Education Endowment Foundation

PROJECT TITLE	READING WITH COMPREHENSION (REACH) Primary, a two-arm cluster randomised trial
DEVELOPER (INSTITUTION)	University of Leeds
EVALUATOR (INSTITUTION)	Sheffield Hallam University
PRINCIPAL INVESTIGATOR(S)	Professor Mike Coldwell
PROTOCOL AUTHOR(S)	Dr Martin Culliney, Dr Josephine Booth, Professor Mike Coldwell, Dr Karen Daniels, Sean Demack
TRIAL DESIGN	Two-arm multisite cluster randomised controlled trial with random allocation at school level
PUPIL AGE RANGE AND KEY STAGE	Year 3 (Age 7-8, KS2) (extended into Y4 due to Covid-19)
NUMBER OF SCHOOLS	79
NUMBER OF PUPILS	780
PRIMARY OUTCOME	Wechsler Individual Achievement Test - Third UK edition for teachers (WIAT III)
SECONDARY OUTCOMES	Diagnostic Test of Word Reading Processes (DTWRP) Understanding Spoken Paragraphs subscale of the Clinical Evaluation of Language Fundamentals (CELF - 5)

Protocol version history

VERSION	DATE	REASON FOR REVISION
1.0	09/07/2019	Original Version
2.0	19/02/2020	Change to primary outcome measure
3.0	04/11/2020	Change to outcome testing arrangements

Page	Protocol version 2 amendments
1	NGRT outcome changed to WIAT III (and throughout main text)
1	School and pupil numbers updated to reflect number randomised
11	Added description of primary outcome piloting process
12	Confirmed that KS1 Reading baseline data collected in sufficient quantity and Phonics data from NPD therefore not necessary
12	Stated that as SHU are now no longer invigilating 25% of secondary outcome tests; SHU now conducting primary outcome tests at same time and able to observe.
22	Fidelity and dosage section updated
23	Added trial registration details
27	Timetable - SHU no longer invigilating 25% of secondary outcome tests; SHU now conducting primary outcome tests at same time and able to observe. Also amended start dates for outcome tests, to begin in April 2020

Page	Protocol version 3 amendments
6-7	Changes in delivery due to Covid-19
13-14	Changes to outcome testing due to Covid-19
25	Changes to data protection section to note updated information provided to parents/carers
26	Changes to staffing
27	Changes to timeline due to Covid-19

Table of contents

Intervention[1]

1. Brief name
REACH (Reading with Comprehension) Primary

2. Why (rationale/theory)
In 2015, the UK government reported that 29% of students did not reach the expected standard in reading at the end of primary school (Department for Education, 2015). According to the Simple View of Reading (Gough & Tunmer, 1986), two components are essential to reading with meaning: decoding and language comprehension. REACH Primary targets both components in one 20 week intervention programme, using established techniques. Reading Intervention (RI), which includes Sound Linkage (Hatcher, Duff & Hulme, 2014), is used to develop word recognition and decoding skills. Oral Language (OL) Intervention (Clarke et al. 2010; Clarke et al. 2013) is used to promote language comprehension and focuses on vocabulary, figurative language, strategy use and narrative.

---

[1] See Logic Model on p21 for a visual representation that supplements this section.

A previous efficacy trial found that school staff reported that REACH had a positive effect on reading skills, reading accuracy and confidence for secondary school pupils. On average, around six months additional progress was reported where children had received both the reading intervention and language intervention, and four months additional progress when experiencing only the reading intervention (EEF, 2016). There was no clear evidence that the interventions improved reading comprehension in particular, as opposed to other skills such as word recognition.

This evaluation design specifically aims to build on implications and findings from the previous trial and evidence review by evaluating the efficacy of the intervention in a primary school setting (EEF, 2016; Clarke et al., 2017).

3. Who (recipients)
The intervention is aimed at Year 3 children who have been identified as performing below national expectations in reading by their school during Year 2. The decision to target Y3 is pragmatic as the intervention is expected to have larger effects in this group due to the phonics component, and to reduce the testing burden on schools by using KS1 scores as a pre-test. Selected pupils will include those with EAL, SEND and FSM to help understand if there are differential outcomes for children with differing needs and starting points. At the pupil level teachers will be asked to identify the 10 lowest performing pupils in Year 3 with regard to the following word recognition and decoding abilities:

1.   Read regular words (e.g. CATCH) aloud accurately and fluently
2.   Read irregular words (e.g. YACHT) aloud accurately and fluently
3.   Read nonsense words (e.g. FLENK) aloud accurately and fluently
4.   Read sentences aloud accurately and fluently

Selection should be made according to these areas of ability. Teachers may wish to use data from national tests to inform their selection, or from standardised measures, or from their own teacher designed assessments, or a combination of these. Pupils are eligible to take part if they show difficulties in one or more of these abilities. There are no exclusionary criteria. The selected sample might include pupils with special educational needs, disabilities or sensory impairments, pupils new to the English language or those who speak English as an additional language.

Schools must be two-form entry or larger due to the number of pupils required for statistical analysis and the workload that can reasonably be expected of participating TAs. Schools must be based in the North East, Yorkshire and Humber, Greater Manchester, or Lincolnshire regions of England, where training will be delivered.

4. What (materials)
In preparation for delivering the REACH Primary programme, Teaching Assistants (TAs) will receive three days of face-to-face training. There will be associated online seminars and a series of tasks linked to these. The first two training days will be scheduled on consecutive days prior to intervention delivery. The designated SLT member should attend the first half day of training. The third day will be 'top-up training' and take place after the first 10 weeks of delivery. TAs will be given a manual, in both printed and digital format, detailing the intervention programme. This will include baseline assessment instructions and scripted

session plans. A training video with modelled example sessions and an online space for sharing resources will also be provided. Ongoing advice will be available from the training team via email and telephone; an online forum (moderated by the training team) will also be in operation, primarily to facilitate peer support and discussion. Where necessary, schools can request additional onsite observation and training from the training team, or alternatively can submit video/audio recorded intervention sessions to the training team for feedback.

Each school in the intervention group will receive a book box – this is a 'starter set' of two books per level for levels 1-30 of the Hatcher (2000) grading system, which takes into account word length, sentence length, and grammatical features of text. The book boxes will be assembled and graded by the delivery team and will cover a range of genres and interests. Graded passages from a range of online sources will also be recommended for each level. Schools are required to contribute £300 to cover the cost of the book box, which is then kept by the schools. Each intervention school will also receive a copy of Sound Linkage and a resource pack (including tactile letters, idiom and jokes cards and THRASS charts). Teaching assistants will need access to tablet computers, photocopying and general teaching materials such as white boards and pens.

5. What (procedures)

Week 1 of the programme involves the administration of baseline assessments. These take place over three 30 minute sessions and cover the following:

- Concepts about print
- Letter identification
- Phoneme identification
- Sound Linkage (Hatcher Duff and Hulme, 2014) –
    o Phonological awareness
    o Syllable Blending
    o Rhyme
    o Phoneme Blending
    o Phoneme Segmentation
    o Phoneme Deletion
    o Phoneme Transposition
    o Spoonerisms
- Early word recognition
- Non word reading
- Burt Reading Test
- Text Reading - Running record
- Free writing
- Dictation

The data from these assessments provide the TAs with a detailed and comprehensive profile of each pupil's strengths and weaknesses which is then used to tailor the content of the RI sessions to each individual.

In Weeks 2-20 of the programme pupils complete two RI sessions and one OL session each week. All sessions last for 30 minutes. Both use a distributed practice approach with routine formats. The methods used in the programme draw on established teaching practices and principles, including social constructivism, authentic literacy pedagogy and dialogic teaching.

The RI sessions comprise:

First section: total duration 10 minutes
      1) Reading an 'easy' book (a book read with at least 95% accuracy)
      2) A running record of an 'instructional' book (a book read with 90–94% accuracy)
Second section: 15 minutes
      3) Letter and word identification activities – using tactile letters, worksheets and games
      4) Phonological activities – taken from Sound Linkage (Hatcher, Duff & Hulme, 2014), reinforced using THRASS chart and tactile letters and digraphs
      5) Cut-up sentence activity – assembling a sentence from cut up sections containing individual words or phrases
      6) Write a sentence activity – creating a new sentence (which is then used in subsequent sessions in the cut-up sentence activity)
Third section: 5 minutes
      7) Introduction to new 'instructional' book
      8) Attempt at reading new 'instructional' book.

The OL sessions are discussion based. Pupils listen to the stories being read aloud by the TA. The stories include a range of complex vocabulary, figurative language, interesting storylines and opportunities for practicing reciprocal teaching strategies. Using the stories as the backbone of each session, the activities pupils complete are:

First section: 10 minutes
      1) Vocabulary – Each session includes one 'word of the day', these are all tier 2 words, which are present in the story. They are explicitly taught using the Multiple Context Learning (MCL) approach (Beck, McKeown & Kucan, 2002). Graphic organisers, definition cards and illustrations are used to consolidate word meanings.
Second section: 10 minutes
      2) Listening Comprehension – Pupils listen to the story and then complete a worksheet or game-like activity inspired by Reciprocal Teaching (RT) (Palincsar & Brown, 1984). The four RT strategies that the pupils practice are Clarification, Summarisation, Prediction and Question Generation.
Third section: 5 minutes
      3) Figurative Language – Pupils discuss figurative language examples present in the story and further develop their understanding of figurative language using illustrated cards containing Idioms, Jokes and Riddles that are related in some way to the story.
Fourth section: 5 minutes
      4) Spoken Narrative – Pupils create a story map collaboratively with the TA. They use the map to discuss key parts of the narrative content and structure, including characters, places, timelines and key events. Oral story retells, using the story map as a prompt are also regularly practised.

6. Who (implementers)
REACH Primary is designed to be delivered by Teaching Assistants (TAs), who have been nominated by the school and received training from the University of Leeds as specified above, following initial agreement from the head teacher. The training tasks are designed to be completed in pairs, with peer support built in to the approach. There will be an agreed

senior leader in the school (such as literacy coordinator, SENCO or HLTA) acting as a point of contact for each school.

7. How (mode of delivery)
Over the course of 20 weeks, each pupil receives 3 x 30 minute sessions of individual support per week. Two sessions per week focus on word recognition/decoding skills and reading aloud fluently and accurately, and the third session each week targets language comprehension skills.

8. Where (setting)
Pupils should be taken out of lessons for their sessions, which should be delivered on a one-to-one basis. Ideally the sessions would be on different days and replace guided reading, rather than supplementing normal lessons. The sessions require a quiet designated space in school, outside of the pupil's regular classroom. Schools will need to commit to providing this for the duration of the project.

9. When and how much (dosage)
Each pupil will receive 3 x 30 minute sessions of individual support per week for 20 weeks. This amounts to 30 hours per pupil.

10. Tailoring
The intervention uses approaches which require personalisation. This is especially true for the RI sessions in which TAs will be responsible for selecting appropriate level texts (using the Hatcher grading system) and choosing from a menu of letter, word and phonological activities. They will do this using the information obtained through the week one baseline assessments. The OL sessions are scripted and based on specific texts; however the activities themselves are designed to be flexible and responsive to pupils' experiences and needs. TAs will be encouraged to include bespoke materials to support these activities (for example, photographs and objects from the local community and other contexts familiar to the pupils). For some pupils tablet technology may be particularly suitable and TAs have the option to personalise the sessions using paper based and online texts as appropriate. Schools should also allow planning time for TAs, recommended at 15/20 minutes per pupil per week.

11. How well (planned)
Face-to-face training will be delivered through local hubs, with participating schools serving as training venues. The intention is that TAs will not have to travel far in order to attend. TAs will have access to all relevant resources. Support will be available via telephone, email and online for TAs throughout, and bespoke on-site training can be requested if necessary. The project aims to complete the TA training as early as possible in the 2019/20 school year in order to give the maximum amount of time for TAs to deliver the intervention. TAs should have access to a designated space for storage of materials for the duration of the project.

Changes in delivery due to Covid-19
At the point that schools in England were closed due to Covid-19 (end March 2020), the majority of participating schools were near to finishing the scheduled REACH Primary delivery and had time to reach the threshold for optimal fidelity described in this protocol (see below). It was therefore agreed that fidelity would be measured according to the criteria set out in the previous version of the protocol, covering only the delivery period originally scheduled (Autumn and Spring terms 2019/20). An additional training session was offered to TAs in intervention schools in September 2020 to support five weeks of extra delivery intended to compensate for the interruption to the programme. There were no material changes to delivery aside from this.

Outcome testing, planned for Summer 2020, is now taking place in November and December 2020, so there is a longer gap between the end of the intervention and testing than was planned. A number of schools have declined the opportunity to continue with the programme and have withdrawn from outcome testing. The implications for the sample size and indicative power calculations are discussed in the updated SAP (version 2.0). Outcome testing is being conducted remotely with test administrators from the evaluation team using video calls to connect with participating schools and collect the required data. The outcome measures remain unchanged. This is detailed below.

Study rationale and background

REAding with CompreHension Primary (REACH Primary) is a targeted intervention for struggling readers that comprises two strands: Reading Intervention (RI) sessions, delivered twice per week, and Oral Language Intervention (OL) sessions, delivered weekly.

While the previous trial showed that the impact of the intervention on pupil reading was positive, it received a low security rating (2 EEF padlocks) as the trial was phased, not all pupils completed tests, and differences in characteristics of pupils between intervention and control schools were observed. In addition, questions over implementation were raised, for instance regarding TA training, delivery and ongoing support, with clear implications for reviewing the content of the training materials and intervention itself. To address this feedback the TA training programme has been redesigned. It now has a more distributed design, with top-up training and a blended learning component (online seminars and training tasks). Furthermore the RI section of the training manual has been revised to make it more user-friendly and the range of resources available to support the RI activities has been increased; TAs will now have access to digital as well as printed resources. These two interrelated aspects of the programme, and their impact on pupil reading progress, require further investigation. In view of this, the IPE design will specifically examine the content and process of TA training and support sessions, the actual extent of intervention received by pupils, and TA perception of how far their training and the delivered intervention supports their confidence and competence in supporting pupil gains in reading.

The Simple View of Reading

The REACH intervention is primarily guided by the two components of the Simple View of Reading, word recognition (decoding) and language comprehension. Gough and Tunmer (1986) define the reading process as:

Decoding (D) x Language Comprehension (LC) = Reading Comprehension (RC)

Whilst the Simple View of Reading draws together the relationship between oral language comprehension and word recognition, reading comprehension (RC) involves the process of constructing meaning from printed words (Hoover and Gough, 1990). This necessitates further consideration of what the process involves.

Reading Comprehension Processes

Children at the early stage of reading, or those who lack efficient decoding skills, often work hard to decode words on the page and this may confound the processes of reading comprehension. Oral language comprehension (including vocabulary knowledge, grammatical understanding, and pragmatic awareness) is a key component of the Simple View of Reading. REACH Primary supports both the language comprehension and decoding dimensions of the reading process. Due to the complexity inherent in reading comprehension, we are particularly interested in identifying perceptions of which elements of

the REACH intervention directly support each dimension, and therefore understanding the processes that support gains in reading comprehension.

Motivational Factors

It is of significant interest that the process evaluation in the previous trial of REACH reported positive attitudinal changes for some children who had undertaken the reading intervention. Conversely, TAs noted difficulty maintaining pupil engagement during oral language comprehension activities. A bidirectional relationship between the will to read and the development of reading skills exists (Morgan and Fuchs, 2007; OECD, 2010). Hempel-Jorgenson et al (2018) suggest a 'virtuous cycle' of increased reading confidence through self-efficacy that in turn increases pupil reading motivation. This is a significant dimension to the study as recent research into reading has heightened recognition of the motivational and behavioural characteristics alongside cognitive characteristics of reading (OECD, 2016; Clark and Teravainen, 2017). Motivation to read is linked to appropriate reading goals, perceptions of self-efficacy and social motivation for reading (Guthrie and Wigfield, 2000).

The proposed trial thus aims to identify teacher and TA perceptions of the impact of components of the intervention that in particular appear to be supporting decoding and comprehension, and positive attitudinal changes. We suggest that increased self-efficacy, confidence and motivation to read is an area worthy of investigation as this may indicate potential for sustained impact of the intervention over time.

Impact Evaluation

Research questions

RQ1.   What is the impact of REACH Primary on pupil reading comprehension ability, as measured by the WIAT III Reading Comprehension subtest?

RQ2.   What is the impact of REACH Primary on pupil word recognition and decoding ability, as measured by GL DTWRP?

RQ3.   What is the impact of REACH Primary on pupil language comprehension, as measured by the 'Understanding Spoken Paragraphs' module of Pearson CELF-5?

Design

The impact of REACH Primary will be evaluated through a two-arm, cluster randomised controlled trial (CRT). The main design elements are summarised in Table 1. Specifically, this is a multisite CRT with three levels (school, TA and pupil). Schools will be randomised within five geographical hub areas. The design acknowledges the clustering of the primary outcome (Wechsler Individual Achievement Test, Third UK edition for teachers, Reading Comprehension subtest, henceforth WIAT III) at the school level and within schools at the TA level. The feasibility of this design will become apparent during the trial period. For example, if TAs commonly support all pupils eligible for the REACH programme (across classes) rather than with a mutually exclusive group of pupils (within different classes), the school and TA levels would not be distinguishable. If this is the case, a two level design will be adopted and this will be published in an updated protocol and Statistical Analysis Plan.

Table 1: Trial design and outcome variables

Trial type and number of arms		Two-arm, cluster randomised
Unit of randomisation		School
Stratification variables (if applicable)		Geographical area, school level mean KS1 reading score for participating pupils
Primary outcome	variable	Pearson WIAT III (Reading Comprehension subtest)
	measure (instrument, scale)	Raw scores
Secondary outcome(s)	variable(s)	Diagnostic Test of Word Reading Processes (decoding of words, non-words and exception words) CELF-5 'Understanding Spoken Paragraphs'
	measure(s) (instrument, scale)	DTWRP: standard age scores CELF-5: paper version, raw scores

Randomisation

Randomisation will take place at the school level for practical and methodological reasons. Practically, for recruitment and implementation of the REACH Primary intervention, a school level approach is preferable. Methodologically, school level randomisation reduces the risk of 'spillover' from the intervention to the control group. The previous REACH trial used within-school randomisation and spillover effects were suspected.

Randomisation will be stratified to ensure balance between the intervention and control groups on two school-level controls; geographical hub area and KS1 reading score. The first stratifying variable will be geographical region, the five categories will be 1) Yorkshire and Humber (North) 2) Yorkshire and Humber (South) 3) North East 4) Lincolnshire and 5) Greater Manchester. This data will be collected directly from schools prior to randomisation. Analysts will not remain blinded to allocation during the evaluation, but randomisation will assign the two groups values of 0 and 1.

Participants

Recruitment will be at the school level. The target is to recruit 80 large primary schools (with two classes in Year 3) from the North East, Yorkshire and Humber, Greater Manchester, and Lincolnshire regions of England. Schools will identify the 10 weakest readers (5 per class) who will be starting Year 3 in September 2019. All participating schools will be required to select five pupils from each class who struggle with reading to take part in the trial. Pupils without KS1 Reading scores are not eligible for the trial as this baseline measure is needed for analysis.

Sample size calculations

The power analysis presents estimated Minimum Detectable Effect Sizes (MDES) for the primary outcome (WIAT III) based on the formulae provided in Bloom et al (2007) and Spybrook et al. (2016). Please see the Appendix for details on these calculations. We present MDES estimates for both 2-level and 3-level multisite CRT designs, as the viability

of a TA level remains unclear (see earlier). This will be determined during the trial period and reported through an updated Protocol or Statistical Analysis Plan. Clustering at the school-level is estimated using a school level Intra Cluster Correlation Coefficient (ICC) of 0.14. Thus, we estimate that 14% of the variation in the primary outcome will lie at the school level. Demack (2019) highlighted how the strength of clustering at a class level is positively correlated with the extent of setting / streaming policies within the sample of schools involved in a trial. Demack (ibid) recommends using a class level ICC value of 0.10 for trials involving KS2 primary classes. However, the REACH trial design does not involve whole classes, it is an intervention targeted at pupils identified as struggling to read. This suggests that pupils selected for REACH are likely to be relatively homogenous across the sample of schools compared with a sample of classes in which some are grouped according to measured/perceived ability. We therefore assume a much weaker TA-level ICC of 0.02; assuming that 2% of the variation in the primary outcome would lie within schools between TA groups. We feel it is reasonable to assume weaker clustering at the TA level for REACH compared with class-level clustering in KS2 more generally and have selected the ICC estimate of 0.02 to reflect this.

Prior to randomisation we will collect KS1 reading test scores from recruited schools. This will be used to examine the structure of this KS1 test data and the strength of clustering at the school and TA levels. Once this data is analysed we will update the ICC estimates and power analyses. Whilst this KS1 data is unlikely to perfectly reflect the structure of the WIAT III outcome data, it is preferable to draw on empirical evidence to inform ICC estimates. At this point in time we do not have these empirical estimates but the KS1 data collected from schools will enable us to do so early in the trial. Table 2 below presents the estimated minimum effect sizes that could be detected as statistically significant ($p<0.05$) with a statistical power of 80%. Please see Appendix A for the formula and calculation.

☐ **School level ICC**	**0.14**
☐ **TA level ICC**	**0.02**
☐ **Pupil-level correlation between baseline & outcome**	**0.74**
☐ **School & TA-level correlation between baseline & outcome**	**0.60**
☐ **Balanced design (half of schools randomly selected to receive intervention)**	
☐ **Five geographical sites; 16 schools per site; Two TAs per school**	
☐ **Five pupils per TA (Overall); Two pupils per TA (FSM)**	
☐ **Six school-level covariates (see Appendix A)**	

Table 2: Sample size calculations

		OVERALL	FSM
MDES		0.24	0.29
Pre-test/ post-test correlations	level 1 (pupil)	0.74	0.74
	level 2 (TA)	0.60	0.60
	level 3 (school)	0.60	0.60
Intracluster correlations (ICCs)	level 2 (TA)	0.02	0.02
	level 3 (school)	0.14	0.14
Alpha		0.05	0.05
Power		0.80	0.80
One-sided or two-sided?		2	2
Average cluster size		5 pupils per TA, 10 per school	2 FSM pupils per TA, 4 per school
Number of schools	Intervention	40	40
	Control	40	40
	Total	80	80
Number of TAs	Intervention	80	80
	Control	80	80
	Total	160	160
Number of pupils	Intervention	400	160
	Control	400	160
	Total	800	320

Increasing the number of schools to 90 would improve sensitivity, with an MDES estimate of between 0.22 sds overall and 0.27 sds for the FSM subsample.

If the TA level is found to not be viable (for example, because TAs within a school work across classes, with all eligible pupils, making the TA and school levels indistinguishable), a multisite 2-level CRT design will be adopted that does not include the TA level. This design results in MDES estimates of 0.23 sds overall and 0.27 sds for the FSM subsample with 80 schools and 0.22 sds overall / 0.26 sds for the FSM subsample with 90 schools.

The aim is to recruit at least 80 schools and this results in MDES estimates notably lower than the +0.33 sd effect size found in the previous evaluation of REACH (EEF, 2016) for both 2-level and 3-level multisite CRT designs.

Please note that indicative power calculations for the sample size now anticipated following Covid-19 delays can be found in the updated SAP.

Outcome measures

For the primary outcome, we will use WIAT III. This test for pupils aged 6-16 measures

untimed reading comprehension. Various types of text are used including fictional stories and informational text. Pupils may read passages aloud or silently and orally respond to literal and inferential comprehension questions read aloud by the examiner. Only the reading comprehension section of WIAT III will be used.

The WIAT III test will be administered in schools by the evaluation team. All personnel will attend a mandatory training session beforehand. Schools must provide a suitable room for testing. TAs from the school will be asked to supervise the tests, as it is expected that the presence of familiar adults will make the pupils feel more comfortable. Pupils will be put into pairs and each pair will go to the allocated test room with a TA, one researcher from the evaluation team and another from the delivery team. One pupil will begin by doing the primary outcome test while the other does the secondary outcome tests (more details of which can be found below). Once finished, they will swap over, so that both primary and secondary outcome tests can be completed within a one-hour slot. This method ensures that there is enough time for all testing to take place during the normal school day without infringing on break times. During the pilot phase, testing was organised in this way and all parties were satisfied that it was a workable format.

The original plan was to use GL New Group Reading Test (NGRT). As a result, power calculations are based on the NGRT. Feasibility was assessed through pilot testing at a volunteer school in November 2019. There was a low level of test completion for NGRT and the test had to be stopped for some children as a result of low levels of engagement. It was agreed that alternative measures should be considered. The WIAT II test was suggested by the delivery team. This was piloted successfully at another volunteer school in January 2020. As the WIAT II is out of print, sufficient copies of the materials for this trial are not available. WIAT III, which features almost identical content, will therefore be used as the primary outcome measure.

Pilot schools will not form part of the trial but have been offered payment along with REACH Primary materials and the opportunity for their TAs to receive REACH Primary training.

The initial evaluation of REACH found a significant impact on reading comprehension as measured by NGRT, which includes both sentence level reading/completion and also a passage comprehension. For REACH Primary, WIAT III will be administered in school by the evaluation team under the supervision of participating TAs. Marking is done during the test as the answer booklet specifies whether each response merits two, one or zero points. Data entry of test results will be done by the evaluation team. Raw scores will be used in analysis.

Two secondary outcomes will be analysed. To measure word recognition/decoding we will use the Diagnostic Test of Word Reading Processes (DTWRP). This has three sets of items that pupils read aloud, 1) 30 exception words which provide a measure of lexical-semantic processes 2) 30 non-words which provide a measure of phonological recoding processes 3) 30 regular words which can be read by either process. This test aligns with the learning outcomes of the RI sessions and allows us to examine whether they are effective in improving word recognition and decoding skills.

To measure language comprehension we propose using the Understanding Spoken Paragraphs subscale of the Clinical Evaluation of Language Fundamentals (CELF -5). Pupils listen to short spoken passages, and after each passage orally respond to a series of comprehension questions. This test aligns with learning outcomes of the C sessions and allows us to examine whether the OL sessions are effective in improving listening

comprehension skills.

These secondary outcome assessments are to be administered in school by research assistants (RAs) working for the delivery team, under TA supervision. RAs will be assigned schools at random and they will be blinded to which condition each school is in. The schools will be informed in advance not to disclose to the RA whether they are in the intervention or control group. We recognise that in reality there may be some instances where pupils inadvertently reveal which condition they are in. RAs will run the tests, including audio recording verbatim during the tests to allow scoring and cross checking.

Once the data has been collected, the RAs will be responsible for scoring the tests and returning the test paper forms in person to the lead RA. All paper test forms will be labelled using codes and stored in a locked filing cabinet. All paper test forms will be scanned by the project administrator and stored as PDFs to facilitate efficient and secure data transfer between teams. Test forms will be shared with the Evaluation Team for independent verification of scoring (see below).

The lead RA will be responsible for data entry and will manually enter the test scores into a master spreadsheet; data entry will be blinded by using non-meaningful codes assigned by the project administrator. The lead RA will not know which condition an individual pupil is in at the point of data entry. The administrator will check the completed spreadsheet once all of the data has been entered, they will make sure that the data is correctly aligned with the relevant codes and check a minimum of 10% of the data by cross checking the scores with the paper test forms.

To guarantee the quality and integrity of the data collected by the Delivery Team, Sheffield Hallam will:

1) Be present during all secondary outcome assessments as they are taking place at the same time and location as the primary outcome assessments
2) Verify 5% of the scores using the tests shared by the Delivery Team. Tests to be reviewed will be randomly selected by the Evaluation Team and scored by a member of staff specialising in literacy measures (Karen Daniels).

Changes to outcome testing due to Covid-19

The WIAT III test will be administered remotely by the evaluation team due to difficulties in personal travel and school visits during the Covid-19 pandemic. Schools will be sent one copy of the stimulus book and the response form to aid completion. Test administrators will also have copies of these two items, and will conduct the assessment via video link. TAs or other school staff will hold the book and turn pages for pupils but have no other involvement. To maintain blinding, schools will be asked to not mention their treatment group or have REACH materials in view. Test administrators will have no knowledge of school allocation.

Tests were originally scheduled for Summer 2020 but school closures led to postponement. The delayed test period is now November-December 2020. As pupils are now in Y4, pupils will begin the WIAT III at age eight. The test is adaptive and pupils who score less than two on the first set of items are routed back to the starting point for age seven so that they have the opportunity to read an easier passage. Raw scores for all pupils are recorded in a single comparable scale (from 0 points to 42).

Secondary outcome assessments will also be conducted online via Microsoft Teams. The

DTWRP has been converted to a series of Powerpoint slides, and the Spoken Paragraphs test for the CELF has been pre-recorded. These will be shown/played via MS Teams screen sharing function. A member of school staff will be present for safeguarding and technology issues, however they will have no involvement in the testing. Two research assistants will be present on MS Teams, with one as the lead tester and one ensuring accurate recording. The testing will be audio recorded to minimise errors or 'mishearing' due to connectivity or sound quality issues. All data inputting will be checked by an RA, and 10% will have a third check by the Research Fellow from UoL to ensure consistency. As with the primary assessments, Research Assistants will have no knowledge of school allocation.

A sample of the secondary outcome assessments will still be checked by the evaluation team as described above but due to the change to remote test administration these tests will be done on a different day to the primary outcome and the evaluation team will not be present to observe the collection of secondary outcome data as per the original plan.
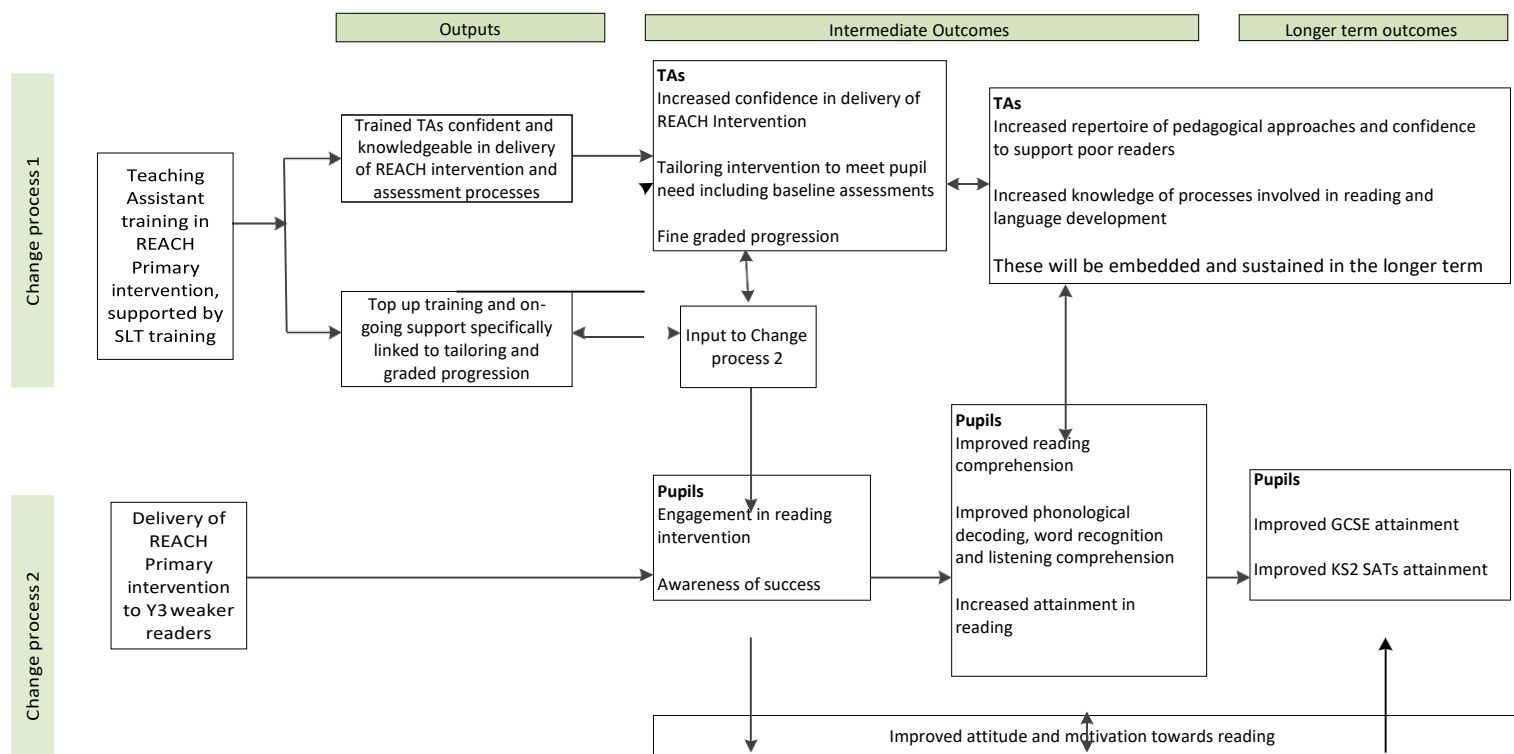
Analysis plan

A multilevel approach will be taken, with pupils clustered into TA groupings (these could correspond to classes) and TA groupings clustered into schools. This approach will be revised to a 2-level model (pupils clustered into schools) if there are problems identifying the TA level. Multilevel linear regression models will be constructed for the primary and secondary outcomes. The baseline measure is KS1 Reading (raw score), obtained directly from schools in June-July 2019 ahead of randomisation. The Phonics Screening Check was to be obtained from NPD and used as a baseline measure in the event that KS1 data from schools was sparse, but this was not necessary.

The first model will only include the school level group identifier (an outcome only model). The second model will also include KS1 Reading as a covariate at the pupil, TA (if 3-level design is adopted) and school level. The final model will also include the variables used within the stratified randomisation. This final model will be the headline ITT impact analysis for the primary outcome. For each model, the coefficient of the school-level dummy variable used to distinguish 'intervention group' pupils within the 40 schools who will receive the REACH Primary programme from 'control group' pupils will be converted into Hedges' g effect size statistics with 95% confidence intervals. The approach detailed in Hedges and Hedberg (2013) will be used, with pupil baseline scores centred around TA mean (assuming 3-level trial is feasible), centred around school mean.

Follow-on ITT exploratory analyses will focus on the impact of the REACH Primary programme on reading attainment for disadvantaged pupils, pupils with English as an additional language, and pupils with special educational needs or disabilities. Relevant indicators from the NPD will be used for these subgroup analyses, including the EVERFSM6 variable for identifying disadvantaged pupils, in accordance with EEF statistical analysis guidance. Pupils in these subgroups are expected to respond better to REACH. The trial has not been powered for these subgroups but the analysis will be conducted on an exploratory basis.

Implementation and process evaluation: REACH Primary Programme Theory of Change and evidence-informed logic model

REACH Primary Evidence-informed Logic Model frame (to be read in conjunction with the supporting information)

Outputs	Intermediate Outcomes	Longer term outcomes

Change process 1

Teaching Assistant training in REACH Primary intervention, supported by SLT training

Trained TAs confident and knowledgeable in delivery of REACH intervention and assessment processes

Top up training and on-going support specifically linked to tailoring and graded progression

TAs
Increased confidence in delivery of REACH Intervention

Tailoring intervention to meet pupil need including baseline assessments

Fine graded progression

Input to Change process 2

TAs
Increased repertoire of pedagogical approaches and confidence to support poor readers

Increased knowledge of processes involved in reading and language development

These will be embedded and sustained in the longer term

Change process 2

Delivery of REACH Primary intervention to Y3 weaker readers

Pupils
Engagement in reading intervention

Awareness of success

Pupils
Improved reading comprehension

Improved phonological decoding, word recognition and listening comprehension

Increased attainment in reading

Pupils
Improved GCSE attainment

Improved KS2 SATs attainment

Improved attitude and motivation towards reading

Contextual factors:
Change process 1 - TA fidelity to REACH programme; accessibility and use of intervention resources; space and time of intervention sessions; volunteer/conscript, prior experience of TA including existing knowledge/experience of approaches to teaching reading; support from senior leadership and class teachers; established or concurrent approaches and practices in the teaching of reading across school/academy trusts; alignment with the current statutory guidance (National Curriculum)
Change process 2 - Selection of intervention pupils, pupil attendance at intervention sessions, suitability of intervention to pupil needs and interests; engagement of pupils during intervention sessions; other pupil characteristics [e.g. SEND]

Complexity: note complex reinforcing feedback loops between any of the outcomes and complex interactions between contextual factors and inputs, outputs, outcomes and impacts

Vision and impact of REACH Primary[2]

For Year 3 pupils:

- Improved reading comprehension
- Improved decoding, word recognition, and language comprehension

For Teaching Assistants (TAs):

- Increased repertoire of pedagogical approaches to supporting poor readers
- Improved confidence in supporting poor readers
- Increased knowledge of skills and processes involved in reading

The evidence informed logic model

The evidence-informed logic model presented above consists of the following components:

- Inputs (the support and training provided in this case)
- Causal process/es (consisting of an implementation pathway underpinned by a causal mechanism, with a set of related medium and shorter term outcomes)
- Contextual factors

We address these in turn below.

Inputs

Training and support provided to teaching assistants, involving five components.

Delivery of intervention to pupils over twenty weeks (subdivided into two ten week blocks).

Training and support provided to teaching assistants

Each school allocated to the intervention group will need to nominate two TAs to deliver the REACH Primary programme. One member of the School Leadership Team is also expected to attend the first half of training day 1. TAs will work with five students each and will be required to complete a training course. The following diagram provides an overview of the training course, which has five components:

1. **Initial training (pre intervention):** TAs will take part in a two day face-to-face training course hosted at a participating school, as well as having a programme manual and access to online resources
2. **Online seminars:** Five 30 minute online seminars will take place during the course of the project to check in and update the delivery team on intervention progress, troubleshoot, and to provide feedback and guidance on the training tasks
3. **Training tasks:** A series of tasks will be set for TAs and TAs will be encouraged to work in pairs to complete the tasks
4. **Top up training (mid intervention):** TAs will take part in one day face-to-face top up training, hosted at a participating school
5. **Support:** Telephone, email and online support for TAs throughout. TAs will have the option to submit recorded intervention sessions (either audio or video) to the research team for feedback and able to request onsite bespoke training where necessary

---

[2]This section, and the IPE design in general, is framed by Coldwell and Maxwell's (2018) research on evidence-informed logic models

REACH Primary intervention delivered to pupils

The intervention is delivered over twenty weeks through 3 x 30 minute sessions of individual support per week. Two sessions per week will focus on word recognition/ decoding skills and reading aloud fluently and accurately. The third session targets language comprehension skills.

Change processes and causal mechanisms

In defining a theory of change we draw on the work of Coldwell and Maxwell (2018, p269), as follows:

> The use of the term "theory" in this tradition is akin to Merton's 'middle-range' theories: those "that lie between the minor but necessary working hypotheses … and the all-inclusive systematic efforts to develop a unified theory" (Merton, 1968, p.39): it is a description of how an intervention leads to change. Weiss (1998, p.57) describes programme theory thus: "the mechanisms that mediate between the delivery (and receipt) of the program and the emergence of the outcomes of interest".

In Figure 1 REACH Primary Evidence-Informed Logic Model (below), we present a logic model that illustrates the 'theory of change' and interrelated aspects of impact for both pupils and TAs, highlighting contextual factors that may influence the impact of the project at each stage. In REACH Primary there are two distinct but interrelated change processes - the first focused on teaching assistant change and the second on pupil change.

Change process 1: Interventions delivered by Teaching Assistants

The causal theory draws on evidence around effective deployment of TAs (as summarised in the EEF TA Deployment Guidance - see Sharples, Webster and Blatchford, 2014); and evidence on effective CPD in schools (see for example Cordingley et al, 2015).

Sharples et al. (2014) identified a number of factors that impact on the effective deployment of TAs. These involved three fields of work in which TAs are commonly involved; everyday classroom conditions, the delivery of structured interventions, and the close integration of learning from work led by both teachers and TAs. The main findings of their report suggest that in everyday classroom conditions, effective deployment of TAs involves being well-prepared for their role in the classroom; adding value to the work of the teacher (and not replacing them) and using talk effectively to promote pupil independence in learning. Finally, TAs are seen as being best deployed in ways that do not always involve supporting low-attaining pupils or separating them from the classroom, teacher and peers. However, TAs working to deliver high quality one-to-one small group work, delivering structured interventions, is seen as beneficial. Such interventions are most effective where these are evidence-based, brief and regularly timetabled over a sustained period, guided by extensive training by well-qualified trainers, and where TAs maintain fidelity to the intervention plan and structure. Monitoring and tracking of pupil progress and assessment which informs subsequent intervention sessions is seen as beneficial. Although not part of the REACH Primary model, Sharples et al's final recommendation is that links between in-class learning and interventions, facilitated by liaison with class teacher, should be made to ensure that interventions are consistent with and extend classroom work and that these links are made explicit to pupils.

Pupil reading skills

It is intended that, in addition to impact on reading comprehension, REACH Primary may lead to improved attitudes towards reading, improved motivation to learn from text and improved KS2 statutory test and later GCSE performance.

The theory guiding REACH Primary is in the first instance guided by the two components of the Simple View of Reading, word recognition (decoding) and language comprehension. A full description of this theory is presented in the 'Study rationale and background' section above.

Pupil engagement in reading interventions

Please see section on motivations (p5, above).

Contextual factors

Successful delivery of the intervention may depend on a range of contextual factors. From our extensive experience of evaluating school-based interventions, we would expect at least four relevant elements:

- School-level factors (senior leader support; alignment with other school priorities *hypothesised to be important in engagement in the programme and training; and in embedding longer term outcomes*)
- TA-level factors (motivation and skills of TAs *hypothesised to be important in engagement in training, and delivery of programme in schools*)
- Programme factors (such as potential variation of quality of training, fidelity to programme *hypothesised to be important in TA delivery*)
- Wider system factors (such as alignment with other policies around reading *hypothesised to be important in engagement in the programme, in delivery of the programme in schools and in embedding longer term outcomes*)

Research questions and data collection

a) Which particular elements of the programme are perceived to support reading comprehension?

b) To what extent is the deployment of Teaching Assistants an effective approach to leading the intervention?

c) In what ways do Teaching Assistants tailor support for pupils that are in line with programme recommendations, and what aspects of the training or resources support this personalisation?

d) What fidelity issues are observed during the trial (including any aspects of intervention delivery not in line with programme recommendations)?

e) Is there evidence that the programme has a positive impact on pupil motivation to read?

f) What does the trial indicate about scaleability?

g) What is the cost per pupil of the intervention?

h) What are the 'business as usual' practices in control schools, and do they change over the course of the programme?

The IPE design will be based on the agreed evidence-informed logic model and will comprise the following:

1. **Further evidence review and early discussion with stakeholders** to build an agreed evidence-informed logic model, and data collection methods. *To address all RQs*

2. **School lead survey** in both control schools (in relation to business as usual) and intervention schools to examine understanding and practices around reading support in schools, including support provided by TAs. Some cost data will be gathered here. *To address RQs a,b,c,d,e, g and h*

3. **TA survey** in intervention schools to examine contextual issues raised in the school visits (and drawn from the delivery team and from the efficacy trial evaluation) that have been experienced as aiding or hindering effective implementation. *To address RQs a,b,c,d,e and g*

4. **Observations of one initial training event and one top-up training event, across three regions,** conducted by experienced literacy specialists, focussed on checking engagement, alignment with expected content, and process. Our previous pilot evaluation of RETAIN for EEF developed a tool to check the focus, content and delivery of CPD against our best current evidence on effective CPD (Maxwell et al, 2018), which could be adapted for this evaluation (see Appendix). Secondary analysis would examine the content of the CPD training and associated materials, to aid with judgment of observation, and also examination of records of attendance at training events and other engagement with support (to assist with judging fidelity). *To address RQs a, d and f*

5. School visits to understand fidelity, to ascertain influences on implementation and, in particular, the extent to which TA practice is aligned with expectations. Our team of literacy experts will conduct 10 school visits during which they will:

- observe at least one TA/pupil session
- interview the TA and school leader responsible for the programme at school level
- gather relevant secondary data such as schemes of work, cost data
- gather data indicating TA attendance at training sessions, completion of training tasks

The focus here will be on ascertaining the extent to which practices align with expectations, contextual factors that influence whether and how the programme works as expected, and intermediate outcomes at TA and pupil level in situ. These observations will inform the post-test TA survey tool. *To address RQs a, b, c, e, f and g*

6. **TA ongoing assessment of pupils' progress** based on the weekly '*running record of an instructional book*' which is usual practice during the intervention that allows for TAs to monitor the accuracy of pupil reading. This will be expanded to included data on pupil motivation. *To address RQs b,c,d,and e*

Collectively, the lines of enquiry detailed below will provide a detailed account of the contextual factors impacting on the REACH Primary intervention.

Table 3: Mapping research questions to IPE methods

Method	RQa	RQb	RQc	RQd	RQe	RQf	RQg	RQh
1. Further evidence review and early discussion with stakeholders	✓	✓	✓	✓	✓	✓	✓	
2. Literacy lead survey	✓	✓	✓	✓	✓		✓	✓
3. TA survey	✓	✓	✓	✓	✓		✓	
4. Training observations of three of the initial training events, and three of the top-up training events	✓			✓		✓		
5. School visits	✓	✓	✓		✓	✓	✓	
6. TA ongoing assessment of pupils' progress		✓	✓	✓	✓			

Table 4: Mapping IPE Methods to elements of the logic model

Method	inputs	change process 1	change process 2	contextual variation
1. Further evidence review and early discussion with stakeholders		✓	✓	
2. Literacy lead survey		✓	✓	✓
3. TA survey		✓	✓	✓
4. Training observations of three of the initial training events, and three of the top-up training events	✓	✓		✓
5. School visits			✓	✓
6. TA ongoing assessment of pupils' progress			✓	✓

Table 5: Mapping dimensions of implementation onto IPE methods

Dimensions of implementation	Source
1. Fidelity/adherence – the extent to which implementers (TAs) adhere to the intended treatment model *(including fidelity to the intervention content and process)*	3. TA Survey 5. School Visits
2. Dosage – how much of the intended intervention has been delivered and/or received *(including pupil attendance at intervention sessions)*	3. TA Survey 5. School Visits
3. Quality – how well different components of an intervention are delivered	3. TA Surveys 5. School Visits
4. Reach – the rate and scope of participation	2. Literacy Lead Survey 5. School visits
5. Responsiveness – the degree to which participants engage with the intervention	2. Literacy Lead Survey 3. TA Survey 4 Observation of training events 5. TA ongoing assessments
6. Programme differentiation – the extent to which intervention activities can be distinguished from other, existing practice	2. Literacy lead Survey 3. TA Survey
7. Monitoring of control/comparison groups – determination of the 'counter- factual' (taking place in the absence of the intervention)	2. Literacy lead survey
8. Adaptation – the nature and extent of changes made to the intervention	3. TA Survey 5. School Visits 6. TA ongoing assessments

5.3 Fidelity and dosage

Compliance will be measured at the TA level, through TA training. Dosage will be measured at the pupil level, through intervention delivery. Full details are provided in Table 6. The TA compliance and pupil dosage measures will be combined to create overall minimal and optimal compliance indicators at the pupil level.

Compliance with the TA training will be assessed according to three criteria that were developed in collaboration with the developers at the University of Leeds:

Criterion 1: The intervention entails three face-to-face training days, and attendance will be used as a compliance measure. If a TA does miss a session, compliance can only be obtained through a visit by a member of the delivery team[1]. Where this does not happen,

---

[1] All schools who missed a training session were offered a visit.

participating TAs are encouraged to seek input from colleagues who have attended the training but this is not deemed an adequate substitute for personally attending and will not count as attendance for the purpose of calculating compliance. Completion of REACH Primary training is the sole criterion for defining minimal compliance for TAs.

Criterion 2: Attendance of four online seminars (5 in total). Five online seminars for participating TAs are offered during the intervention; at least four of these must be attended, as indicated by seminar attendance lists and video watching statistics, for optimal compliance.

Criterion 3: A minimum of five of the eight gap tasks must be completed for optimal compliance. TAs must complete gap tasks within 14 days of the expected completion date. To verify completion, a combination of statistics on who has viewed the online material and those who have answered a short 'quiz' at the end of each session is used to verify completion.

The three TA level criteria will be drawn together to construct two TA level binary variables; the first defining minimal compliance (1=completing REACH training; 0=not completing) and the second defining optimal compliance (1=completing REACH training, online seminars and gap tasks).

Dosage, assessed at pupil level, will be measured by the number of intervention sessions delivered. REACH Primary comprises two distinct components: Reading (38 sessions) and the Comprehension (19 sessions). The total pupil level dosage will be calculated using equation 1.1:

Equation 1.1: $$\text{Dosage}_{\text{pupil}} = \frac{\left[\frac{\text{Reading Sessions}}{38} + \frac{\text{Comprehension Sessions}}{19}\right]}{2}$$

Pupils must complete 34 Reading Intervention sessions, and 17 Comprehension sessions (in addition to the sessions scheduled for the first week of the intervention) to be considered compliant. This will enable the construction of a single binary pupil level measure of compliance (1=completing 89%+ Reading AND 89%+ Comprehension REACH sessions). Minimal compliance is completing the equivalent of the first term's material, 47% of the RI and C sessions (nine Comprehension sessions and 18 Reading Intervention sessions).

The recommended time for each session is 30 minutes. Each consists of a structured activity schedule. It is therefore expected that TAs use the full amount of time allocated for the sessions so that it is possible to cover all of the necessary content. The evaluators are collecting data from TAs on the duration of each session. Those lasting for less than 25 minutes will not be counted as complete, as this is considered by the developers to be the minimum length of time required to cover all constituent elements of the sessions.

Table 6: Compliance indicators, TA and pupil level.

Activity	Description	Data source	Measurement (minimal)	Measurement (optimal)
TA Training (TA level)				
TA Face to Face Training	Content covered through attending sessions or a visit from a researcher	Attendance list	All three sessions attended	All three sessions attended
TA Gap Tasks	Completed within 14 days of expected date	Short test at the end proves completion	N/A	Five of the eight tasks must be completed
Online Seminars	Sessions attended or video summaries watched after the event	Attendance lists in online seminar; video watching statistics	N/A	Four of the five sessions must be completed
Intervention delivery (pupil level)				
Reading Intervention	Pupil completes weekly session, which lasts at least 25 minutes	TA log	18 of 38 sessions completed, plus week 1	34 of 38 sessions completed, plus week 1
Comprehension Sessions	Four core activities completed per session	TA log	9 of 19 sessions must be completed	17 of 19 sessions must be completed

The TA and pupil level dimensions of compliance will be drawn together to create the overall minimum and optimal compliance measures as follows:

TA minimal compliance (0 or 1) * pupil minimal dose (0 or 1) = Minimal compliance (0 or 1)
TA optimal compliance (0 or 1) * pupil optimal dose (0 or 1) = Optimal compliance (0 or 1)

These variables will be used to estimate the Complier Average Causal Effect (CACE). The purpose of the Complier Average Causal Effect (CACE) analysis is to estimate the impact of REACH for pupils deemed to have 'complied' with the intervention.

CACE will be estimated using two stage least squares (2SLS) regression (Gerber & Green, 2012). The first stage will model the pupil-level compliance variables (first minimal then optimal) using the same explanatory variables listed in Table 3 for the headline ITT analyses along with additional school level

items that are available via the school census as included in Table 4. This will be a multilevel logistic regression model used to generate predicted minimal/optimal compliance (1 or 0) for use in the second stage model. The second stage models will use predicted compliance in place of the group identifier variable in the ITT analyses specified above to generate the CACE estimates for REACH Primary.
Two CACE estimates will be calculated:

- First, using the predicted minimal compliance variable; 1=pupils who attended 47%+ Reading and 47%+ Comprehension sessions led by a TA who completed the REACH training; 0=control pupils plus pupils in intervention schools who attended <47% Reading and <47% Comprehension sessions OR the TA did not complete REACH training.
- Second, using the predicted optimal compliance variable; 1=pupils who attended 89%+ Reading and 89%+ Comprehension sessions led by a TA who completed the REACH training, 4+ online seminars and gap tasks; 0=control pupils plus pupils in intervention schools who attended <89% Reading and <89% Comprehension sessions OR the TA did not complete training / attend online seminars / complete gap tasks.

Please note that the specified measure of compliance is at the TA level but the final / overall measure of compliance is at the pupil level. Therefore the same approach for obtaining the CACE estimate for the specified 4-level MSCRT will be used if the TA level is not viable and the design becomes a 3-level MSCRT.

Cost evaluation

This will be addressed as follows:

Direct marginal costs:

- Information on fees, trainer pay, venue hire (from EEF and delivery partner)
- Any additional paid TA time (via literacy lead survey)

Time:

- Time spent on intervention (TA post-intervention survey)
- Cover costs for training attendees (TA post-intervention survey)

Prerequisites:

- Space for training (EEF/developers)
- Space for one-to-one reading support sessions in school
- Book box (£300 per intervention school)

Cost calculations based on this data will be conducted in line with EEF cost guidance.

Ethics and registration

- The University of Leeds and SHU have obtained full ethical approval through their respective university ethics committees.

- Schools are responsible for selecting pupils who are deemed weak readers and could benefit from additional support. Parents/carers are then sent an information sheet containing details of the project. The Memorandum of Understanding, to be signed by all schools to confirm their participation, includes a condition that schools must ensure that parents are informed that their child has been identified as in need of extra reading support.

- The trial is publicly registered: http://www.isrctn.com/ISRCTN15145068

Data protection

SHU and the University of Leeds will act jointly as data controllers for the duration of the project. The EEF will be data controller once the data have been submitted to its data archive. GL Assessment will be data processors for the purpose of marking outcome test papers. A data sharing agreement will detail the personal data to be shared, and a Fair Processing Notice will be sent to all participating schools as per GDPR requirements.

For the purpose of research, pupil data supplied to the evaluation team by schools will be linked with information about the pupils from the National Pupil Database (NPD) and shared between SHU, the University of Leeds, the Department for Education, the EEF, EEF's data contractor for their archive and, in an anonymised form, with other research teams and potentially the UK Data Archive. Further matching to NPD data may take place during subsequent research.

For transparency, the precise terms of this data sharing will be stated in a Fair Processing Notice, specifying the personal data to be processed (pupil names, pupil ID numbers, FSM

status, EAL status, SEN status), in line with General Data Protection Regulation (GDPR) guidelines that came into force from 25 May 2018. For information on how SHU treats data from research participants, please see:
https://www.shu.ac.uk/about-this-website/privacy-policy/privacy-notices/privacy-notice-for-research

The University of Leeds and SHU will strictly comply with current legislation, including the GDPR. Under GDPR Article 6, Paragraph 1e, the legal basis for this project is it being a 'public task', as the research is being conducted to evaluate the impact of an approach to reading support that has potential benefits for pupils participating in the trial and beyond. Therefore, parents/carers are free to withdraw their child from data collection and analysis at any time. Information on this will be provided for schools and parents/carers. If a parent/carer decides to withdraw, this would mean that no data on their child would be included in the evaluation and the child would not be required to take the outcome assessments in summer 2020. Updated information letters for parents/carers were issued in September 2020 to describe the changes to the trial due to Covid-19, including the opportunity to withdraw from outcome testing.

Special category data (EAL, SEN and FSM status, accessed from the NPD) is to be collected and processed for the purpose of scientific research as permitted under GDPR Article 9 (j). Specifically, we are processing this data to determine if the REACH Primary programme has different effects on different subgroups of pupils. The EEF was established with a remit to break the link between family background and educational attainment, and all EEF projects conduct subgroup analysis on FSM pupils. We are interested in the effects on EAL and SEN pupils as we believe that the project could be particularly beneficial to them.

The Fair Processing Notice and Project Information Sheet contain further details and can be found on the SHU web page for REACH Primary:

https://www.shu.ac.uk/about-us/academic-departments/institute-of-education/research/projects/reach-primary

Personnel

Professor Mike Coldwell: Project director - overarching responsibility for the successful completion of the project. Mike is a very experienced director of multiple method evaluations, with specific interest in evaluation methodology. He currently directs the EEF Integrating English Trial, and previously co-directed the South and West Yorkshire TA Scale -Up Campaign evaluation.

Dr Karen Daniels: IPE and Literacy lead - responsible for the successful completion of all aspects of the IPE. Karen leads developments in English across Primary and Early Years Initial Teacher Education routes. Her research interests include early literacy pedagogy and the social, emotional and cognitive dimensions of reading. She has worked with clients including Booktrust, Learning School Alliances and Academy trusts. Karen will be supported by colleagues in the Primary English team on the observation and school visit elements of the IPE.

Sean Demack: Statistical lead - responsible for the successful completion of the statistical design and analysis. Sean is a Principal Research Fellow in CDARE and has extensive experience in quantitative research including the design and analyses of educational RCTs.

Sean was co-director on the EEF ScratchMaths trial and currently is directing the EEF Realistic Maths trial.

Dr Josephine Booth: Project and Trial Manager (shared) - day to oversight of all elements of the project including testing management. Jo is a highly experienced project manager, performing this role on many projects over the past eight years including the three year, €3.6m EU-funded Chain Reaction science education study involving partners from 12 countries.

Dr Martin Culliney: Project and Trial Manager (shared) - and trial statistician, conducting quantitative analysis primarily on the trial strand. Martin has worked on projects commissioned by government and third sector funders such as DfE, DEFRA, BIS and ACAS, leading on survey design, data management and analysis, including the current Integrating English EEF evaluation.

Dr Paula Clarke: Intervention developer and project manager of the delivery team. Paula will lead the recruitment of schools, prepare the TA training and intervention materials and deliver all of the face to face training. Paula is an Associate Professor and has over 10 years' experience developing and delivering school based interventions to support the development of reading comprehension skills.

Dr Shirley-Anne Paul: Project Co-Investigator - will support the development work on the intervention materials and training, will lead on the setting up of the project online elements, and advise the PDRA on project management strategies for collecting secondary outcome data across 80 school sites. Shirley coordinated the EEF REACH trial and has systematically reviewed reading interventions for secondary school pupils. She is a qualified Project Manager and is a Trial Coordinator at York Trials Unit.

Risks

Risk	Solutions	Perceived level of risk
Problems with compliance or recruitment	Incentives should ensure participation, previous trials have had good retention rates	Low
TAs do not complete full training	Training course is only three days. Fidelity will be measured through developer attendance records	Low
TAs do not apply REACH principles in one-to-one sessions	Fidelity will be measured during process evaluation through TA surveys and school visits	Low
Problems with NPD access	SHU highly experienced with the application process, NPD only used for baseline here, delays very unlikely to affect main report deadline	Low
Staff departures	SHU has very low staff turnover. Experienced statisticians and specialists in literacy available if necessary	Low
	This is a higher risk for UoL given the central role of the project lead	Moderate
Difficulties administering outcome tests	Pilot testing to identify any issues. Both SHU and UoL have detailed plans to recruit and train staff for testing	Moderate

Timeline

Dates	Activity	Staff responsible/ leading
Jan 19	MoU and parental info sheet to Leeds/EEF	SHU
Jan 19	Ethics submission	SHU
Jan 19	Project info sheet to SHU/EEF	Leeds
Feb 19	Evidence review	SHU
Mar 19	Protocol first draft	SHU
Apr 19	Data collection tools (literacy lead and teacher)	SHU
Jan-May 19	Revise intervention, produce resources	Leeds
Feb-June 19	Recruit schools, identify TAs	Leeds (with SHU support)
May 19	Baseline testing (KS1)	Schools
Jun- Jul 19	Collect baseline data from schools	SHU
Jun- Jul 19	Pre-intervention survey (literacy lead)	SHU
w/c 15 Jul 19	Randomisation, schools informed of allocation	SHU
w/c 16 Sep 19	TA training	Leeds
Sep 19	TA training observations x 6 (inc. day 3, scheduled for Spring 20)	SHU
Sep 19	Pre-intervention survey (TAs in intervention schools)	SHU
Oct 19	SAP/trial registration	SHU
Oct 19-May 20	Intervention delivery in schools	Leeds
Oct 19-May 20	Case studies in 10 schools	SHU
Feb-May 20	Post-intervention surveys (TAs in intervention schools and literacy leads in all schools)	SHU
Nov-Dec 20 (Delayed from Summer 20)	Primary outcome assessment (WIAT III)	SHU
Nov-Dec 20 (Delayed from Summer 20)	Secondary outcome assessments (CELF, DTWRP)	Leeds
Mar 21 (Delayed from Oct 20)	Report first draft	SHU

References

Bloom, H., Richburg-Hayes, L. and Rebeck Black, A. (2007) 'Using covariates to improve precision for studies that randomise schools to evaluate educational interventions' *Education Evaluation and Policy Analysis* 29(1) pp30-59

Clark, C. & Teravainen, A. (2017) *What it means to be a reader at age 11: valuing skills, affective components and behavioral processes: An outline of the evidence*. London: National Literacy Trust

Clark, C. & Rumbold, K. (2016) Reading for Pleasure. A Research Report. National Literacy Trust.

Clarke, P., Snowling, M., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science*, 21(8), pp1106–1116.

Clarke, P. J., Truelove, E., Hulme, C., & Snowling, M. J. (2013). Developing reading comprehension. Chichester, UK: Wiley-Blackwell.

Clarke, P. Paul, S.A, Smith, G., Snowling, J. & Hulme, C. (2017). *Reading Intervention for Poor Readers at the Transition to Secondary School*, Scientific Studies of Reading, 21:5, pp408-427.

Coldwell, M. and Maxwell, B. (2018) Using evidence-informed logic models to bridge methods in educational evaluation. *Review of Education*, 6 (3), pp 267-300.

Cordingley, P., Higgins, S., Greany, T., Buckler, N., Coles-Jordan, D., Crisp, B., & Coe, R. (2015). *Developing great teaching: lessons from the international reviews into effective professional development.* London: Teacher Development Trust. Available at: https://tdtrust.org/wp-content/uploads/2015/10/DGT-Full-report.pdf

Department for Education (2015) *Reading: the next steps. Supporting higher standards in schools encouraging children to read widely and for pleasure: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/409409/Reading_the_next_steps.pdf*

Demack, S. (2019) Does the classroom level matter in the design of educational trials? EEF Research Paper No. 3 available at https://educationendowmentfoundation.org.uk/public/files/Publications/Does_the_classroom_level_matter.pdf

Education Endowment Foundation (2016) *REACH Evaluation report and executive summary.* Available at: https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/reach/

Gough, P. and Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*, p6–10.

Guthrie, J. T. and Wigfield, A. (2000). Engagement and motivation in reading. In M.L. Kamil, P.B. Mosenthal, P. D. Pearson, and R. Barr (Eds.), *Handbook of Reading Research* (3rd Ed.). New York: Longman.

Guthrie, J. T., Wigfield, A. & Wei You (2007) Instructional Contexts for Engagement and Motivation in Reading. Contemporary, *Education Psychology,* 32, pp. 282-313

Hatcher, P., Duff, F., & Hulme, C. Sound Linkage: An Integrated Programme for Overcoming Reading Difficulties, 3rd Edition. London: Wiley Blackwell.

Hempel-Jorgensen, A., Cremin, T., Harris, D. & Chamberlain, L. (2018). Pedagogy for reading for pleasure in low socio-economic primary schools: beyond 'pedagogy of poverty'? *Literacy,* 52(5), pp 86-94.

Hoover, A. and Gough, P. (1990) The Simple View of Reading. *Reading and Writing,* 2 (2), pp127-160.

Merton, R. K. (1968) *Social theory and social structure*. London: Simon and Schuster.

Morgan, P. and Fuchs, D. (2007) Is there a bidirectional relationship between children's reading skills and reading motivation? *Exceptional children*, 73.2, pp. 165-183.

Ng, C. and Graham, S (2018) Improving literacy engagement: enablers, challenges and catering for students from disadvantaged backgrounds. *Journal of Research in Reading*, 41.4, pp. 615-624.

OECD (2010) PISA 2009 Results: Learning to Learn- Student Engagement, Strategies and Practices (Volume 111) Paris: OECD.

OECD (2016) PISA 2018 Draft analytical frameworks. Available at: https://www.oecd.org/pisa/data/PISA-2018-draft-frameworks.pdf.

Sharples, J., Webster, R., & Blatchford, P. (2014). Making best use of teaching assistants: Guidance report. London: Education Endowment Foundation. Available at: https://educationendowmentfoundation.org.uk/public/files/Publications/Campaigns/TA_Guidance_Report_MakingBestUseOfTeachingAssisstants.pdf.

Weiss, C. H. (1997) Theory-based evaluation: Past, present, and future. *New directions for evaluation*, 76, pp. 41-55.

Appendix A: REACH Power Analysis

From Spybrook et al. (2016), the MDES equation for a 4-level MSCRT assuming zero effect size variability across clusters but also including covariate explanatory power at class (TA) and pupil levels is:

$$MDES_{4LMSCRT} \sim M_{(M(K-L-2))} \sqrt{\frac{1}{P(1-P)}} \sqrt{\frac{ICC_{sch}(1-R_{sch}^2)}{MK} + \frac{ICC_{TA}(1-R_{TA}^2)}{MKJ} + \frac{(1-ICC_{sch}-ICC_{TA})(1-R_{pup}^2)}{MKJn}}$$

Where…

- P is the proportion of schools who receive the intervention (=0.50)
- R_{Sch}^2 is the school-level covariate explanatory power (=0.60^2= 0.36)
- R_{TA}^2 is the TA-level covariate explanatory power (=0.36)
- R_{pup}^2 is the pupil-level covariate explanatory power (=0.55)
- ICC_{Sch} is the school level Intra Cluster Correlation coefficient (=0.14)
- ICC_{TA} is the class level Intra Cluster Correlation coefficient (=0.02)
- M is the number of geographical sites (=5)
- K is the number of schools per site (=16)
- J is the number of TA per school (=2)
- n is the number of pupils per TA (=5)
- L is the number of school level covariates (=6)[3]
- $_{((K-L-2))}$ is the t-distribution multiplier with M(K-L-2); 5(16-6-2); (40) degrees of freedom. Assuming a two-tailed test with a statistical significance of 0.05 (α/2=0.025) and statistical power of (1-β=0.80). M_{40}= 2.8718.

This results in an MDES estimate of 0.24 standard deviations. For the FSM analyses, the number of pupils per TA is reduced to two and if all other factors are assumed to be the same as above, the FSM MDES estimate is 0.29 standard deviations.

These estimates assume a 3-level CRT design that is blocked by geographical area. The viability of the TA level is unknown at this point. If it is found that the TA level is not viable, the design will become a 2-level CRT blocked by geographical area.

A 2-level multisite CRT design results in an MDES estimate of 0.23 standard deviations overall and 0.27 sds for the FSM subsample.

---

[3] The school level covariates will be: Intervention/control group membership; Geographical Hub (4 dummy variables) and the School level KS1 attainment used within the randomisation.

Appendix B: Mapping to indicative characteristics of effective CPD: frame from RETAIN evaluation

Critical features of effective CPD	Indicative characteristics of effective CPD	Findings from interviews, observations of session videos and professional review of resources
Content focus	Overview	☐
	Curriculum content that helps teachers understand how pupils learn, both generally and in specific subject areas: including subject-specific pedagogy and enables participants to access the theory and evidence underlying the relevant pedagogy, subject knowledge, and strategies.	☐
	A logical thread between the various components of the programme.	☐
	A focus on learner progression, starting points and next steps, including formative assessment, to enable teachers to see the impact of their learning and work on their pupil.	☐
	Content includes alternative pedagogies for pupils with different needs.	☐
	Content takes account of different teachers' starting points.	☐
Active learning	Overview	
	Opportunities are provided for teachers to reveal and discuss their beliefs activities and test ideas from different perspectives. This includes helping participants believe that better outcomes are possible, particularly among schools where achievement has been depressed over time.	☐
	Teachers are engaged in analysis of and reflection around the underpinning rationale for practices changes, and the supporting evidence.	☐

	Activities include explicit discussions, following the initial input, about how to translate CPD content to the classroom. This includes teachers making links between professional learning and pupil learning explicit through discussion of pupil progression and analysis of assessment data	☐
	Teachers implement what they have learned by experimenting in the classroom.	☐
	Specialists support teachers through modelling, providing observation and feedback, and coaching.	☐
Relevance	The content and activities have overt relevance to participants' day-to-day experiences with, and aspirations for, their pupils.	☐
Duration and rhythm	The duration (total time and the spread over time) is sufficient to lead to, and the 'rhythm' of follow-up, support and consolidation enables, changes in teachers' practices.	☐
Collaborative participation	Teachers engage in peer learning with colleagues attending the programme.	☐
	Teachers share and discuss learning with colleagues in their own school.	☐
	The design of collaboration participation leads to positive outcomes for teachers and minimises the negative outcomes that can be associated with collaborative activity.	☐
Shared sense of purpose about professional development	There is a shared sense of purpose about professional development between teachers and their schools.	☐
	Senior leaders in schools ensure that enabling mechanisms are in pace to support teachers in implementing what they have learned from the programme and share that learning with school colleagues.	☐