:
  - o Nature of support (from SLT, other teachers etc)
  - o Nature of support needed

- Barriers to tutoring as perceived by teacher:

  - o How these could be tackled

8. Monitoring pupil progress

Aim: To establish how pupil progress is monitered both within the lessons (by tutors) and more broadly by the department

- Tutors monitoring pupil progress:

  - o Use of pupil record-keeping from tutors
  - o How easy/difficult tutors find this process

- Teachers monitoring pupil progress:

  - o Use of formal assessment
  - o Use of informal assessment
  - o How involved tutors are in informal assessment

Impact of the programme and the future

9. Impacts

Aim: To provide an understanding of any soft or hard skills gained from the programme, for tutors and tutees. To gather information about the positive and negative impacts on the maths department and the school more broadly

- Impacts on tutors:

  - o Reading ability
  - o Perceptions of own ability
  - o Confidence
  - o Attitudes to school
  - o Social impact
  - o Enjoyment of reading

- Impacts on English department and school:

  - o Any impacts on English department
  - o Staffing
  - o Resources
  - o Time spent planning
  - o Transferring techniques to broader curriculum

10. Views of programme and recommendations

Aim: To gather views on the programme as a whole, whether they would recommend it and how it can be improved.

- Does the programme meet its aims:

    o How/ Why not?

- Any additional support they would have liked:

    o At what stage?
    o In what form?

- Any improvements:

    o Materials
    o Support structures
    o Programme itself
    o Anything else

- Recommend use to other schools:

    o Why/ why not?

11. Sustainability

- What would they change if the programme was being rolled out nationally?
- What would be the biggest challenges other schools might face?
- Will the school continue with peer tutoring after this project?

    o Why?
    o Why not?

Appendix C: Main analysis of the outcomes

Table C1: Analysis of overall reading ability (primary outcome) - Year 7 pupils – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.846	0.018	46.150	0.000	0.810	0.881
FSM eligibility	-0.205	0.649	-0.320	0.752	-1.478	1.067
Gender	1.050	0.491	2.140	0.032	0.088	2.012
Birth month - February	-1.616	1.218	-1.330	0.184	-4.004	0.771
Birth month - March	-0.039	1.185	-0.030	0.974	-2.363	2.284
Birth month - April	1.213	1.211	1.000	0.317	-1.162	3.587
Birth month - May	-0.671	1.202	-0.560	0.577	-3.026	1.684
Birth month - June	0.343	1.170	0.290	0.770	-1.951	2.636
Birth month - July	2.304	1.216	1.900	0.058	-0.078	4.687
Birth month - August	1.810	1.175	1.540	0.123	-0.493	4.114
Birth month - September	-0.510	1.183	-0.430	0.666	-2.829	1.809
Birth month - October	0.273	1.157	0.240	0.814	-1.994	2.540
Birth month - November	0.426	1.210	0.350	0.725	-1.946	2.799
Birth month - December	-0.321	1.236	-0.260	0.795	-2.743	2.101
Allocation	-0.280	0.662	-0.420	0.672	-1.577	1.017
School 2	-0.452	1.433	-0.320	0.752	-3.260	2.355
School 4	1.820	1.273	1.430	0.153	-0.675	4.315
School 5	2.108	1.367	1.540	0.123	-0.570	4.787
School 6	2.867	1.490	1.920	0.054	-0.052	5.786
School 7	-3.412	1.688	-2.020	0.043	-6.719	-0.104
School 8	0.428	1.729	0.250	0.805	-2.961	3.817
School 9	1.014	1.458	0.700	0.487	-1.844	3.873
School 10	2.224	1.209	1.840	0.066	-0.146	4.594
School 3	0.071	1.390	0.050	0.959	-2.652	2.795
Intercept	14.351	2.208	6.500	0.000	10.024	18.678
Number of pupils	1300					
Number of classes	60					

Table C2: Analysis of overall reading ability (primary outcome) - Year 9 pupils – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.903	0.018	49.650	0.000	0.868	0.939
FSM eligibility	-0.686	0.657	-1.040	0.296	-1.975	0.602
Gender	-0.211	0.507	-0.420	0.677	-1.205	0.783
Birth month - February	0.137	1.246	0.110	0.913	-2.306	2.580
Birth month - March	1.387	1.213	1.140	0.253	-0.990	3.764
Birth month - April	1.298	1.260	1.030	0.303	-1.173	3.768
Birth month - May	1.680	1.220	1.380	0.168	-0.711	4.071
Birth month - June	-0.695	1.200	-0.580	0.563	-3.047	1.657
Birth month - July	-0.080	1.181	-0.070	0.946	-2.394	2.235
Birth month - August	1.520	1.208	1.260	0.208	-0.847	3.887
Birth month - September	1.849	1.195	1.550	0.122	-0.492	4.191
Birth month - October	1.155	1.237	0.930	0.351	-1.270	3.580
Birth month - November	2.465	1.223	2.020	0.044	0.068	4.862
Birth month - December	0.870	1.204	0.720	0.470	-1.489	3.229
Allocation	-0.912	0.595	-1.530	0.125	-2.078	0.254
School 2	0.793	1.297	0.610	0.541	-1.748	3.334
School 4	-0.472	1.129	-0.420	0.676	-2.685	1.740
School 5	1.540	1.194	1.290	0.197	-0.800	3.880
School 6	1.429	1.284	1.110	0.266	-1.087	3.946
School 7	-5.913	1.458	-4.060	0.000	-8.771	-3.056
School 8	-4.212	1.502	-2.800	0.005	-7.157	-1.268
School 9	2.922	1.251	2.340	0.020	0.470	5.375
School 10	1.876	1.080	1.740	0.082	-0.240	3.992
School 3	-0.971	1.290	-0.750	0.452	-3.499	1.558
Intercept	9.335	2.178	4.290	0.000	5.066	13.605
Number of pupils	1265					
Number of classes	60					

Table C3: Analysis of passage comprehension (secondary outcome) - Year 7 pupils – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.741	0.020	37.670	0.000	0.703	0.780
FSM eligibility	-2.272	2.820	-0.810	0.421	-7.800	3.256
Gender	9.439	2.138	4.420	0.000	5.249	13.629
Birth month - February	-4.754	5.298	-0.900	0.369	-15.138	5.629
Birth month - March	0.645	5.156	0.130	0.900	-9.460	10.750
Birth month - April	1.817	5.270	0.340	0.730	-8.513	12.147
Birth month - May	-5.304	5.231	-1.010	0.311	-15.558	4.949
Birth month - June	1.199	5.091	0.240	0.814	-8.778	11.176
Birth month - July	7.861	5.288	1.490	0.137	-2.503	18.225
Birth month - August	6.042	5.114	1.180	0.237	-3.981	16.065
Birth month - September	2.155	5.148	0.420	0.675	-7.935	12.246
Birth month - October	2.006	5.032	0.400	0.690	-7.856	11.868
Birth month - November	-0.465	5.268	-0.090	0.930	-10.789	9.859
Birth month - December	0.044	5.375	0.010	0.993	-10.491	10.579
Allocation	-0.447	2.867	-0.160	0.876	-6.066	5.172
School 2	-2.803	6.209	-0.450	0.652	-14.973	9.367
School 4	9.884	5.515	1.790	0.073	-0.926	20.693
School 5	10.762	5.922	1.820	0.069	-0.845	22.369
School 6	15.617	6.457	2.420	0.016	2.961	28.273
School 7	-13.115	7.316	-1.790	0.073	-27.455	1.225
School 8	5.370	7.494	0.720	0.474	-9.317	20.058
School 9	10.198	6.320	1.610	0.107	-2.189	22.586
School 10	11.642	5.238	2.220	0.026	1.376	21.908
School 3	0.367	6.022	0.060	0.951	-11.435	12.169
Intercept	85.717	8.397	10.210	0.000	69.259	102.175
Number of pupils	1300					
Number of classes	60					

Table C4: Analysis of passage comprehension (secondary outcome) - Year 9 pupils – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.789	0.023	34.650	0.000	0.745	0.834
FSM eligibility	-2.561	3.182	-0.800	0.421	-8.797	3.675
Gender	6.708	2.465	2.720	0.007	1.877	11.540
Birth month - February	-0.740	6.025	-0.120	0.902	-12.548	11.068
Birth month - March	5.622	5.865	0.960	0.338	-5.872	17.117
Birth month - April	4.554	6.097	0.750	0.455	-7.396	16.505
Birth month - May	4.535	5.899	0.770	0.442	-7.026	16.096
Birth month - June	-3.222	5.803	-0.560	0.579	-14.594	8.151
Birth month - July	-2.153	5.722	-0.380	0.707	-13.368	9.062
Birth month - August	2.685	5.855	0.460	0.647	-8.791	14.160
Birth month - September	4.883	5.773	0.850	0.398	-6.433	16.199
Birth month - October	3.472	5.981	0.580	0.562	-8.251	15.195
Birth month - November	10.898	5.910	1.840	0.065	-0.686	22.481
Birth month - December	-0.845	5.825	-0.150	0.885	-12.262	10.572
Allocation	-5.218	3.253	-1.600	0.109	-11.593	1.157
School 2	7.787	7.012	1.110	0.267	-5.957	21.531
School 4	-2.701	6.207	-0.440	0.663	-14.866	9.463
School 5	8.926	6.568	1.360	0.174	-3.947	21.799
School 6	6.057	7.041	0.860	0.390	-7.744	19.858
School 7	-37.728	7.887	-4.780	0.000	-53.187	-22.270
School 8	-23.779	8.262	-2.880	0.004	-39.971	-7.587
School 9	17.265	6.897	2.500	0.012	3.746	30.784
School 10	10.762	5.937	1.810	0.070	-0.874	22.398
School 3	-5.826	6.980	-0.830	0.404	-19.506	7.855
Intercept	77.279	10.063	7.680	0.000	57.555	97.003
Number of pupils	1265					
Number of classes	60					

Table C5: Analysis of sentence completion (secondary outcome) - Year 7 pupils – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.741	0.022	33.920	0.000	0.698	0.784
FSM eligibility	-2.953	2.534	-1.170	0.244	-7.920	2.014
Gender	-6.138	1.924	-3.190	0.001	-9.908	-2.368
Birth month - February	-2.747	4.782	-0.570	0.566	-12.119	6.625
Birth month - March	0.414	4.638	0.090	0.929	-8.677	9.505
Birth month - April	6.878	4.750	1.450	0.148	-2.432	16.188
Birth month - May	-3.574	4.718	-0.760	0.449	-12.820	5.672
Birth month - June	5.432	4.594	1.180	0.237	-3.572	14.435
Birth month - July	4.314	4.754	0.910	0.364	-5.004	13.632
Birth month - August	1.116	4.608	0.240	0.809	-7.915	10.147
Birth month - September	-2.997	4.632	-0.650	0.518	-12.076	6.081
Birth month - October	1.618	4.544	0.360	0.722	-7.287	10.524
Birth month - November	6.519	4.739	1.380	0.169	-2.769	15.808
Birth month - December	-0.426	4.851	-0.090	0.930	-9.934	9.081
Allocation	-1.604	2.249	-0.710	0.476	-6.013	2.804
School 2	-2.323	4.955	-0.470	0.639	-12.034	7.389
School 4	0.002	4.333	0.000	1.000	-8.490	8.495
School 5	-1.599	4.654	-0.340	0.731	-10.721	7.523
School 6	2.615	5.112	0.510	0.609	-7.404	12.635
School 7	-12.807	5.875	-2.180	0.029	-24.321	-1.293
School 8	-9.505	5.885	-1.620	0.106	-21.040	2.030
School 9	-9.593	4.985	-1.920	0.054	-19.363	0.177
School 10	-1.819	4.096	-0.440	0.657	-9.847	6.208
School 3	-6.444	4.751	-1.360	0.175	-15.755	2.867
Intercept	104.188	8.554	12.180	0.000	87.422	120.953
Number of pupils	1304					
Number of classes	60					

Table C6: Analysis of sentence completion (secondary outcome) - Year 9 pupils – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.829	0.022	37.480	0.000	0.786	0.872
FSM eligibility	-6.289	2.774	-2.270	0.023	-11.727	-0.852
Gender	-7.256	2.151	-3.370	0.001	-11.472	-3.040
Birth month - February	-1.372	5.305	-0.260	0.796	-11.769	9.025
Birth month - March	-3.686	5.164	-0.710	0.475	-13.808	6.436
Birth month - April	1.704	5.381	0.320	0.752	-8.842	12.249
Birth month - May	2.765	5.199	0.530	0.595	-7.425	12.954
Birth month - June	-4.025	5.102	-0.790	0.430	-14.025	5.975
Birth month - July	-7.809	5.035	-1.550	0.121	-17.678	2.059
Birth month - August	-6.239	5.129	-1.220	0.224	-16.291	3.814
Birth month - September	5.779	5.091	1.140	0.256	-4.200	15.757
Birth month - October	-0.549	5.276	-0.100	0.917	-10.889	9.792
Birth month - November	1.604	5.195	0.310	0.757	-8.577	11.786
Birth month - December	7.360	5.132	1.430	0.151	-2.697	17.418
Allocation	-0.433	2.161	-0.200	0.841	-4.668	3.801
School 2	1.219	4.811	0.250	0.800	-8.211	10.650
School 4	5.432	4.057	1.340	0.181	-2.518	13.383
School 5	2.516	4.296	0.590	0.558	-5.905	10.936
School 6	2.401	4.640	0.520	0.605	-6.692	11.494
School 7	-13.484	5.380	-2.510	0.012	-24.028	-2.940
School 8	-6.543	5.422	-1.210	0.228	-17.170	4.085
School 9	1.717	4.499	0.380	0.703	-7.101	10.534
School 10	4.650	3.883	1.200	0.231	-2.962	12.261
School 3	-6.911	4.739	-1.460	0.145	-16.200	2.378
Intercept	78.904	9.327	8.460	0.000	60.624	97.183
Number of pupils	1275					
Number of classes	60					

Appendix D Subgroup analysis of the outcomes

Table D1a: Analysis of overall reading ability (primary outcome) - Year 7 pupils (Only FSM)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.04 (-0.21 to 0.13)
Regression coefficient (95% CI)	-0.49 (-4.33 to 3.34)	-0.60 (-2.76 to 1.60)
P-value	0.801	0.589
ICC (SE)	0.049 (0.057)	0.008 (0.052)
Variance class level (SE)	9.13 (10.70)	0.50 (3.34)
Variance pupil level (SE)	177.79 (18.25)	63.42 (6.57)
Total sample size (pupils)	246 (58)	246 (58)

Table D1b: Analysis of overall reading ability (primary outcome) - Year 7 pupils, FSM interaction model

	Adjusted analysis
Regression coefficient of the interaction between free school meals and experimental cell (95% CI)	0.068 (-2.42 to 2.56)
P-value	0.957
ICC (SE)	0.037 (0.015)
Variance class level (SE)	2.88 (1.18)
Variance pupil level (SE)	75.32 (3.02)
Variance component on Free School Meals (SE)	0.00 (0.00)
Total sample size (classes)	1,300 (60)

Table D2: Analysis of overall reading ability (primary outcome) - Year 7 pupils (Only FSM) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.816	0.040	20.380	0.000	0.738	0.895
Gender	0.545	1.086	0.500	0.616	-1.584	2.673
Birth month - February	-1.634	2.839	-0.580	0.565	-7.199	3.931
Birth month - March	2.627	2.635	1.000	0.319	-2.537	7.792
Birth month - April	5.763	2.477	2.330	0.020	0.907	10.619
Birth month - May	4.937	2.477	1.990	0.046	0.081	9.793
Birth month - June	5.868	2.690	2.180	0.029	0.596	11.140
Birth month - July	2.381	2.606	0.910	0.361	-2.726	7.488
Birth month - August	5.423	2.577	2.100	0.035	0.372	10.474
Birth month - September	2.023	2.634	0.770	0.443	-3.141	7.186
Birth month - October	2.261	2.440	0.930	0.354	-2.522	7.044
Birth month - November	1.934	2.625	0.740	0.461	-3.211	7.078
Birth month - December	4.489	2.487	1.800	0.071	-0.386	9.364
Allocation	-0.597	1.105	-0.540	0.589	-2.763	1.569
School 2	2.050	2.457	0.830	0.404	-2.765	6.865
School 4	4.451	2.233	1.990	0.046	0.074	8.828
School 5	3.288	2.142	1.530	0.125	-0.911	7.487
School 6	3.794	2.348	1.620	0.106	-0.808	8.396
School 7	-1.388	2.350	-0.590	0.555	-5.995	3.219
School 8	-1.022	2.969	-0.340	0.731	-6.841	4.797
School 9	2.641	1.995	1.320	0.186	-1.269	6.551
School 10	7.961	2.230	3.570	0.000	3.590	12.331
School 3	1.762	2.645	0.670	0.505	-3.422	6.946
Intercept	12.573	4.135	3.040	0.002	4.468	20.678
Number of pupils	246					
Number of classes	58					

Table D3: Analysis of overall reading ability (primary outcome) - Year 7 Interaction model – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.846	0.018	46.150	0.000	0.810	0.881
FSM eligibility	-0.238	0.890	-0.270	0.789	-1.983	1.507
Gender	1.049	0.491	2.140	0.033	0.086	2.011
Birth month - February	-1.615	1.218	-1.330	0.185	-4.003	0.772
Birth month - March	-0.042	1.186	-0.040	0.972	-2.366	2.283
Birth month - April	1.213	1.211	1.000	0.317	-1.162	3.587
Birth month - May	-0.674	1.203	-0.560	0.575	-3.031	1.684
Birth month - June	0.343	1.170	0.290	0.769	-1.950	2.637
Birth month - July	2.304	1.216	1.900	0.058	-0.078	4.687
Birth month - August	1.810	1.175	1.540	0.123	-0.493	4.114
Birth month - September	-0.508	1.184	-0.430	0.668	-2.828	1.813
Birth month - October	0.273	1.157	0.240	0.813	-1.994	2.540
Birth month - November	0.426	1.210	0.350	0.725	-1.947	2.798
Birth month - December	-0.322	1.236	-0.260	0.795	-2.744	2.100
Allocation	-0.293	0.706	-0.420	0.678	-1.677	1.091
School 2	-0.454	1.433	-0.320	0.751	-3.263	2.354
School 4	1.819	1.273	1.430	0.153	-0.676	4.314
School 5	2.106	1.367	1.540	0.124	-0.574	4.785
School 6	2.865	1.490	1.920	0.055	-0.056	5.785
School 7	-3.410	1.688	-2.020	0.043	-6.718	-0.103
School 8	0.425	1.730	0.250	0.806	-2.966	3.815
School 9	1.014	1.458	0.700	0.487	-1.845	3.872
School 10	2.221	1.210	1.840	0.066	-0.151	4.593
School 3	0.072	1.389	0.050	0.959	-2.651	2.795
FSM * Allocation Interaction	0.068	1.270	0.050	0.957	-2.422	2.558
Intercept	14.360	2.216	6.480	0.000	10.018	18.703
Number of pupils	1300					
Number of classes	60					

Table D4a: Analysis of overall reading ability (primary outcome) - Year 9 pupils (Only FSM)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.06 (-0.14 to 0.02)
Regression coefficient (95% CI)	-2.55 (-6.94 to 1.84)	-0.84 (-3.13 to 1.44)
P-value	0.255	0.468
ICC (SE)	0.081 (0.061)	0.045 (0.070)
Variance class level (SE)	18.20 (14.10)	2.90 (4.57)
Variance pupil level (SE)	207.07 (21.33)	61.70 (6.74)
Total sample size (pupils)	246 (60)	246 (60)

Table D4b: Analysis of overall reading ability (primary outcome) - Year 9 pupils, FSM interaction model

	Adjusted analysis
Regression coefficient of the interaction between free school meals and experimental cell (95% CI)	-0.07 (-2.81 to 2.67)
P-value	0.958
ICC (SE)	0.016 (0.012)
Variance class level (SE)	1.28 (0.93)
Variance pupil level (SE)	78.00 (3.24)
Variance component on Free School Meals (SE)	3.82 (5.00)
Total sample size (pupils)	1,265 (60)

Table D5: Analysis of overall reading score (primary outcome) - Year 9 pupils (Only FSM) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.906	0.039	23.520	0.000	0.830	0.981
Gender	-0.047	1.071	-0.040	0.965	-2.147	2.053
Birth month - February	1.350	2.620	0.520	0.606	-3.784	6.485
Birth month - March	3.525	2.245	1.570	0.116	-0.875	7.926
Birth month - April	4.669	2.435	1.920	0.055	-0.103	9.441
Birth month - May	3.556	2.280	1.560	0.119	-0.912	8.025
Birth month - June	3.165	2.270	1.390	0.163	-1.285	7.614
Birth month - July	0.850	2.391	0.360	0.722	-3.836	5.535
Birth month - August	6.034	2.389	2.530	0.012	1.351	10.718
Birth month - September	4.543	2.320	1.960	0.050	-0.004	9.091
Birth month - October	6.555	2.555	2.570	0.010	1.548	11.562
Birth month - November	3.164	2.564	1.230	0.217	-1.861	8.189
Birth month - December	5.267	2.458	2.140	0.032	0.450	10.084
Allocation	-0.844	1.164	-0.720	0.468	-3.125	1.437
School 2	5.436	2.366	2.300	0.022	0.798	10.074
School 4	0.379	2.292	0.170	0.869	-4.113	4.872
School 5	-0.160	2.392	-0.070	0.947	-4.849	4.528
School 6	3.042	2.658	1.140	0.252	-2.168	8.252
School 7	-2.528	2.251	-1.120	0.261	-6.940	1.883
School 8	-6.185	2.937	-2.110	0.035	-11.941	-0.429
School 9	3.143	2.164	1.450	0.146	-1.098	7.384
School 10	3.038	2.397	1.270	0.205	-1.660	7.737
School 3	2.822	2.373	1.190	0.234	-1.829	7.473
Intercept	4.506	4.317	1.040	0.297	-3.956	12.967
Number of pupils	246					
Number of classes	60					

Table D6: Analysis of overall reading ability (primary outcome) - Year 9 Interaction model – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.903	0.018	49.640	0.000	0.867	0.939
FSM eligibility	-0.641	0.999	-0.640	0.521	-2.599	1.316
Gender	-0.224	0.507	-0.440	0.659	-1.217	0.769
Birth month - February	0.134	1.246	0.110	0.914	-2.308	2.576
Birth month - March	1.415	1.212	1.170	0.243	-0.960	3.790
Birth month - April	1.312	1.261	1.040	0.298	-1.159	3.783
Birth month - May	1.661	1.220	1.360	0.173	-0.730	4.052
Birth month - June	-0.704	1.199	-0.590	0.557	-3.054	1.646
Birth month - July	-0.139	1.179	-0.120	0.906	-2.450	2.172
Birth month - August	1.499	1.207	1.240	0.214	-0.867	3.866
Birth month - September	1.844	1.194	1.540	0.122	-0.496	4.184
Birth month - October	1.072	1.237	0.870	0.386	-1.353	3.496
Birth month - November	2.435	1.223	1.990	0.046	0.038	4.833
Birth month - December	0.792	1.205	0.660	0.511	-1.570	3.154
Allocation	-0.891	0.639	-1.390	0.163	-2.142	0.361
School 2	0.580	1.303	0.440	0.656	-1.974	3.134
School 4	-0.499	1.125	-0.440	0.657	-2.704	1.706
School 5	1.560	1.187	1.310	0.189	-0.767	3.887
School 6	1.341	1.279	1.050	0.294	-1.165	3.848
School 7	-6.220	1.494	-4.160	0.000	-9.147	-3.292
School 8	-4.273	1.496	-2.860	0.004	-7.205	-1.340
School 9	2.938	1.261	2.330	0.020	0.467	5.410
School 10	1.833	1.073	1.710	0.088	-0.270	3.936
School 3	-1.166	1.291	-0.900	0.366	-3.696	1.364
FSM * Allocation Interaction	-0.073	1.398	-0.050	0.958	-2.813	2.667
Intercept	9.461	2.175	4.350	0.000	5.198	13.724
Number of pupils	1265					
Number of classes	60					

Table D7a: Analysis of passage comprehension (secondary outcome) - Year 7 pupils (Only FSM)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.02 (-0.20 to 0.16)
Regression coefficient (95% CI)	-0.34 (-14.99 to 14.31)	-1.21 (-10.49 to 8.06)
P-value	0.964	0.798
ICC (SE)	0.028 (0.058)	0.000 (0.00)
Variance class level (SE)	82.03 (171.59)	0.000 (0.000)
Variance pupil level (SE)	2863.09 (299.14)	1208.21 (108.94)
Total sample size (pupils)	247 (58)	246 (58)

Table D7b: Analysis of passage comprehension (secondary outcome) - Year 7 pupils, FSM interaction model

	Adjusted analysis
Regression coefficient of the interaction between free school meals and experimental cell (95% CI)	2.46 (-8.37 to 13.29)
P-value	0.656
ICC (SE)	0.036 (0.015)
Variance class level (SE)	53.48 (22.37)
Variance pupil level (SE)	1425.47 (57.27)
Variance component on Free School Meals (SE)	0.000 (0.000)
Total sample size (pupils)	1,300 (60)

Table D8: Analysis of passage comprehension (secondary outcome) - Year 7 pupils (Only FSM) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.701	0.040	17.390	0.000	0.622	0.780
Gender	5.467	4.725	1.160	0.247	-3.795	14.729
Birth month - February	-4.162	12.375	-0.340	0.737	-28.417	20.092
Birth month - March	18.884	11.410	1.650	0.098	-3.480	41.248
Birth month - April	30.642	10.735	2.850	0.004	9.601	51.683
Birth month - May	22.042	10.764	2.050	0.041	0.945	43.139
Birth month - June	19.172	11.697	1.640	0.101	-3.754	42.099
Birth month - July	16.233	11.282	1.440	0.150	-5.880	38.346
Birth month - August	22.684	11.139	2.040	0.042	0.852	44.517
Birth month - September	14.712	11.486	1.280	0.200	-7.801	37.225
Birth month - October	14.003	10.620	1.320	0.187	-6.811	34.818
Birth month - November	15.950	11.333	1.410	0.159	-6.262	38.163
Birth month - December	21.979	10.851	2.030	0.043	0.710	43.247
Allocation	-1.213	4.731	-0.260	0.798	-10.485	8.059
School 2	2.688	10.545	0.250	0.799	-17.981	23.356
School 4	25.349	9.600	2.640	0.008	6.534	44.164
School 5	17.075	9.185	1.860	0.063	-0.928	35.078
School 6	18.208	10.075	1.810	0.071	-1.538	37.954
School 7	-2.415	10.040	-0.240	0.810	-22.092	17.262
School 8	6.098	12.740	0.480	0.632	-18.872	31.069
School 9	15.669	8.501	1.840	0.065	-0.993	32.331
School 10	39.189	9.572	4.090	0.000	20.429	57.949
School 3	-1.711	11.387	-0.150	0.881	-24.029	20.607
Intercept	73.886	14.813	4.990	0.000	44.854	102.918
Number of pupils	246					
Number of classes	58					

Table D9: Analysis of passage comprehension (secondary outcome) - Year 7 Interaction model – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.741	0.020	37.670	0.000	0.703	0.780
FSM eligibility	-3.448	3.865	-0.890	0.372	-11.023	4.128
Gender	9.403	2.139	4.400	0.000	5.210	13.596
Birth month - February	-4.714	5.298	-0.890	0.374	-15.098	5.670
Birth month - March	0.562	5.159	0.110	0.913	-9.549	10.674
Birth month - April	1.809	5.270	0.340	0.731	-8.520	12.138
Birth month - May	-5.397	5.235	-1.030	0.303	-15.658	4.864
Birth month - June	1.218	5.090	0.240	0.811	-8.759	11.195
Birth month - July	7.862	5.288	1.490	0.137	-2.501	18.226
Birth month - August	6.033	5.113	1.180	0.238	-3.989	16.055
Birth month - September	2.248	5.152	0.440	0.663	-7.851	12.346
Birth month - October	2.017	5.031	0.400	0.689	-7.844	11.878
Birth month - November	-0.493	5.268	-0.090	0.925	-10.818	9.831
Birth month - December	0.006	5.375	0.000	0.999	-10.530	10.542
Allocation	-0.925	3.060	-0.300	0.762	-6.924	5.073
School 2	-2.876	6.209	-0.460	0.643	-15.045	9.293
School 4	9.842	5.513	1.790	0.074	-0.964	20.648
School 5	10.669	5.923	1.800	0.072	-0.939	22.277
School 6	15.529	6.457	2.400	0.016	2.873	28.185
School 7	-13.070	7.314	-1.790	0.074	-27.405	1.264
School 8	5.257	7.494	0.700	0.483	-9.432	19.946
School 9	10.163	6.318	1.610	0.108	-2.219	22.545
School 10	11.530	5.241	2.200	0.028	1.257	21.802
School 3	0.391	6.019	0.060	0.948	-11.405	12.188
FSM * Allocation Interaction	2.458	5.525	0.440	0.656	-8.370	13.287
Intercept	86.034	8.427	10.210	0.000	69.517	102.551
Number of pupils	1300					
Number of classes	60					

Table D10a: Analysis of passage comprehension (secondary outcome) - Year 9 pupils (Only FSM)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.03 (-0.22 to 0.16)
Regression coefficient (95% CI)	-7.72 (-25.50 to 10.07)	-1.88 (-13.61 to 9.86)
P-value	0.395	0.754
ICC (SE)	0.041 (0.056)	0.058 (0.067)
Variance class level (SE)	172.47 (237.86)	95.56 (112.35)
Variance pupil level (SE)	4011.03 (413.01)	1547.54 (165.59)
Total sample size (pupils)	247 (59)	246 (59)

Table D10b: Analysis of passage comprehension (secondary outcome) - Year 9 pupils, FSM interaction model

	Adjusted analysis
Regression coefficient of the interaction between free school meals and experimental cell (95% CI)	4.19 (-9.36 to 17.74)
P-value	0.544
ICC (SE)	0.034 (0.015)
Variance class level (SE)	63.05 (28.03)
Variance pupil level (SE)	1807.60 (75.37)
Variance component on Free School Meals (SE)	63.05 (28.03.51)
Total sample size (pupils)	1,265 (60)

Table D11: Analysis of passage comprehension (secondary outcome) - Year 9 pupils (Only FSM) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.834	0.047	17.830	0.000	0.743	0.926
Gender	10.005	5.401	1.850	0.064	-0.580	20.589
Birth month - February	-0.074	13.164	-0.010	0.995	-25.875	25.726
Birth month - March	18.176	11.284	1.610	0.107	-3.940	40.292
Birth month - April	19.204	12.261	1.570	0.117	-4.827	43.234
Birth month - May	14.954	11.465	1.300	0.192	-7.517	37.426
Birth month - June	5.401	11.410	0.470	0.636	-16.963	27.764
Birth month - July	-7.339	12.021	-0.610	0.542	-30.900	16.222
Birth month - August	24.146	12.084	2.000	0.046	0.462	47.830
Birth month - September	19.874	11.665	1.700	0.088	-2.988	42.737
Birth month - October	27.335	12.832	2.130	0.033	2.185	52.485
Birth month - November	14.431	12.884	1.120	0.263	-10.822	39.683
Birth month - December	30.260	12.348	2.450	0.014	6.059	54.462
Allocation	-1.877	5.986	-0.310	0.754	-13.610	9.856
School 2	24.881	12.178	2.040	0.041	1.013	48.749
School 4	-0.970	11.775	-0.080	0.934	-24.049	22.109
School 5	4.750	12.272	0.390	0.699	-19.302	28.802
School 6	14.158	13.634	1.040	0.299	-12.565	40.880
School 7	-19.448	11.659	-1.670	0.095	-42.300	3.403
School 8	-51.249	15.104	-3.390	0.001	-80.852	-21.647
School 9	22.130	11.212	1.970	0.048	0.154	44.105
School 10	14.470	12.205	1.190	0.236	-9.452	38.392
School 3	8.669	12.182	0.710	0.477	-15.208	32.546
Intercept	39.939	19.368	2.060	0.039	1.978	77.900
Number of pupils	246					
Number of classes	59					

Table D12: Analysis of passage comprehension (secondary outcome) - Year 9 Interaction model – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.789	0.023	34.670	0.000	0.745	0.834
FSM eligibility	-4.802	4.919	-0.980	0.329	-14.443	4.840
Gender	6.631	2.462	2.690	0.007	1.806	11.456
Birth month - February	-0.651	6.021	-0.110	0.914	-12.451	11.149
Birth month - March	5.887	5.857	1.010	0.315	-5.591	17.366
Birth month - April	4.493	6.096	0.740	0.461	-7.454	16.441
Birth month - May	4.305	5.898	0.730	0.465	-7.255	15.864
Birth month - June	-3.412	5.795	-0.590	0.556	-14.771	7.947
Birth month - July	-2.416	5.712	-0.420	0.672	-13.611	8.779
Birth month - August	2.506	5.850	0.430	0.668	-8.960	13.973
Birth month - September	4.692	5.766	0.810	0.416	-6.610	15.993
Birth month - October	2.826	5.979	0.470	0.636	-8.892	14.544
Birth month - November	10.517	5.908	1.780	0.075	-1.062	22.095
Birth month - December	-1.446	5.831	-0.250	0.804	-12.874	9.982
Allocation	-6.050	3.439	-1.760	0.078	-12.790	0.689
School 2	7.405	7.043	1.050	0.293	-6.398	21.208
School 4	-2.832	6.189	-0.460	0.647	-14.961	9.298
School 5	9.115	6.532	1.400	0.163	-3.687	21.918
School 6	5.504	7.011	0.780	0.432	-8.238	19.246
School 7	-39.998	8.075	-4.950	0.000	-55.824	-24.173
School 8	-23.492	8.226	-2.860	0.004	-39.616	-7.368
School 9	17.064	6.944	2.460	0.014	3.454	30.674
School 10	10.712	5.899	1.820	0.069	-0.850	22.275
School 3	-6.427	6.981	-0.920	0.357	-20.110	7.255
FSM * Allocation Interaction	4.191	6.915	0.610	0.544	-9.362	17.743
Intercept	78.186	10.039	7.790	0.000	58.509	97.863
Number of pupils	1265					
Number of classes	60					

Table D13a: Analysis of sentence completion (secondary outcome) - Year 7 pupils (Only FSM)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.15 (-0.39 to 0.09)
Regression coefficient (95% CI)	-7.67 (-21.62 to 6.28)	-7.07 (-18.07 to 3.94)
P-value	0.281	0.208
ICC (SE)	0.068 (0.059)	0.082 (0.055)
Variance class level (SE)	159.90 (141.93)	107.89 (75.14)
Variance pupil level (SE)	2179.63 (223.28)	1213.66 (121.68)
Total sample size (pupils)	248 (58)	248 (58)

Table D13b: Analysis of sentence completion (secondary outcome) - Year 7 pupils, FSM interaction model

	Adjusted analysis
Regression coefficient of the interaction between free school meals and experimental cell (95% CI)	-3.87 (-15.47 to 7.70)
P-value	0.511
ICC (SE)	0.009 (0.011)
Variance class level (SE)	10.12 (13.38)
Variance pupil level (SE)	1152.17 (46.85)
Variance component on Free School Meals (SE)	134.50 (77.71)
Total sample size (pupils)	1,304 (60)

Table D14: Analysis of sentence completion (secondary outcome) - Year 7 pupils (Only FSM) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.687	0.052	13.300	0.000	0.586	0.788
Gender	-1.371	4.852	-0.280	0.778	-10.880	8.139
Birth month - February	-1.595	12.593	-0.130	0.899	-26.277	23.086
Birth month - March	3.994	11.491	0.350	0.728	-18.527	26.516
Birth month - April	3.379	11.121	0.300	0.761	-18.418	25.177
Birth month - May	6.127	11.110	0.550	0.581	-15.649	27.902
Birth month - June	21.652	11.909	1.820	0.069	-1.688	44.992
Birth month - July	-1.060	11.693	-0.090	0.928	-23.977	21.857
Birth month - August	2.365	11.468	0.210	0.837	-20.112	24.843
Birth month - September	-2.335	11.570	-0.200	0.840	-25.012	20.343
Birth month - October	2.350	10.894	0.220	0.829	-19.002	23.703
Birth month - November	3.017	11.741	0.260	0.797	-19.994	26.029
Birth month - December	7.158	11.096	0.650	0.519	-14.589	28.906
Allocation	-7.068	5.614	-1.260	0.208	-18.070	3.935
School 2	9.056	12.252	0.740	0.460	-14.957	33.068
School 4	-3.889	11.139	-0.350	0.727	-25.721	17.943
School 5	4.076	10.852	0.380	0.707	-17.193	25.346
School 6	9.114	11.896	0.770	0.444	-14.203	32.430
School 7	-7.261	12.086	-0.600	0.548	-30.949	16.427
School 8	-13.262	14.924	-0.890	0.374	-42.513	15.989
School 9	-4.401	10.508	-0.420	0.675	-24.996	16.195
School 10	5.808	11.153	0.520	0.603	-16.051	27.667
School 3	-10.094	13.085	-0.770	0.440	-35.740	15.553
Intercept	112.265	18.544	6.050	0.000	75.919	148.611
Number of pupils	248					
Number of classes	58					

Table D15: Analysis of sentence completion (secondary outcome) - Year 7 Interaction model – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.746	0.022	34.460	0.000	0.703	0.788
FSM eligibility	-1.367	4.099	-0.330	0.739	-9.401	6.667
Gender	-6.180	1.922	-3.210	0.001	-9.948	-2.412
Birth month - February	-3.082	4.759	-0.650	0.517	-12.409	6.244
Birth month - March	0.297	4.614	0.060	0.949	-8.747	9.342
Birth month - April	6.319	4.736	1.330	0.182	-2.964	15.602
Birth month - May	-3.873	4.711	-0.820	0.411	-13.107	5.361
Birth month - June	5.013	4.571	1.100	0.273	-3.946	13.971
Birth month - July	4.304	4.742	0.910	0.364	-4.990	13.597
Birth month - August	0.512	4.591	0.110	0.911	-8.485	9.510
Birth month - September	-3.373	4.619	-0.730	0.465	-12.426	5.681
Birth month - October	1.290	4.527	0.290	0.776	-7.582	10.163
Birth month - November	5.577	4.725	1.180	0.238	-3.684	14.838
Birth month - December	-0.403	4.833	-0.080	0.934	-9.875	9.069
Allocation	-1.148	2.286	-0.500	0.616	-5.629	3.333
School 2	-2.897	4.755	-0.610	0.542	-12.216	6.423
School 4	0.595	4.104	0.140	0.885	-7.449	8.639
School 5	-1.862	4.461	-0.420	0.676	-10.606	6.881
School 6	2.478	4.923	0.500	0.615	-7.172	12.127
School 7	-12.984	5.880	-2.210	0.027	-24.508	-1.459
School 8	-8.891	5.607	-1.590	0.113	-19.880	2.098
School 9	-9.872	4.938	-2.000	0.046	-19.550	-0.195
School 10	-1.637	3.868	-0.420	0.672	-9.217	5.943
School 3	-5.783	4.499	-1.290	0.199	-14.600	3.034
FSM * Allocation Interaction	-3.886	5.910	-0.660	0.511	-15.470	7.698
Intercept	102.748	8.460	12.140	0.000	86.166	119.330
Number of pupils	1304					
Number of classes	60					

Table D16a: Analysis of sentence completion (secondary outcome) - Year 9 pupils (Only FSM)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.13 (-0.31 to 0.05)
Regression coefficient (95% CI)	-10.00 (-23.95 to 3.95)	-7.17 (-16.81 to 2.47)
P-value	0.160	0.145
ICC (SE)	0.005 (0.060)	0.000 (0.000)
Variance class level (SE)	14.46 (178.58)	0.00 (0.00)
Variance pupil level (SE)	2951.76 (317.18)	1370.10 (122.79)
Total sample size (pupils)	249 (60)	249 (60)

Table D16b: Analysis of sentence completion (secondary outcome) - Year 9 pupils, FSM interaction model

	Adjusted analysis
Regression coefficient of the interaction between free school meals and experimental cell (95% CI)	-8.03 (-19.09 to 3.02)
P-value	0.154
ICC (SE)	0.000 (0.000)
Variance class level (SE)	0.00 (0.00)
Variance pupil level (SE)	1439.37 (58.58)
Variance component on Free School Meals (SE)	29.56 (70.41)
Total sample size (pupils)	1,275 (60)

Table D17: Analysis of sentence completion (secondary outcome) - Year 9 pupils (Only FSM) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.804	0.052	15.460	0.000	0.702	0.906
Gender	-12.696	4.940	-2.570	0.010	-22.378	-3.015
Birth month - February	-12.547	11.938	-1.050	0.293	-35.945	10.851
Birth month - March	-6.650	10.324	-0.640	0.519	-26.885	13.585
Birth month - April	5.796	11.257	0.510	0.607	-16.268	27.859
Birth month - May	0.790	10.572	0.070	0.940	-19.931	21.511
Birth month - June	9.914	10.595	0.940	0.349	-10.852	30.681
Birth month - July	-1.766	11.112	-0.160	0.874	-23.546	20.013
Birth month - August	2.553	11.082	0.230	0.818	-19.168	24.274
Birth month - September	10.365	10.818	0.960	0.338	-10.838	31.568
Birth month - October	2.594	11.752	0.220	0.825	-20.439	25.628
Birth month - November	16.591	11.723	1.420	0.157	-6.385	39.567
Birth month - December	13.292	11.360	1.170	0.242	-8.972	35.557
Allocation	-7.172	4.918	-1.460	0.145	-16.812	2.467
School 2	12.746	10.053	1.270	0.205	-6.957	32.449
School 4	2.245	9.736	0.230	0.818	-16.838	21.328
School 5	5.132	10.245	0.500	0.616	-14.947	25.211
School 6	3.885	11.390	0.340	0.733	-18.438	26.209
School 7	-1.423	9.192	-0.150	0.877	-19.439	16.593
School 8	1.249	12.576	0.100	0.921	-23.401	25.898
School 9	-0.513	8.919	-0.060	0.954	-17.995	16.969
School 10	5.331	10.202	0.520	0.601	-14.664	25.326
School 3	-0.389	10.201	-0.040	0.970	-20.383	19.606
Intercept	80.468	20.562	3.910	0.000	40.168	120.768
Number of pupils	249					
Number of classes	60					

Table D18: Analysis of sentence completion (secondary outcome) - Year 9 Interaction model – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.829	0.022	37.490	0.000	0.785	0.872
FSM eligibility	-2.162	4.075	-0.530	0.596	-10.150	5.825
Gender	-7.191	2.148	-3.350	0.001	-11.402	-2.981
Birth month - February	-1.384	5.299	-0.260	0.794	-11.769	9.001
Birth month - March	-3.561	5.158	-0.690	0.490	-13.671	6.549
Birth month - April	2.002	5.378	0.370	0.710	-8.538	12.542
Birth month - May	2.869	5.195	0.550	0.581	-7.312	13.051
Birth month - June	-3.820	5.096	-0.750	0.454	-13.808	6.169
Birth month - July	-7.775	5.028	-1.550	0.122	-17.629	2.079
Birth month - August	-6.254	5.125	-1.220	0.222	-16.298	3.790
Birth month - September	5.857	5.085	1.150	0.249	-4.109	15.823
Birth month - October	-0.220	5.273	-0.040	0.967	-10.556	10.115
Birth month - November	1.753	5.193	0.340	0.736	-8.425	11.930
Birth month - December	7.617	5.132	1.480	0.138	-2.442	17.675
Allocation	1.145	2.404	0.480	0.634	-3.566	5.856
School 2	0.446	4.861	0.090	0.927	-9.081	9.973
School 4	5.572	4.079	1.370	0.172	-2.422	13.566
School 5	2.414	4.316	0.560	0.576	-6.044	10.873
School 6	2.530	4.665	0.540	0.588	-6.613	11.673
School 7	-14.158	5.494	-2.580	0.010	-24.926	-3.390
School 8	-6.653	5.452	-1.220	0.222	-17.339	4.034
School 9	1.765	4.561	0.390	0.699	-7.174	10.704
School 10	4.407	3.899	1.130	0.258	-3.236	12.049
School 3	-7.339	4.771	-1.540	0.124	-16.691	2.013
FSM * Allocation Interaction	-8.032	5.639	-1.420	0.154	-19.085	3.021
Intercept	78.257	9.334	8.380	0.000	59.962	96.552
Number of pupils	1275					
Number of classes	60					

Appendix E: Subgroup analysis of the outcomes (gender)

Table E1a: Analysis of overall reading ability (primary outcome) - Year 7 pupils (Only boys)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.05 (-0.17 to 0.07)
Regression coefficient (95% CI)	-0.22 (-4.37 to 3.94)	-0.77 (-2.50 to 0.97)
P-value	0.919	0.388
ICC (SE)	0.183 (0.042)	0.035 (0.023)
Variance class level (SE)	46.01 (12.23)	3.09 (2.08)
Variance pupil level (SE)	204.94 (12.06)	86.26 (5.07)
Total sample size (pupils)	638 (60)	635 (60)

Table E1b: Analysis of overall reading ability (primary outcome) - Year 7 pupils (Only girls)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	0.02 (-0.08 to 0.12)
Regression coefficient (95% CI)	-1.38 (-4.62 to 1.87)	0.21 (-1.07 to 1.48)
P-value	0.405	0.751
ICC (SE)	0.118 (0.034)	0.005 (0.018)
Variance class level (SE)	23.44 (7.45)	0.29 (1.17)
Variance pupil level (SE)	174.91 (10.04)	64.50 (3.71)
Total sample size (pupils)	667 (60)	665 (60)

Table E1c: Analysis of overall reading ability (primary outcome) - Year 7 pupils, Gender interaction model

	Adjusted analysis
Regression coefficient of the interaction between free school meals and experimental cell (95% CI)	1.12 (-0.80 to 3.04)
P-value	0.253
ICC (SE)	0.036 (0.015)
Variance class level (SE)	2.84 (1.17)
Variance pupil level (SE)	75.26 (3.02)
Variance component on Gender (SE)	0.00 (0.00)
Total sample size (classes)	1,300 (60)

Table E2: Analysis of overall reading ability (primary outcome) - Year 7 pupils (Only boys) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.847	0.027	31.860	0.000	0.795	0.899
Free School Meals (FSM)	0.793	1.020	0.780	0.437	-1.206	2.793
Birth month - February	-0.327	1.868	-0.180	0.861	-3.989	3.334
Birth month - March	-1.180	1.880	-0.630	0.530	-4.865	2.506
Birth month - April	1.699	1.792	0.950	0.343	-1.813	5.212
Birth month - May	0.134	1.859	0.070	0.942	-3.509	3.777
Birth month - June	0.868	1.839	0.470	0.637	-2.736	4.471
Birth month - July	3.172	1.804	1.760	0.079	-0.363	6.707
Birth month - August	2.040	1.830	1.110	0.265	-1.546	5.625
Birth month - September	-1.173	1.802	-0.650	0.515	-4.704	2.358
Birth month - October	-0.590	1.777	-0.330	0.740	-4.073	2.892
Birth month - November	0.117	1.806	0.060	0.948	-3.423	3.657
Birth month - December	-0.192	1.939	-0.100	0.921	-3.992	3.608
Allocation	-0.766	0.887	-0.860	0.388	-2.504	0.972
School 2	-3.122	1.974	-1.580	0.114	-6.990	0.746
School 4	-0.324	1.730	-0.190	0.851	-3.714	3.066
School 5	0.566	1.844	0.310	0.759	-3.048	4.179
School 6	1.227	1.971	0.620	0.534	-2.637	5.091
School 7	-5.105	2.198	-2.320	0.020	-9.413	-0.797
School 8	-1.271	2.431	-0.520	0.601	-6.036	3.493
School 9	-1.184	2.060	-0.570	0.565	-5.220	2.853
School 10	1.266	1.591	0.800	0.426	-1.853	4.385
School 3	-0.627	1.780	-0.350	0.724	-4.117	2.862
Intercept	15.502	3.165	4.900	0.000	9.298	21.705
Number of pupils	635					
Number of classes	60					

Table E3: Analysis of overall reading ability (primary outcome) - Year 7 pupils (Only girls) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.860	0.024	35.480	0.000	0.812	0.907
Gender	-1.160	0.832	-1.390	0.163	-2.791	0.471
Birth month - February	-2.881	1.578	-1.830	0.068	-5.974	0.212
Birth month - March	1.574	1.496	1.050	0.293	-1.358	4.505
Birth month - April	0.225	1.628	0.140	0.890	-2.967	3.417
Birth month - May	-0.980	1.541	-0.640	0.525	-4.001	2.041
Birth month - June	-0.081	1.490	-0.050	0.957	-3.001	2.839
Birth month - July	1.043	1.625	0.640	0.521	-2.143	4.228
Birth month - August	1.701	1.494	1.140	0.255	-1.227	4.629
Birth month - September	0.470	1.539	0.310	0.760	-2.546	3.487
Birth month - October	1.460	1.498	0.970	0.330	-1.476	4.396
Birth month - November	0.753	1.609	0.470	0.640	-2.400	3.905
Birth month - December	-0.420	1.571	-0.270	0.789	-3.498	2.659
Allocation	0.207	0.652	0.320	0.751	-1.071	1.484
School 2	2.328	1.441	1.610	0.106	-0.497	5.152
School 4	3.929	1.233	3.190	0.001	1.512	6.347
School 5	3.746	1.332	2.810	0.005	1.135	6.356
School 6	4.791	1.533	3.130	0.002	1.787	7.794
School 7	-1.729	1.914	-0.900	0.366	-5.479	2.022
School 8	2.521	1.640	1.540	0.124	-0.694	5.735
School 9	3.266	1.384	2.360	0.018	0.554	5.979
School 10	3.112	1.189	2.620	0.009	0.781	5.443
School 3	0.900	1.474	0.610	0.542	-1.990	3.790
Intercept	12.439	2.803	4.440	0.000	6.945	17.933
Number of pupils	665					
Number of classes	60					

Table E4: Analysis of overall reading ability (primary outcome) - Year 7 Gender interaction model – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	-0.234	0.650	-0.360	0.719	-1.507	1.039
FSM eligibility	0.514	0.679	0.760	0.449	-0.817	1.845
Gender	-1.634	1.218	-1.340	0.179	-4.021	0.752
Birth month - February	-0.043	1.185	-0.040	0.971	-2.365	2.280
Birth month - March	1.196	1.211	0.990	0.324	-1.178	3.569
Birth month - April	-0.625	1.202	-0.520	0.603	-2.981	1.730
Birth month - May	0.347	1.170	0.300	0.767	-1.946	2.640
Birth month - June	2.347	1.216	1.930	0.054	-0.036	4.729
Birth month - July	1.788	1.175	1.520	0.128	-0.515	4.090
Birth month - August	-0.542	1.183	-0.460	0.647	-2.861	1.777
Birth month - September	0.267	1.156	0.230	0.817	-1.999	2.534
Birth month - October	0.447	1.210	0.370	0.712	-1.925	2.818
Birth month - November	-0.350	1.235	-0.280	0.777	-2.771	2.071
Birth month - December	-0.849	0.826	-1.030	0.304	-2.468	0.770
Allocation	-0.461	1.428	-0.320	0.747	-3.260	2.337
School 2	1.795	1.269	1.420	0.157	-0.691	4.282
School 4	2.130	1.362	1.560	0.118	-0.539	4.799
School 5	2.858	1.484	1.930	0.054	-0.052	5.767
School 6	-3.387	1.683	-2.010	0.044	-6.685	-0.089
School 7	0.454	1.723	0.260	0.792	-2.923	3.832
School 8	1.034	1.453	0.710	0.477	-1.814	3.883
School 9	2.237	1.205	1.860	0.063	-0.124	4.598
School 10	0.076	1.385	0.050	0.956	-2.638	2.790
School 3	1.119	0.980	1.140	0.253	-0.801	3.039
Gender * Allocation	14.524	2.211	6.570	0.000	10.190	18.857
Intercept	-0.234	0.650	-0.360	0.719	-1.507	1.039
Number of pupils	1300					
Number of classes	60					

Table E5a: Analysis of overall reading ability (primary outcome) - Year 9 pupils (Only boys)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.01 (-0.11 to 0.09)
Regression coefficient (95% CI)	1.26 (-1.67 to 4.19)	-0.16 (-1.69 to 1.37)
P-value	0.398	0.836
ICC (SE)	0.039 (0.023)	0.006 (0.017)
Variance class level (SE)	10.07 (5.94)	0.50 (1.57)
Variance pupil level (SE)	248.66 (14.28)	90.12 (5.17)
Total sample size (pupils)	662 (60)	662 (60)

Table E5b: Analysis of overall reading ability (primary outcome) - Year 9 pupils (Only girls)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.12 (-0.22 to -0.02)
Regression coefficient (95% CI)	-4.68 (-7.55 to -1.80)	-1.67 (-3.02 to -0.32)
P-value	0.001	0.016
ICC (SE)	0.047 (0.025)	0.004 (0.020)
Variance class level (SE)	10.28 (5.53)	0.27 (1.31)
Variance pupil level (SE)	208.67 (12.56)	65.36 (3.97)
Total sample size (pupils)	604 (60)	603 (60)

Table E5c: Analysis of overall reading ability (primary outcome) - Year 9 pupils, Gender interaction model

	Adjusted analysis
Regression coefficient of the interaction between free school meals and experimental cell (95% CI)	-1.80 (-3.79 to 0.18)
P-value	0.075
ICC (SE)	0.018 (0.012)
Variance class level (SE)	1.44 (0.93)
Variance pupil level (SE)	78.36 (3.19)
Variance component on Gender (SE)	0.00 (0.00)
Total sample size (classes)	1,265 (60)

Table E6: Analysis of overall reading ability (primary outcome) - Year 9 pupils (Only boys) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.895	0.026	34.360	0.000	0.844	0.946
Free School Meals (FSM)	-0.888	1.006	-0.880	0.377	-2.858	1.083
Birth month - February	-0.022	1.866	-0.010	0.991	-3.678	3.635
Birth month - March	0.785	1.841	0.430	0.670	-2.823	4.394
Birth month - April	0.073	1.810	0.040	0.968	-3.474	3.620
Birth month - May	1.102	1.785	0.620	0.537	-2.396	4.600
Birth month - June	0.888	1.834	0.480	0.628	-2.706	4.483
Birth month - July	-0.361	1.796	-0.200	0.841	-3.882	3.160
Birth month - August	0.671	1.812	0.370	0.711	-2.880	4.223
Birth month - September	2.039	1.818	1.120	0.262	-1.524	5.601
Birth month - October	0.394	1.787	0.220	0.825	-3.108	3.897
Birth month - November	2.945	1.840	1.600	0.109	-0.661	6.551
Birth month - December	0.278	1.767	0.160	0.875	-3.186	3.741
Allocation	-0.161	0.781	-0.210	0.836	-1.692	1.369
School 2	1.552	1.709	0.910	0.364	-1.798	4.903
School 4	-0.541	1.476	-0.370	0.714	-3.433	2.352
School 5	2.782	1.659	1.680	0.094	-0.469	6.032
School 6	2.219	1.702	1.300	0.192	-1.116	5.554
School 7	-6.422	1.937	-3.320	0.001	-10.219	-2.625
School 8	-4.592	1.919	-2.390	0.017	-8.352	-0.832
School 9	4.136	1.611	2.570	0.010	0.979	7.292
School 10	2.465	1.426	1.730	0.084	-0.330	5.260
School 3	0.939	1.702	0.550	0.581	-2.397	4.276
Intercept	9.492	3.026	3.140	0.002	3.562	15.423
Number of pupils	662					
Number of classes	60					

Table E7: Analysis of overall reading ability (primary outcome) - Year 9 pupils (Only girls) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.915	0.025	36.580	0.000	0.866	0.964
Gender	-0.495	0.853	-0.580	0.561	-2.166	1.176
Birth month - February	0.284	1.623	0.180	0.861	-2.896	3.464
Birth month - March	1.957	1.566	1.250	0.211	-1.112	5.025
Birth month - April	3.504	1.732	2.020	0.043	0.109	6.899
Birth month - May	2.253	1.636	1.380	0.168	-0.953	5.459
Birth month - June	-1.871	1.546	-1.210	0.226	-4.901	1.158
Birth month - July	0.229	1.533	0.150	0.881	-2.775	3.233
Birth month - August	2.461	1.577	1.560	0.119	-0.631	5.552
Birth month - September	1.907	1.544	1.240	0.217	-1.119	4.932
Birth month - October	2.134	1.690	1.260	0.207	-1.178	5.447
Birth month - November	2.138	1.592	1.340	0.179	-0.981	5.257
Birth month - December	1.862	1.610	1.160	0.248	-1.294	5.018
Allocation	-1.668	0.690	-2.420	0.016	-3.021	-0.316
School 2	0.106	1.548	0.070	0.946	-2.928	3.140
School 4	-0.381	1.286	-0.300	0.767	-2.902	2.140
School 5	0.535	1.281	0.420	0.676	-1.976	3.046
School 6	0.628	1.444	0.440	0.663	-2.201	3.458
School 7	-5.417	1.699	-3.190	0.001	-8.748	-2.087
School 8	-3.493	1.750	-2.000	0.046	-6.924	-0.062
School 9	1.342	1.446	0.930	0.354	-1.493	4.176
School 10	1.280	1.215	1.050	0.292	-1.100	3.661
School 3	-3.383	1.513	-2.240	0.025	-6.348	-0.417
Intercept	8.485	3.005	2.820	0.005	2.595	14.374
Number of pupils	603					
Number of classes	60					

Table E8: Analysis of overall reading ability (primary outcome) - Year 9 Gender interaction model – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.901	0.018	49.380	0.000	0.865	0.936
FSM eligibility	-0.687	0.657	-1.050	0.295	-1.974	0.600
Gender	0.671	0.709	0.950	0.344	-0.719	2.061
Birth month - February	0.142	1.245	0.110	0.909	-2.298	2.582
Birth month - March	1.372	1.211	1.130	0.257	-1.002	3.746
Birth month - April	1.333	1.259	1.060	0.290	-1.135	3.800
Birth month - May	1.612	1.219	1.320	0.186	-0.777	4.001
Birth month - June	-0.754	1.199	-0.630	0.529	-3.104	1.596
Birth month - July	-0.141	1.180	-0.120	0.905	-2.453	2.172
Birth month - August	1.379	1.209	1.140	0.254	-0.990	3.748
Birth month - September	1.850	1.193	1.550	0.121	-0.488	4.189
Birth month - October	1.102	1.236	0.890	0.373	-1.321	3.524
Birth month - November	2.432	1.222	1.990	0.046	0.038	4.827
Birth month - December	0.923	1.203	0.770	0.443	-1.434	3.280
Allocation	-0.054	0.767	-0.070	0.944	-1.558	1.449
School 2	0.847	1.299	0.650	0.514	-1.699	3.394
School 4	-0.521	1.132	-0.460	0.645	-2.739	1.696
School 5	1.523	1.197	1.270	0.203	-0.822	3.869
School 6	1.487	1.287	1.160	0.248	-1.036	4.009
School 7	-5.904	1.460	-4.040	0.000	-8.766	-3.041
School 8	-4.239	1.506	-2.820	0.005	-7.190	-1.288
School 9	2.942	1.254	2.350	0.019	0.484	5.401
School 10	1.860	1.082	1.720	0.086	-0.260	3.981
School 3	-1.022	1.293	-0.790	0.429	-3.555	1.512
Gender * Allocation	-1.802	1.014	-1.780	0.075	-3.788	0.185
Intercept	9.225	2.178	4.240	0.000	4.956	13.493
Number of pupils	1265					
Number of classes	60					

Table E9a: Analysis of passage comprehension (secondary outcome) - Year 7 pupils (Only boys)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.03 (-0.16 to 0.10)
Regression coefficient (95% CI)	0.08 (-15.62 to 15.77)	-1.75 (-8.97 to 5.46)
P-value	0.992	0.634
ICC (SE)	0.170 (0.040)	0.023 (0.021)
Variance class level (SE)	637.50 (174.86)	38.06 (36.44)
Variance pupil level (SE)	3116.69 (183.27)	1650.09 (97.16)
Total sample size (pupils)	639 (60)	635 (60)

Table E9b: Analysis of passage comprehension (secondary outcome) - Year 7 pupils (Only girls)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	0.02 (-0.1 to 0.14)
Regression coefficient (95% CI)	-3.90 (-16.15 to 8.34)	0.89 (-4.97 to 6.76)
P-value	0.532	0.766
ICC (SE)	0.109 (0.033)	0.018 (0.020)
Variance class level (SE)	321.75 (104.90)	21.33 (23.60)
Variance pupil level (SE)	2637.41 (151.18)	1180.93 (67.67)
Total sample size (pupils)	667 (60)	665 (60)

Table E9c: Analysis of passage comprehension (secondary outcome) - Year 7 pupils, Gender interaction model

	Adjusted analysis
Regression coefficient of the interaction between free school meals and experimental cell (95% CI)	3.11 (-5.25 to 11.46)
P-value	0.466
ICC (SE)	0.036 (0.015)
Variance class level (SE)	53.14 (22.32)
Variance pupil level (SE)	1424.28 (57.26)
Variance component on Gender (SE)	0.00 (0.00)
Total sample size (classes)	1,300 (60)

Table E10: Analysis of passage comprehension (secondary outcome) - Year 7 pupils (Only boys) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.738	0.028	26.000	0.000	0.683	0.794
Free School Meals (FSM)	2.912	4.439	0.660	0.512	-5.788	11.613
Birth month - February	0.134	8.143	0.020	0.987	-15.826	16.094
Birth month - March	-1.135	8.184	-0.140	0.890	-17.176	14.905
Birth month - April	4.193	7.812	0.540	0.591	-11.118	19.504
Birth month - May	0.721	8.119	0.090	0.929	-15.192	16.634
Birth month - June	5.418	8.018	0.680	0.499	-10.296	21.133
Birth month - July	12.985	7.856	1.650	0.098	-2.413	28.383
Birth month - August	9.485	7.987	1.190	0.235	-6.170	25.140
Birth month - September	-1.344	7.862	-0.170	0.864	-16.753	14.064
Birth month - October	-3.247	7.747	-0.420	0.675	-18.431	11.937
Birth month - November	-1.757	7.881	-0.220	0.824	-17.203	13.690
Birth month - December	1.027	8.452	0.120	0.903	-15.540	17.593
Allocation	-1.753	3.682	-0.480	0.634	-8.969	5.463
School 2	-19.315	8.256	-2.340	0.019	-35.496	-3.134
School 4	-0.968	7.194	-0.130	0.893	-15.068	13.133
School 5	1.183	7.663	0.150	0.877	-13.836	16.202
School 6	6.662	8.214	0.810	0.417	-9.438	22.762
School 7	-22.216	9.169	-2.420	0.015	-40.186	-4.246
School 8	-7.428	10.163	-0.730	0.465	-27.347	12.491
School 9	-3.659	8.601	-0.430	0.671	-20.516	13.198
School 10	6.483	6.589	0.980	0.325	-6.430	19.397
School 3	-4.539	7.376	-0.620	0.538	-18.996	9.918
Intercept	93.210	11.884	7.840	0.000	69.917	116.503
Number of pupils	635					
Number of classes	60					

Table E11: Analysis of passage comprehension (secondary outcome) - Year 7 pupils (Only girls) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.773	0.026	29.410	0.000	0.721	0.824
Gender	-7.078	3.567	-1.980	0.047	-14.069	-0.087
Birth month - February	-9.983	6.776	-1.470	0.141	-23.263	3.298
Birth month - March	4.920	6.431	0.770	0.444	-7.684	17.524
Birth month - April	-3.278	6.995	-0.470	0.639	-16.988	10.433
Birth month - May	-10.135	6.618	-1.530	0.126	-23.107	2.837
Birth month - June	-3.451	6.403	-0.540	0.590	-16.000	9.098
Birth month - July	0.786	6.989	0.110	0.910	-12.913	14.484
Birth month - August	2.333	6.423	0.360	0.716	-10.256	14.922
Birth month - September	6.517	6.612	0.990	0.324	-6.443	19.476
Birth month - October	8.033	6.437	1.250	0.212	-4.584	20.650
Birth month - November	0.058	6.917	0.010	0.993	-13.498	13.614
Birth month - December	-2.751	6.753	-0.410	0.684	-15.986	10.484
Allocation	0.892	2.992	0.300	0.766	-4.973	6.757
School 2	12.847	6.537	1.970	0.049	0.035	25.660
School 4	20.469	5.653	3.620	0.000	9.389	31.549
School 5	20.160	6.100	3.310	0.001	8.205	32.116
School 6	25.338	6.981	3.630	0.000	11.654	39.021
School 7	-5.845	8.586	-0.680	0.496	-22.672	10.983
School 8	18.322	7.524	2.440	0.015	3.575	33.068
School 9	22.993	6.355	3.620	0.000	10.538	35.448
School 10	15.886	5.454	2.910	0.004	5.196	26.576
School 3	4.833	6.726	0.720	0.472	-8.350	18.016
Intercept	78.536	10.637	7.380	0.000	57.689	99.384
Number of pupils	665					
Number of classes	60					

Table E12: Analysis of overall passage comprehension (secondary outcome) - Year 7 Gender interaction model – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.742	0.020	37.680	0.000	0.703	0.781
FSM eligibility	-2.351	2.822	-0.830	0.405	-7.882	3.180
Gender	7.948	2.958	2.690	0.007	2.151	13.745
Birth month - February	-4.805	5.297	-0.910	0.364	-15.187	5.578
Birth month - March	0.635	5.155	0.120	0.902	-9.468	10.739
Birth month - April	1.771	5.270	0.340	0.737	-8.558	12.099
Birth month - May	-5.177	5.233	-0.990	0.323	-15.434	5.081
Birth month - June	1.213	5.090	0.240	0.812	-8.762	11.189
Birth month - July	7.985	5.290	1.510	0.131	-2.383	18.353
Birth month - August	5.984	5.113	1.170	0.242	-4.038	16.007
Birth month - September	2.062	5.149	0.400	0.689	-8.030	12.154
Birth month - October	1.990	5.031	0.400	0.692	-7.871	11.850
Birth month - November	-0.414	5.267	-0.080	0.937	-10.738	9.909
Birth month - December	-0.039	5.376	-0.010	0.994	-10.575	10.497
Allocation	-2.028	3.590	-0.570	0.572	-9.064	5.007
School 2	-2.828	6.198	-0.460	0.648	-14.977	9.321
School 4	9.815	5.505	1.780	0.075	-0.975	20.606
School 5	10.821	5.911	1.830	0.067	-0.764	22.406
School 6	15.590	6.445	2.420	0.016	2.957	28.223
School 7	-13.048	7.305	-1.790	0.074	-27.365	1.269
School 8	5.445	7.480	0.730	0.467	-9.215	20.106
School 9	10.254	6.308	1.630	0.104	-2.110	22.618
School 10	11.678	5.227	2.230	0.025	1.432	21.923
School 3	0.379	6.010	0.060	0.950	-11.400	12.159
Gender * Allocation	3.109	4.263	0.730	0.466	-5.246	11.465
Intercept	86.236	8.425	10.240	0.000	69.723	102.750
Number of pupils	1304					
Number of classes	60					

Table E13a: Analysis of passage comprehension (secondary outcome) - Year 9 pupils (Only boys)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	0.01 (-0.12 to 0.14)
Regression coefficient (95% CI)	4.35 (-7.78 to 16.49)	0.30 (-7.99 to 8.59)
P-value	0.482	0.843
ICC (SE)	0.043 (0.023)	0.026 (0.021)
Variance class level (SE)	182.80 (101.04)	58.51 (46.31)
Variance pupil level (SE)	4176.23 (239.17)	2163.68 (124.25)
Total sample size (pupils)	665 (60)	662 (60)

Table E13b: Analysis of passage comprehension (secondary outcome) - Year 9 pupils (Only girls)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.2 (-0.214 to -0.186)
Regression coefficient (95% CI)	-19.25 (-30.05 to -8.45)	-10.89 (-17.48 to -4.30)
P-value	0.000	0.001
ICC (SE)	0.045 (0.025)	0.017 (0.022)
Variance class level (SE)	140.55 (79.98)	23.21 (30.95)
Variance pupil level (SE)	2994.40 (180.68)	1380.52 (83.79)
Total sample size (pupils)	604 (60)	603 (60)

Table E13c: Analysis of passage comprehension (secondary outcome) - Year 9 pupils, Gender interaction model

	Adjusted analysis
Regression coefficient of the interaction between free school meals and experimental cell (95% CI)	-11.94 (-21.52 to -2.35)
P-value	0.015
ICC (SE)	0.036 (0.015)
Variance class level (SE)	67.98 (28.04)
Variance pupil level (SE)	1816.03 (73.90)
Variance component on Gender (SE)	0.00 (0.00)
Total sample size (classes)	1,265 (60)

Table E14: Analysis of passage comprehension (secondary outcome) - Year 9 pupils (Only boys) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.772	0.032	23.910	0.000	0.709	0.835
Free School Meals (FSM)	-5.147	4.953	-1.040	0.299	-14.854	4.560
Birth month - February	-3.835	9.220	-0.420	0.677	-21.905	14.235
Birth month - March	4.057	9.092	0.450	0.655	-13.762	21.877
Birth month - April	0.173	8.931	0.020	0.985	-17.332	17.678
Birth month - May	3.744	8.827	0.420	0.671	-13.557	21.044
Birth month - June	9.212	9.050	1.020	0.309	-8.526	26.950
Birth month - July	-4.294	8.894	-0.480	0.629	-21.725	13.137
Birth month - August	-3.583	8.955	-0.400	0.689	-21.134	13.968
Birth month - September	8.469	8.989	0.940	0.346	-9.149	26.086
Birth month - October	-1.942	8.842	-0.220	0.826	-19.272	15.389
Birth month - November	14.498	9.080	1.600	0.110	-3.298	32.294
Birth month - December	-5.072	8.736	-0.580	0.562	-22.195	12.050
Allocation	0.303	4.232	0.070	0.943	-7.992	8.597
School 2	13.860	9.189	1.510	0.131	-4.150	31.870
School 4	-4.853	8.045	-0.600	0.546	-20.620	10.915
School 5	16.304	8.933	1.830	0.068	-1.205	33.812
School 6	11.994	9.224	1.300	0.193	-6.084	30.073
School 7	-44.893	10.380	-4.320	0.000	-65.238	-24.549
School 8	-25.908	10.491	-2.470	0.014	-46.469	-5.347
School 9	22.383	8.824	2.540	0.011	5.089	39.677
School 10	15.809	7.753	2.040	0.041	0.613	31.005
School 3	5.233	9.137	0.570	0.567	-12.674	23.141
Intercept	78.725	13.884	5.670	0.000	51.514	105.936
Number of pupils	662					
Number of classes	60					

Table E15: Analysis of passage comprehension (secondary outcome) - Year 9 pupils (Only girls) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.823	0.031	26.210	0.000	0.762	0.885
Gender	0.630	3.948	0.160	0.873	-7.108	8.368
Birth month - February	3.163	7.483	0.420	0.673	-11.504	17.829
Birth month - March	8.456	7.224	1.170	0.242	-5.702	22.615
Birth month - April	14.956	8.005	1.870	0.062	-0.734	30.646
Birth month - May	5.640	7.549	0.750	0.455	-9.156	20.436
Birth month - June	-11.579	7.136	-1.620	0.105	-25.565	2.407
Birth month - July	0.625	7.084	0.090	0.930	-13.259	14.510
Birth month - August	10.175	7.304	1.390	0.164	-4.141	24.490
Birth month - September	3.060	7.106	0.430	0.667	-10.869	16.988
Birth month - October	11.669	7.800	1.500	0.135	-3.619	26.958
Birth month - November	9.216	7.337	1.260	0.209	-5.165	23.598
Birth month - December	7.467	7.438	1.000	0.315	-7.112	22.046
Allocation	-10.894	3.362	-3.240	0.001	-17.485	-4.304
School 2	1.836	7.476	0.250	0.806	-12.816	16.488
School 4	0.544	6.307	0.090	0.931	-11.818	12.906
School 5	3.218	6.320	0.510	0.611	-9.170	15.606
School 6	0.535	7.116	0.080	0.940	-13.411	14.481
School 7	-29.883	8.220	-3.640	0.000	-45.994	-13.771
School 8	-20.006	8.536	-2.340	0.019	-36.737	-3.276
School 9	10.511	7.080	1.480	0.138	-3.365	24.387
School 10	5.882	5.955	0.990	0.323	-5.789	17.554
School 3	-20.027	7.322	-2.740	0.006	-34.378	-5.676
Intercept	74.930	13.548	5.530	0.000	48.378	101.483
Number of pupils	603					
Number of classes	60					

Table E16: Analysis of passage comprehension (secondary outcome) - Year 9 Gender interaction model – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.785	0.023	34.420	0.000	0.740	0.830
FSM eligibility	-2.560	3.174	-0.810	0.420	-8.781	3.661
Gender	12.578	3.440	3.660	0.000	5.836	19.320
Birth month - February	-0.708	6.010	-0.120	0.906	-12.487	11.071
Birth month - March	5.520	5.850	0.940	0.345	-5.946	16.987
Birth month - April	4.774	6.083	0.780	0.433	-7.149	16.697
Birth month - May	4.075	5.887	0.690	0.489	-7.463	15.614
Birth month - June	-3.630	5.791	-0.630	0.531	-14.980	7.719
Birth month - July	-2.588	5.711	-0.450	0.650	-13.782	8.605
Birth month - August	1.710	5.854	0.290	0.770	-9.764	13.184
Birth month - September	4.922	5.759	0.850	0.393	-6.366	16.209
Birth month - October	3.140	5.968	0.530	0.599	-8.558	14.837
Birth month - November	10.690	5.896	1.810	0.070	-0.866	22.247
Birth month - December	-0.475	5.813	-0.080	0.935	-11.868	10.918
Allocation	0.462	4.005	0.120	0.908	-7.387	8.311
School 2	8.165	7.025	1.160	0.245	-5.605	21.934
School 4	-3.008	6.221	-0.480	0.629	-15.202	9.185
School 5	8.819	6.583	1.340	0.180	-4.083	21.720
School 6	6.478	7.058	0.920	0.359	-7.356	20.312
School 7	-37.649	7.900	-4.770	0.000	-53.134	-22.165
School 8	-23.950	8.280	-2.890	0.004	-40.178	-7.721
School 9	17.411	6.913	2.520	0.012	3.861	30.960
School 10	10.665	5.950	1.790	0.073	-0.997	22.327
School 3	-6.176	6.993	-0.880	0.377	-19.883	7.530
Gender * Allocation	-11.936	4.890	-2.440	0.015	-21.520	-2.352
Intercept	76.291	10.058	7.590	0.000	56.578	96.004
Number of pupils	1265					
Number of classes	60					

Table E17a: Analysis of sentence completion (secondary outcome) - Year 7 pupils (Only boys)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.03 (-0.18 to 0.12)
Regression coefficient (95% CI)	-3.52 (-16.55 to 9.51)	-1.56 (-8.87 to 5.75)
P-value	0.596	0.676
ICC (SE)	0.139 (0.039)	0.045 (0.026)
Variance class level (SE)	405.14 (124.87)	65.89 (39.35)
Variance pupil level (SE)	2509.62 (147.71)	1413.91 (83.43)
Total sample size (pupils)	642 (60)	639 (60)

Table E17b: Analysis of sentence completion (secondary outcome) - Year 7 pupils (Only girls)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.03 (-0.15 to 0.12)
Regression coefficient (95% CI)	-5.06 (-14.72 to 4.61)	-1.13 (-5.75 to 2.36)
P-value	0.305	0.632
ICC (SE)	0.111 (0.035)	0.000 (0.000)
Variance class level (SE)	202.16 (68.61)	0.000 (0.000)
Variance pupil level (SE)	1624.20 (93.53)	890.22 (48.82)
Total sample size (pupils)	667 (60)	665 (60)

Table E17c: Analysis of sentence completion (secondary outcome) - Year 7 pupils, Gender interaction model

	Adjusted analysis
Regression coefficient of the interaction between free school meals and experimental cell (95% CI)	0.89 (-6.63 to 8.41)
P-value	0.816
ICC (SE)	0.017 (0.013)
Variance class level (SE)	19.67 (15.20)
Variance pupil level (SE)	1169.21 (47.13)
Variance component on Gender (SE)	0.000 (0.000)
Total sample size (classes)	1,304 (60)

Table E18: Analysis of sentence completion (secondary outcome) - Year 7 pupils (Only boys) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.735	0.032	23.150	0.000	0.672	0.797
Free School Meals (FSM)	-4.782	4.103	-1.170	0.244	-12.825	3.260
Birth month - February	4.761	7.575	0.630	0.530	-10.085	19.607
Birth month - March	0.283	7.599	0.040	0.970	-14.610	15.177
Birth month - April	9.892	7.275	1.360	0.174	-4.366	24.150
Birth month - May	-7.517	7.536	-1.000	0.319	-22.288	7.254
Birth month - June	9.917	7.466	1.330	0.184	-4.717	24.550
Birth month - July	7.637	7.288	1.050	0.295	-6.647	21.921
Birth month - August	-1.388	7.427	-0.190	0.852	-15.944	13.169
Birth month - September	1.814	7.274	0.250	0.803	-12.443	16.071
Birth month - October	3.841	7.206	0.530	0.594	-10.283	17.965
Birth month - November	6.154	7.292	0.840	0.399	-8.139	20.447
Birth month - December	3.967	7.867	0.500	0.614	-11.451	19.385
Allocation	-1.559	3.728	-0.420	0.676	-8.866	5.748
School 2	0.198	8.273	0.020	0.981	-16.016	16.412
School 4	0.119	7.273	0.020	0.987	-14.136	14.373
School 5	-1.338	7.747	-0.170	0.863	-16.522	13.846
School 6	3.440	8.275	0.420	0.678	-12.778	19.658
School 7	-12.110	9.190	-1.320	0.188	-30.123	5.903
School 8	-11.318	10.086	-1.120	0.262	-31.086	8.451
School 9	-8.087	8.635	-0.940	0.349	-25.011	8.837
School 10	-0.735	6.716	-0.110	0.913	-13.897	12.428
School 3	-4.847	7.520	-0.640	0.519	-19.586	9.893
Intercept	104.055	12.797	8.130	0.000	78.974	129.136
Number of pupils	639					
Number of classes	60					

Table E19: Analysis of sentence completion (secondary outcome) - Year 7 pupils (Only girls) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.757	0.030	25.360	0.000	0.699	0.816
Gender	-1.127	3.091	-0.360	0.715	-7.185	4.930
Birth month - February	-9.664	5.857	-1.650	0.099	-21.143	1.816
Birth month - March	1.218	5.548	0.220	0.826	-9.656	12.092
Birth month - April	3.954	6.040	0.650	0.513	-7.885	15.793
Birth month - May	0.359	5.720	0.060	0.950	-10.852	11.570
Birth month - June	2.784	5.522	0.500	0.614	-8.040	13.607
Birth month - July	1.007	6.030	0.170	0.867	-10.813	12.826
Birth month - August	3.905	5.539	0.700	0.481	-6.952	14.761
Birth month - September	-6.838	5.700	-1.200	0.230	-18.011	4.335
Birth month - October	-0.008	5.571	0.000	0.999	-10.927	10.912
Birth month - November	7.512	5.966	1.260	0.208	-4.182	19.205
Birth month - December	-3.132	5.829	-0.540	0.591	-14.557	8.292
Allocation	-1.128	2.358	-0.480	0.632	-5.749	3.493
School 2	-2.861	5.243	-0.550	0.585	-13.138	7.416
School 4	0.303	4.473	0.070	0.946	-8.463	9.070
School 5	-1.570	4.825	-0.330	0.745	-11.026	7.886
School 6	3.169	5.555	0.570	0.568	-7.719	14.057
School 7	-11.655	6.988	-1.670	0.095	-25.352	2.041
School 8	-7.676	5.945	-1.290	0.197	-19.328	3.976
School 9	-10.102	5.010	-2.020	0.044	-19.921	-0.283
School 10	-1.847	4.302	-0.430	0.668	-10.280	6.585
School 3	-6.279	5.350	-1.170	0.241	-16.764	4.207
Intercept	93.449	11.138	8.390	0.000	71.619	115.279
Number of pupils	665					
Number of classes	60					

Table E20: Analysis of sentence completion (secondary outcome) - Year 7 Gender interaction model – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.741	0.022	33.920	0.000	0.699	0.784
FSM eligibility	-2.980	2.537	-1.170	0.240	-7.952	1.992
Gender	-6.563	2.656	-2.470	0.013	-11.769	-1.357
Birth month - February	-2.761	4.782	-0.580	0.564	-12.134	6.612
Birth month - March	0.411	4.638	0.090	0.929	-8.680	9.502
Birth month - April	6.865	4.750	1.450	0.148	-2.445	16.176
Birth month - May	-3.539	4.720	-0.750	0.453	-12.789	5.711
Birth month - June	5.435	4.594	1.180	0.237	-3.568	14.438
Birth month - July	4.349	4.756	0.910	0.361	-4.973	13.671
Birth month - August	1.098	4.608	0.240	0.812	-7.934	10.130
Birth month - September	-3.022	4.633	-0.650	0.514	-12.103	6.059
Birth month - October	1.615	4.544	0.360	0.722	-7.291	10.520
Birth month - November	6.535	4.739	1.380	0.168	-2.754	15.823
Birth month - December	-0.448	4.852	-0.090	0.926	-9.957	9.062
Allocation	-2.057	2.975	-0.690	0.489	-7.889	3.774
School 2	-2.330	4.952	-0.470	0.638	-12.036	7.376
School 4	-0.015	4.331	0.000	0.997	-8.503	8.473
School 5	-1.582	4.652	-0.340	0.734	-10.699	7.535
School 6	2.615	5.109	0.510	0.609	-7.398	12.628
School 7	-12.785	5.872	-2.180	0.029	-24.294	-1.276
School 8	-9.489	5.882	-1.610	0.107	-21.017	2.039
School 9	-9.579	4.982	-1.920	0.055	-19.343	0.186
School 10	-1.807	4.093	-0.440	0.659	-9.830	6.215
School 3	-6.440	4.747	-1.360	0.175	-15.745	2.865
Gender * Allocation	0.891	3.836	0.230	0.816	-6.627	8.409
Intercept	104.378	8.598	12.140	0.000	87.526	121.229
Number of pupils	1304					
Number of classes	60					

Table E21a: Analysis of sentence completion (secondary outcome) - Year 9 pupils (Only boys)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	0.04 (-0.07 to 0.15)
Regression coefficient (95% CI)	3.57 (-6.42 to 13.56)	2.56 (-3.76 to 8.90)
P-value	0.483	0.427
ICC (SE)	0.018 (0.020)	0.000 (0.000)
Variance class level (SE)	64.53 (71.57)	0.000 (0.000)
Variance pupil level (SE)	3526.28 (201.88)	1653.51 (90.41)
Total sample size (pupils)	669 (60)	669 (60)

Table E21b: Analysis of sentence completion (secondary outcome) - Year 9 pupils (Only girls)

	Unadjusted analysis	Adjusted analysis
Effect Size (CI)	-	-0.06 (-0.18 to 0.06)
Regression coefficient (95% CI)	-12.71 (-21.45 to -3.97)	-3.15 (-8.76 to -2.46)
P-value	0.004	0.271
ICC (SE)	0.016 (0.020)	0.000 (0.000)
Variance class level (SE)	40.25 (52.47)	0.000 (0.000)
Variance pupil level (SE)	2527.07 (151.92)	1177.06 (67.62)
Total sample size (pupils)	607 (60)	606 (60)

Table E21c: Analysis of sentence completion (secondary outcome) - Year 9 pupils, Gender interaction model

	Adjusted analysis
Regression coefficient of the interaction between free school meals and experimental cell (95% CI)	-7.10 (-15.55 to 1.36)
P-value	0.100
ICC (SE)	0.000 (0.000)
Variance class level (SE)	0.000 (0.000)
Variance pupil level (SE)	0.000 (0.000)
Variance component on Gender (SE)	0.000 (0.000)
Total sample size (classes)	1,275 (60)

Table E22: Analysis of sentence completion (secondary outcome) - Year 9 pupils (Only boys) – Regression output

|	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.810	0.030	26.970	0.000	0.751	0.868
Free School Meals (FSM)	-2.784	4.257	-0.650	0.513	-11.128	5.560
Birth month - February	1.250	7.921	0.160	0.875	-14.275	16.775
Birth month - March	-5.079	7.863	-0.650	0.518	-20.490	10.331
Birth month - April	3.861	7.732	0.500	0.618	-11.293	19.015
Birth month - May	2.739	7.594	0.360	0.718	-12.146	17.624
Birth month - June	-7.263	7.757	-0.940	0.349	-22.468	7.941
Birth month - July	-6.768	7.669	-0.880	0.377	-21.798	8.262
Birth month - August	-7.667	7.682	-1.000	0.318	-22.723	7.389
Birth month - September	8.850	7.771	1.140	0.255	-6.382	24.081
Birth month - October	2.529	7.631	0.330	0.740	-12.427	17.485
Birth month - November	7.876	7.818	1.010	0.314	-7.446	23.199
Birth month - December	8.052	7.550	1.070	0.286	-6.746	22.850
Allocation	2.560	3.224	0.790	0.427	-3.760	8.879
School 2	3.048	7.070	0.430	0.666	-10.808	16.905
School 4	9.592	6.064	1.580	0.114	-2.292	21.477
School 5	0.154	6.850	0.020	0.982	-13.272	13.580
School 6	-0.206	7.013	-0.030	0.977	-13.950	13.539
School 7	-17.815	8.030	-2.220	0.027	-33.553	-2.077
School 8	-10.288	7.927	-1.300	0.194	-25.825	5.248
School 9	7.223	6.630	1.090	0.276	-5.772	20.217
School 10	5.811	5.875	0.990	0.323	-5.704	17.327
School 3	-5.730	6.978	-0.820	0.412	-19.407	7.947
Intercept	81.949	12.652	6.480	0.000	57.152	106.746
Number of pupils	662					
Number of classes	60					

Table E23: Analysis of sentence completion (secondary outcome) - Year 9 pupils (Only girls) – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.850	0.033	25.560	0.000	0.785	0.915
Gender	-10.225	3.566	-2.870	0.004	-17.214	-3.235
Birth month - February	-4.507	6.872	-0.660	0.512	-17.976	8.962
Birth month - March	-3.385	6.604	-0.510	0.608	-16.327	9.558
Birth month - April	-0.358	7.355	-0.050	0.961	-14.773	14.056
Birth month - May	2.879	6.935	0.420	0.678	-10.713	16.471
Birth month - June	-2.439	6.560	-0.370	0.710	-15.295	10.418
Birth month - July	-8.786	6.499	-1.350	0.176	-21.525	3.952
Birth month - August	-4.089	6.685	-0.610	0.541	-17.191	9.013
Birth month - September	5.064	6.516	0.780	0.437	-7.707	17.836
Birth month - October	-4.696	7.142	-0.660	0.511	-18.694	9.302
Birth month - November	-5.286	6.702	-0.790	0.430	-18.422	7.850
Birth month - December	6.243	6.820	0.920	0.360	-7.123	19.609
Allocation	-3.151	2.863	-1.100	0.271	-8.763	2.461
School 2	-1.121	6.451	-0.170	0.862	-13.766	11.523
School 4	0.151	5.328	0.030	0.977	-10.291	10.593
School 5	3.705	5.283	0.700	0.483	-6.649	14.060
School 6	5.848	5.972	0.980	0.327	-5.857	17.554
School 7	-9.006	7.006	-1.290	0.199	-22.737	4.725
School 8	0.196	7.277	0.030	0.979	-14.068	14.459
School 9	-4.360	6.000	-0.730	0.467	-16.120	7.399
School 10	3.859	5.002	0.770	0.440	-5.946	13.663
School 3	-8.548	6.254	-1.370	0.172	-20.805	3.709
Intercept	68.059	13.936	4.880	0.000	40.746	95.373
Number of pupils	603					
Number of classes	60					

Table E24: Analysis of sentence completion (secondary outcome) - Year 9 Gender interaction model – Regression output

	Coefficient	Std. Err.	z	P>z	95% Conf. Interval	
Pre-intervention score	0.826	0.022	37.310	0.000	0.783	0.870
FSM eligibility	-6.286	2.771	-2.270	0.023	-11.718	-0.855
Gender	-3.785	3.011	-1.260	0.209	-9.687	2.117
Birth month - February	-1.361	5.299	-0.260	0.797	-11.747	9.025
Birth month - March	-3.726	5.159	-0.720	0.470	-13.837	6.386
Birth month - April	1.853	5.376	0.340	0.730	-8.683	12.389
Birth month - May	2.485	5.196	0.480	0.632	-7.699	12.669
Birth month - June	-4.310	5.100	-0.850	0.398	-14.305	5.685
Birth month - July	-8.039	5.032	-1.600	0.110	-17.901	1.823
Birth month - August	-6.780	5.134	-1.320	0.187	-16.842	3.283
Birth month - September	5.827	5.086	1.150	0.252	-4.141	15.795
Birth month - October	-0.736	5.271	-0.140	0.889	-11.068	9.595
Birth month - November	1.492	5.190	0.290	0.774	-8.679	11.663
Birth month - December	7.557	5.127	1.470	0.141	-2.493	17.607
Allocation	2.946	2.979	0.990	0.323	-2.893	8.785
School 2	1.423	4.808	0.300	0.767	-8.000	10.846
School 4	5.232	4.054	1.290	0.197	-2.714	13.179
School 5	2.428	4.292	0.570	0.572	-5.984	10.841
School 6	2.595	4.636	0.560	0.576	-6.492	11.681
School 7	-13.397	5.374	-2.490	0.013	-23.931	-2.864
School 8	-6.619	5.417	-1.220	0.222	-17.236	3.997
School 9	1.812	4.494	0.400	0.687	-6.997	10.621
School 10	4.592	3.879	1.180	0.237	-3.011	12.195
School 3	-7.045	4.735	-1.490	0.137	-16.325	2.236
Gender * Allocation	-7.095	4.313	-1.650	0.100	-15.548	1.357
Intercept	78.309	9.324	8.400	0.000	60.035	96.583
Number of pupils	1265					
Number of classes	60					

Appendix F: Security classification of trial findings

Evaluation design   Implementation   Analysis and interpretation
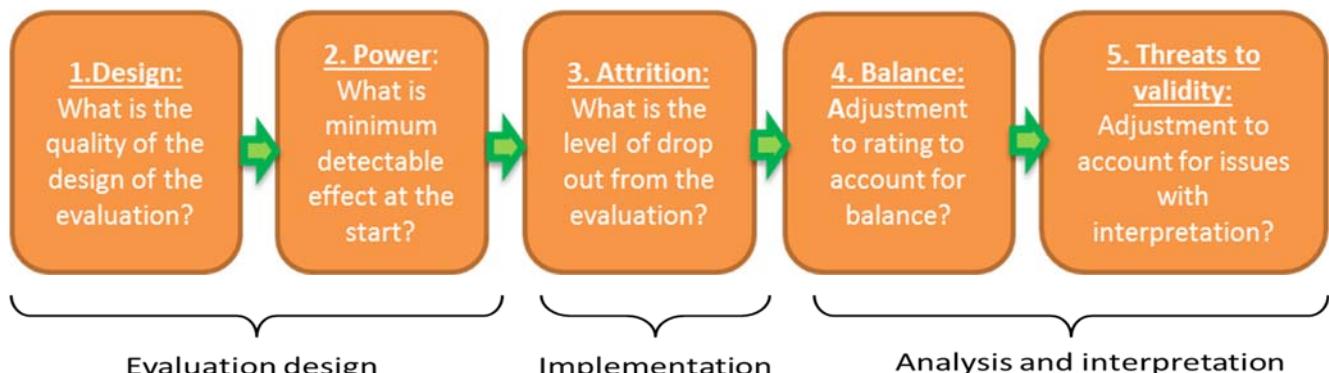
Rating	1. Design	2. Power (MDES)	3. Attrition	4. Balance	5. Threats to validity
5 🔒	Fair and clear experimental design (RCT)	< 0.2	< 10%	Well-balanced on observables	No threats to validity
4 🔒	Fair and clear experimental design (RCT, RDD)	< 0.3	< 20%		
3 🔒	Well-matched comparison (quasi-experiment)	< 0.4	< 30%		
2 🔒	Matched comparison (quasi-experiment)	< 0.5	< 40%		
1 🔒	Comparison group with poor or no matching	< 0.6	< 50%		
0 🔒	No comparator	> 0.6	> 50%	Imbalanced on observables	Significant threats

The final security rating for this trial is 4 🔒. This means that the conclusions have moderate to high security.

This evaluation was designed as a randomised controlled trail. The sample size was designed to detect a MDES of less than 0.2, by design. At the unit of randomisation (class), there was zero attrition, and extremely low attrition at the pupil level also. Some of the baseline measures of attainment were a little imbalanced at baseline (at the level of ES=0.1), but since many measures were compared and attrition was low it is likely that these small differences arose by chance. The post-tests were administered by the schools by teachers who were aware of the treatment allocation. As there was no independent invigilation (of even a sample of schools), the padlock rating is reduced to 4 🔒.

Appendix G: 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.

Education
Endowment
Foundation