

Teacher Choices: The Story Time Trial Evaluation Protocol

Evaluator: National Foundation for Educational
Research (NFER)

Principal investigator: Dr Ben Styles



Template last updated: August 2019

PROJECT TITLE	The Story Time Trial
DEVELOPER (INSTITUTION)	Education Endowment Foundation (EEF)
EVALUATOR (INSTITUTION)	National Foundation for Educational Research (NFER)
PRINCIPAL INVESTIGATOR(S)	Dr Ben Styles
PROTOCOL AUTHOR(S)	Maria Galvis, Palak Roy and Jennie Harland
TRIAL DESIGN	Two-armed randomised controlled trial (RCT) with randomisation at the school level
TRIAL TYPE	Feasibility of RCT
PUPIL AGE RANGE AND KEY STAGE	Ages 8-10 (Years 4 and 5)
NUMBER OF SCHOOLS	100 schools
NUMBER OF PUPILS	Approximately 9,000 pupils (based on an average cohort size of 45 pupils and two year groups per school)
PRIMARY OUTCOME MEASURE AND SOURCE	Pupils' comprehension skills (as measured by reading and listening comprehension of the text selected for the intervention, NFER's bespoke assessment)
SECONDARY OUTCOME MEASURE AND SOURCE	Attainment Pupils' wider reading comprehension skills (as measured by the reading comprehension of an unseen text, NFER's bespoke assessment) Non-cognitive Pupils' engagement in reading lessons, confidence in reading, liking of reading measured via a pupil survey using the scales from the PIRLS ¹ (Progress in International Reading Literacy Study)

¹ For information about the instrument and psychometric properties of the three scales see: Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.). (2017). Methods and Procedures in PIRLS 2016. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/publications/pirls/2016-methods/P16_MP_Chap14_Context_Questionnaire_Scales.pdf

Protocol version history

VERSION	DATE	REASON FOR REVISION
1.0 [original]	5 th March 2021	N/A
2.0	06/05/22	The trial has been rescheduled twice. Under the first attempt, the Story time sessions were planned to take place in the Spring term 2020 (academic year 2019-20) and under the second attempt these were planned for the summer term of 2021 (academic year 2020-21). Due to Covid-related school closures in Spring terms of 2020 and 2021, it was jointly decided by NFER and EEF to delay the trial activities. This protocol version describes the story time sessions involving years 4 and 5 pupils and teachers in the autumn term 2021 (academic year 2021-22). The changes are described in the study rationale and background section.

Contents

Teacher Choices study rationale and background	4
Teacher Choices trial 2 ‘The Story Time trial’ rationale and background	10
Intervention	11
Feasibility of measurement and impact evaluation	17
Research questions	17
Design	17
Randomisation	18
Participants	18
Sample size calculations	18
Outcome measures	19
Compliance	21
Analysis	22
Implementation and process evaluation	23
IPE research focus	23
Research methods	23
Process analysis	25
Cost information	25
Ethics and registration	26
Data protection	26
Personnel	27
Risks	28

Timeline	30
References	32

Teacher Choices study rationale and background

'The Story Time' is the second trial in the Teacher Choices project that was scheduled to run during the summer term of the academic year 2019-2020. Due to the Covid-19 pandemic, the UK Government ordered all schools to close from 20th March 2020 to all pupils (except vulnerable children and the children of key workers). Consequently, the trial activities were paused. At this time, the schools had been recruited and randomly allocated to one of the two groups. However, schools had not seen the Teacher Guidance documents. Therefore, it was agreed between NFER and EEF that the trial will be delayed, additional schools will be recruited (if required) and randomised such that the Story Time sessions will be delivered in the first half of the summer term 2021 (April-May 2021). The recruitment activities commenced in the autumn term 2020. Further recruitment and data collection was planned for the spring term 2021. However, schools closed for in-person learning once again in the spring term 2021, hence the trial activities were paused again.

NFER and EEF decided to delay the trial activities so that the Story Time sessions will be delivered in the second half of the autumn term 2021 (November-December 2021). From June to September 2021, the schools that were recruited in 2020 will be re-invited to take part in the trial and additional schools will be recruited (if required).

Before the school closures in March 2020, in addition to recruiting schools to take part in the trial and collecting participants' baseline data, NFER had also completed the development of the end-point assessment. The teacher guidance was developed by EEF and consultation on the teacher guidance was completed by NFER researchers. The original trial included an additional strand of work, involving North Yorkshire Coast (NYC) Research School (RS) in supporting NFER with school recruitment, consultation on the teacher guidance document and providing feedback on the end-point assessment. The process evaluation of the trial intended to explore differences between schools recruited through the RS network and other schools. However, for the delayed trial, it has been decided that we will not involve a Research School and therefore, it will not be applicable to explore differences by recruitment method. In the original trial, NFER was collecting Key stage 1 (KS1) Age-related expectations (ARE) in English reading directly from schools. This was the baseline measure for the trial. However, this data is not available for the current Year 4 trial cohort (due to KS1 cancellation in May 2020). As a result, it was agreed that NFER will collect reading or English assessment data for those schools that used commercial assessments in the spring or summer term 2021.

Rationale for the evaluation

In 2018, EEF undertook a review of their grant-making process with the aim of understanding how to make their projects timelier, more relevant for schools and to ensure that the results were as useful as possible to schools and other decision-makers. Through this process, they identified that teachers and headteachers were keen to answer research questions which can more directly feed into existing teaching practice. These questions were often not related to the impact of manualised programmes, which require schools to purchase particular resources or training, but were instead about the everyday choices that teachers have to make when planning their lessons and deciding on the pedagogical approaches they take in their classrooms.

As a result of this review, EEF commissioned NFER to conduct two classroom-based trials to explore the feasibility of running this type of trials and to understand more about the best designs and methods to use, with the intention of EEF then commissioning more of these in the future. These classroom-based trials are called 'Teacher Choices trials'. These trials will explore the feasibility of running robust evaluations of teachers' everyday choices in the classroom. The first Teacher Choices trial is called the 'A Winning Start trial'. This protocol relates to the second trial, which is called 'The Story Time Trial'.

The key aims of the 'Teacher Choices' trials are to ascertain the feasibility of running such trials in schools and to provide recommendations for future scale-up. The research questions will focus specifically on how the selected everyday choices are implemented in the classroom, how the everyday choices can be defined, how the trial methodology is followed including if/how the choices can be tested in a robust way using teacher-tests and the most effective way to communicate the results to teachers to support classroom practice/improvement.

The Teacher Choices trials aim to answer the research questions centred on the feasibility of the research method as well as the implementation, process and impact evaluation of the teacher choices. The research questions addressed in the first Teacher Choices trial 'A Winning Start' trial (TC1) and its second-run 'A Winning Start trial (second run)' can be found in the relevant protocols (EEF, 2020).

The following list presents the research questions addressed by this trial- 'The Story Time Trial' or TC2:

1. What are the barriers and challenges in implementing a 'Teacher Choices' trial?
 - a. Is it possible to recruit schools to Teacher Choices trials?
 - b. Do schools understand the trial design and adhere to the requirements set out in the protocol, i.e. randomisation, data collection and testing?
 - c. How much information do teachers require in order to follow the evaluation methodology?
2. How is the choice implemented?
 - a. Is the level of teacher guidance on the 'choice' appropriate to interpret and implement the selected approach and to establish compliance with the intended 'choice'?
 - b. Do teachers adapt the approach to suit their context?
 - c. What is teachers' usual practice?
3.
 - a. Is it possible to derive a suitable baseline measure by combining the scores of different commercial age-standardised English or reading assessments?
 - b. Can a story or text-specific test be used to measure the impact of this teacher choice? How should the test be designed? What is the level of informal assessment trialling required to design a test that assesses text-specific and wider reading comprehension skills? What psychometric properties does the test need in order to be used as a valid and reliable outcome measure for future trials?
4. What can we learn from this trial about the expected effect sizes, school-level intra-cluster correlations, pre-post correlations and sample sizes for future teacher choice trials?
5.
 - a. Does a given teacher choice improve pupil attainment outcomes in text-aligned listening and reading comprehension skills, and in wider reading comprehension skills on an unseen text (as measured by a test designed by NFER)?
 - b. Does a given teacher choice improve pupils' engagement in reading lessons, confidence in reading and liking of reading (as measured via a pupil survey)?
6. Do we observe any evaluation-driven effects i.e. does the trial change the way schools and teachers behave?

7. How should we best report this work to ensure that the research is clearly represented, and does not over-claim, but also ensures accessibility for teachers?

Due to the type of research questions and the relative simplicity of the choices, the evaluation designs (for both Teacher Choices trials) will not follow standard EEF guidance during the trials. These designs will enable the trials to run over shorter timeframes to ensure that the research questions are answered in a timely way. During these trials, the EEF will take the lead in defining these everyday choices for teachers. NFER will design and run the evaluations to test these choices in a robust way. For this trial, the two theories of change focus on reading and listening comprehension skills as outcomes in different ways. NFER assessment experts will not only create the end-point assessment to measure outcomes but will also have a role in designing aspects of the intervention itself that rely on questioning.

Teacher Choices

It is envisaged that the everyday teacher choices being investigated as part of these trials and in the future would be of high interest to schools. They should involve a real choice that can be made by classroom teachers, be easy to define so that teachers know what they are expected to do as part of the project and relatively easy to implement without lots of training and resources. Teachers would receive a simple input, which could be easily replicated in the future for teachers wishing to implement the choice outside of the test condition.

Table 1 sets out how this trial will contribute to the overall aims of the teacher choices feasibility study.

Table 1: Teacher Choices study – research questions, methods and analyses for ‘The Story Time Trial’

Research Area	Research Question	Study methodology area	Data collection methods	Participants/data sources	Data analysis methods
Feasibility of implementing the RCT methodology for teacher choices evaluations	RQ1: What are the enabling and challenging factors in implementing a teacher choices trial? - Is it possible to recruit schools/teachers? - Do schools understand the trial design and adhere to its requirements (such as randomisation, data collection and testing)? - How much information do teachers require in order to follow the evaluation methodology?	IPE	Recruitment Teacher feedback survey Telephone and case study interviews	NFER recruitment and monitoring data Teachers N=200 Interviews with head/SLT and teachers (8 case studies; 8 telephone interviews)	Descriptive analysis Formative view
Implementation and process evaluation of the approach/choice	RQ2: How is the choice implemented? - Is the level of teacher guidance and support on the ‘choices’ appropriate to interpret and implement the selected approach and to establish ‘compliance’ with the intended ‘choice’? - Do teachers adapt the approach to suit their context? - What is their usual practice?	IPE	Consultation on teacher guidance Teacher feedback survey Telephone and case study interviews	Approx. 6 school staff Teachers N=200 Observations (subject to COVID-19 restrictions guidance) and interviews with head/SLT and teachers (8 case studies; 8 telephone interviews)	Thematic analysis Perspectives analysis Observational analysis Triangulation Formative view

Research Area	Research Question	Study methodology area	Data collection methods	Participants/data sources	Data analysis methods
Feasibility of testing / measurement	RQ3a: Is it possible to derive a suitable baseline measure by combining the scores of different commercial age-standardised English or reading assessments?	Feasibility of measuring impact	Commercial English or reading assessments	Pupil N= approx. 4,500 (data available for subset of schools)	Create a standardised common scale measurement of reading comprehension and/or English
	RQ3b: Can a story or text-specific test be used to measure the impact of this teacher choice? How should the test be designed? What is the level of informal assessment trialling required to design a text-specific test? What psychometric properties does the test need in order to be used as a valid outcome measure for future trials?		NFER's bespoke end-point assessment	Pupil N= approx. 9,000	Undertake reliability of the end-point assessment
Trial design	RQ4: What can we learn from the trials about the expected effect sizes, intra-cluster correlations, pre-post correlations and sample sizes for future teacher choice trials?	Measurement, sample sizes	NFER's bespoke end-point assessment	Pupil N=approx. 9,000	Analyse the pre-post outcome measures from the end-assessment
Impact	RQ5a. Does a given teacher choice improve pupil attainment outcomes in text-aligned listening and reading comprehension skills, and in wider reading comprehension skills on an unseen text (as measured by a test designed by NFER)?	Impact	NFER's bespoke end-point assessment	Pupil N=approx. 9,000	Analyse the pre-post outcome measures from the end-assessment
	RQ5b. Does a given teacher choice improve pupils' engagement in reading lessons, confidence in reading and liking of reading (as measured via a pupil survey)?				

Research Area	Research Question	Study methodology area	Data collection methods	Participants/data sources	Data analysis methods
Feasibility of RCT methodology	RQ6: Do we observe any evaluation-driven effects i.e. does the evaluation (or the trial) itself change the way schools and teachers behave?	IPE	Telephone and case study interviews	Observations (subject to COVID-19 restrictions guidance) and interviews with head/SLT and teachers (8 case studies; 8 telephone interviews)	Perspectives analysis Observational analysis
Appropriateness of reporting	RQ7: How should we best report this work to ensure that the research is clearly represented, and does not over-claim, but also ensures accessibility for teachers?	IPE	Focus group	Up to 8 teachers (from primary and secondary schools; trial and non-trial teachers ²)	Formative view

² The focus groups will take place in spring 2022 and findings from these will be reported separately from the Teacher Choices trials.

Teacher Choices trial 2 'The Story Time trial' rationale and background

Background and process of selection

This protocol relates to the second classroom-based Teacher Choices trial: 'The story time trial'. It aims to evaluate two popular ways that primary teachers read aloud to the whole classroom: sustained reading (GO! approach) and reading with pauses to ask questions (STOP! approach).

To select the second Teacher Choices trial, EEF piloted a process where a number of small seminar sessions were delivered with discussion groups of teachers to elicit questions of interest. In addition, members from the EEF Programmes team suggested areas of interest and possible questions. From this process, EEF along with the participating teachers identified six testable questions that were considered to be testable questions. Using an online survey called the 'Teacher Tapp', EEF asked primary school teachers to identify the question they would most like to know the answer to. The teacher responses were almost unanimous across all year groups, which confirmed that most primary teachers were interested in knowing which reading aloud approach was most effective- continuous or interrupted.

Rationale

Reading aloud to children is considered a regular part of classroom practice. In most primary schools, the class teacher will share stories by reading aloud to their children regularly (Cremin et al, 2014). In some classes, teachers will read the text, without stopping, immersing the children in the story and seeking to share their pleasure in the text. In others, teachers will stop regularly asking a series of questions designed to prompt the children to think about what they are hearing and scaffold the metacognitive processes (Oakhill et al, 2014).

Reading the story without pausing is a popular way of reading aloud that builds on an established tradition of sharing stories, exposing children to a wide range of texts and seeking to trigger pupils' enjoyment of, and motivation to, read. Some of the potential mechanisms identified to support learning through continuous reading include:

- a fluent reader could scaffold understanding of a challenging text due to the proper use of juncture, pitch and stress;
- being exposed to teachers' modelling of prosodic skill could aid both word reading and text comprehension;
- giving the pupils access to more challenging vocabulary (breadth) used in new contexts (depth);
- allowing pupils to follow the text and develop their own situational model;
- pupils being exposed to more text in the same amount of class time as they are not interrupting reading with questions

In contrast, there is evidence to suggest that reading the text and pausing to ask questions may help to model to children the implicit processes that accomplished readers undertake when reading, helping them to develop these metacognitive skills for themselves. Some of the potential mechanisms identified to support learning through reading with questioning include:

- drawing pupils' attention to salient features of the text that are important for understanding, helping them notice more helpful parts of the text
- asking questions to pupils could foster a deeper understanding and in turn help them learn from the text
- pupils could learn how to compose questions themselves from the teacher model and apply this (listening) to other contexts

This evaluation will run as a two-armed Randomised Control Trial (RCT) where participating primary schools will be randomly allocated to one of the two reading approaches (GO! or STOP!). All schools will receive a short implementation guidance and EEF's support emails and ad hoc calls (if requested). Year 4 and Year 5 teachers will use the allocated reading approach with their classes every day for 15 minutes, over a period of approximately three weeks. At the end of the intervention period, teachers will administer an end-point assessment and a pupil survey. The assessment will measure the primary outcome and the secondary outcome of attainment, the survey will measure the non-cognitive secondary outcomes for the trial.

The process evaluation seeks to understand how the teachers implement the choices, including views on the level and nature of guidance needed for compliance. It also aims to understand how the trial methodology itself is followed by teachers, including adherence to randomisation condition and trial requirements such as testing and returning data; and views on the amount and nature of 'trial information' provided. As Trial 2 is using defined questions/content and text-aligned assessment (rather than a range of content and tests as in Trial 1), it will also be possible to explore whether the mechanisms in each of the theories of change for the two choices being studied appear to be working.

Intervention

The following framework outlines what is comprised in 'The story time trial' evaluation. In addition, a TIDIER framework for each reading approach has been developed and agreed between EEF and NFER. These will be included in the final report.

Brief name

The story time trial

Why

The study aims to investigate the impact of two different approaches where classroom teachers read aloud to the whole class in primary schools: sustained reading (GO! approach) and reading with pauses to ask questions (STOP! approach). We aim to improve our understanding of the mechanisms that help children develop comprehension skills within a whole class scenario.

The rationale for continuous reading is that pupils develop love of reading by listening to a fluent reader, are engaged with more challenging texts than they could access themselves and are immersed in the story. On the other hand, reading and asking questions helps pupils develop a metacognitive approach to reading and models how an expert reader accesses a text. Different question types can be used to develop different areas of comprehension e.g. vocabulary, story structure, character, setting or identifying misconceptions.

Who (recipients)

Year 4 and Year 5 teachers are the recipients of the guidance and the implementers of the intervention. Year 4 and Year 5 pupils from the participating schools are the recipients of the reading aloud approach.

What

Materials

Year 4 and Year 5 teachers from the participating schools will be given a short guide with Do's and Don'ts to enable them to implement the approach that their school is randomised to (GO! or STOP!). In addition to the guidance, EEF will approach schools via email to offer support on implementation of each approach. EEF will also offer further emails or calls, if requested by schools.

NFER will provide the book that teachers should read as part of the intervention. A copy of the book *The Iron Woman* by Ted Hughes will be sent to each teacher along with the relevant guidance document. We will also send a set of questions/prompts to be used by the teachers applying the STOP! approach.

Procedures

The teachers will experience only one of the reading approaches over a period of approximately three weeks. During this period, teachers will use the randomly allocated approach in daily 15-minutes sessions, until they have read the whole book. In each session, teachers using the GO! approach will read aloud continuously for 15 minutes, while teachers using the STOP! approach will read aloud and pause to ask questions/prompts that will be provided. There is no stipulation on time spent on reading/questions but a suggestion is 8-10 minutes reading interspersed with 7-5 minutes of questions. If the teachers complete reading *The Iron Woman* to their class in less than three weeks, they can continue using their allocated approach by selecting and reading a different age-appropriate book.

Who (implementers)

EEF will develop the teacher guidance. Year 4 and Year 5 classroom teachers are the implementers of the reading aloud approach. Teaching Assistants (TA) and school leaders are not eligible to take part. Teachers may or may not have used the approach before.

At the end of the trial, and once they have completed all the data requirements, participating schools will receive a £100 token that can be used to expand the CPD library. This will be accompanied by a list of EEF recommended titles, although the choice of texts will be up to the each school.

How

After randomisation, we will send a copy of the book and a guidance booklet to teachers to have the opportunity to familiarise with the text and the approach. Teachers on the STOP! group will also receive a set of questions for discussion during the sessions. Important considerations for implementation include:

- Each daily session should be no longer than 15 minutes.
- Teachers must ensure that all the text is covered during the intervention period.
- Teachers in the GO! group may well complete the book within the 3-week window (as they are reading continuously). They may select another book of their choice

and continue reading in the same way for the remaining sessions. Teachers in the STOP! group can continue with another book of their choice should they complete the book within the 3-week window; they will need to prepare questions covering vocabulary, plot, characters etc.

- Teachers must choose a suitable and regular time of day for the reading aloud sessions, for example before or after break or lunch. The end of the day may not be suitable as children tend to be less focused.
- Teachers can choose the layout of the class for the session as long as pupils are sitting (on the carpet or their desks)
- It is important that pupils do not have any distractions for the 15 minutes that the teacher is reading to them.

Where

The intervention will be delivered in all Year 4 and Year 5 classrooms.

When and how much

The intervention will take place in the second half of the autumn term 2021 (November-December 2021).

The teacher should use the randomly allocated reading approach once a day, for approximately 15 minutes, during the intervention period (up to three weeks). Although there is not a determined number of sessions in which the intervention should be delivered, it is important that teachers keep the approach consistent for the whole of the specified book. This is expected to be around 2-3 weeks. If teachers complete the book within the 3-week window (which may well be the case for the GO! approach as they are reading continuously but may also occur with the STOP! approach) teachers can select a book of their choice and continue to use the approach for the remaining sessions.

Teachers on the STOP! group should ask between three and four prompts/questions spread out during the session. NFER will provide these questions. Although it is not stipulated how much time should be spent in reading/questions per session, the suggestion is to spend eight-ten minutes reading interspersed with five-seven minutes for questions.

Tailoring

Teachers must keep the reading time close to 15 minutes in each session. If the book is completed in fewer sessions, teachers can continue using the same approach with another age-appropriate book.

Teachers must read the guidance document for the allocated approach and make a credible attempt to comply with the intervention as described in the guidelines. Teachers using the GO! approach must avoid contamination by not asking questions to pupils.

Both reading aloud approaches are designed to be whole-class activities and should involve all pupils in the class (including pupils with special educational needs and disabilities (SEND) and English as an Additional Language (EAL)). Teachers on the STOP! group should use the question prompts provided.

How well

We will explore the extent and nature of fidelity at three levels: i) implementation fidelity (i.e. teachers' implementation of the approach itself), ii) implementation support (i.e. the provision

of the guidance materials and EEF's email support helpline for teachers), and iii) class-level dosage (i.e. pupils' exposure to the approach). These are outlined further below.

- i) We will explore the extent of **implementation fidelity** in terms of: whether teachers have read the guidance document, whether teachers have read the book and if allocated to STOP! have familiarised themselves with the prompts provided, have planned a suitable and regular time for delivery, and whether teachers have used their allocated approach daily (as far as practically possibly) for 15 minutes until they have completed the book. This is the minimum expected level of implementation 'fidelity' to the teacher choice approaches being trialled in the Story Time Trial.
- ii) We will also explore **other aspects of implementation support and characteristics including**: previous experience of the book, usual reading aloud practice, whether teachers had queries on the guidance document or implementation, ability to stick to the allocated approach (e.g. without questions/discussion in the GO! condition), any additional questions/prompts teachers may have used in the STOP! condition, and whether teachers moved on to another book in the same condition and how they approached this for any remainder sessions.

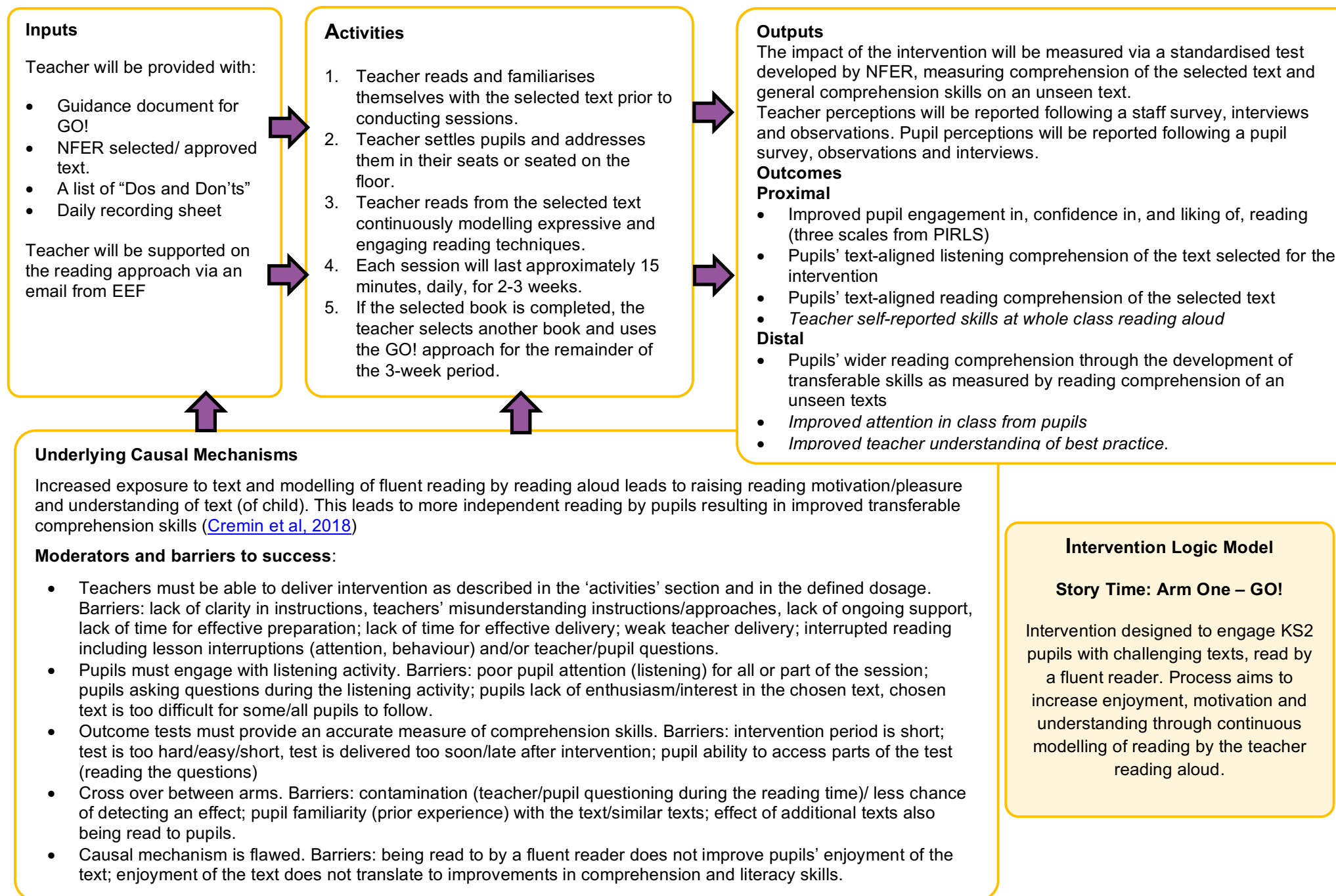
We will also explore **class-level dosage** as a way of understanding pupils' exposure to the approach. Teachers are expected to use their allocated approach for daily 15 minutes sessions (excepting unforeseen issues or pre-planned trips etc.) until they complete the book. This is expected to be within 2–3 weeks. Teachers can continue with another book using the same approach for any remaining sessions within the 3-week window. Teachers can voluntarily log in a 'daily recording sheet' how many sessions/minutes they use the approach for with the selected book, and with any additional book. This daily recording sheet can help teachers respond accurately to a set of questions in the online survey, administered at the end of the term. Please see the compliance section for further details.

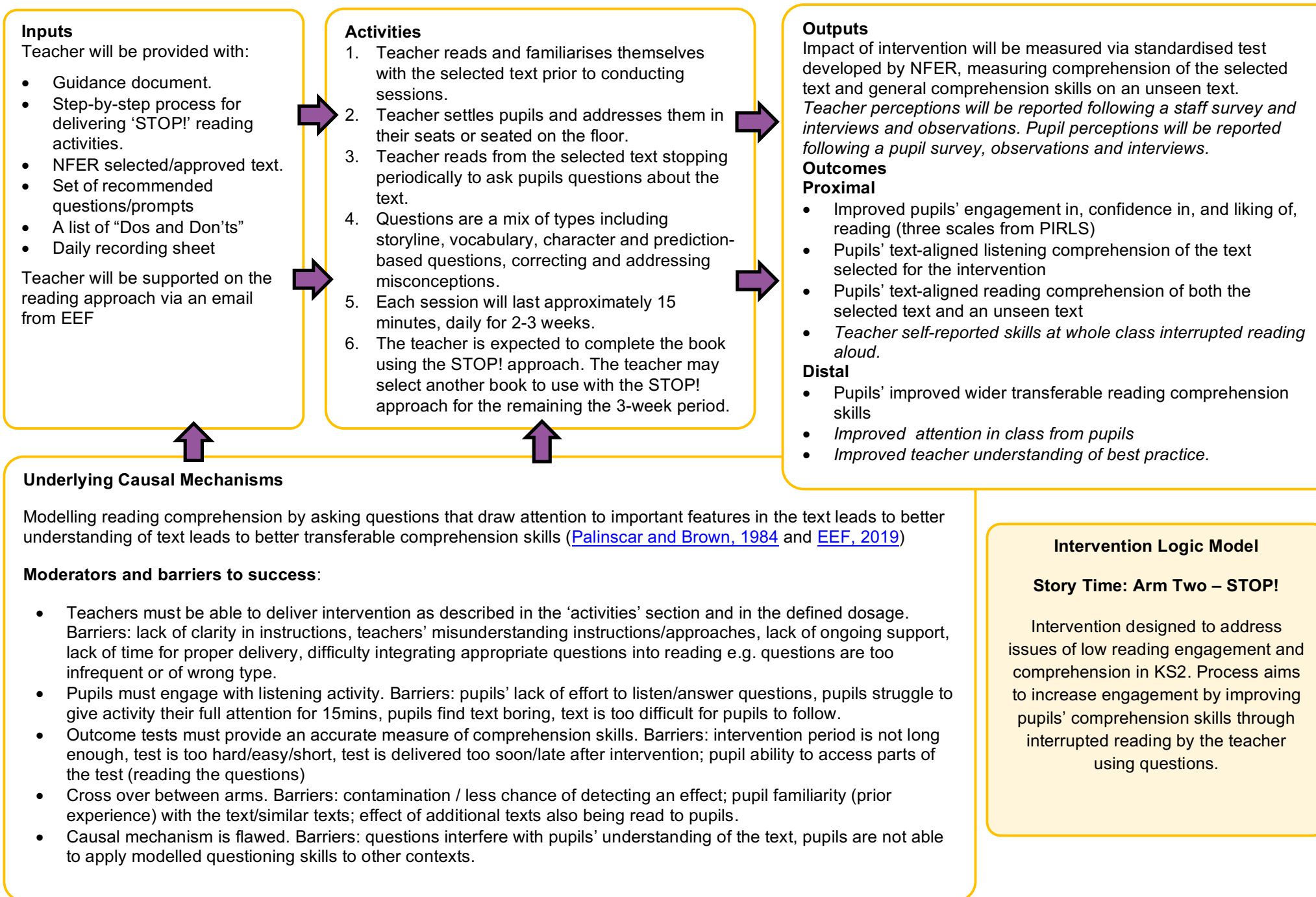
Costs

Costs in terms of planning time for each condition will be explored. We will also note the cost of the book.

Figure 1: Logic model for the continuous reading GO! approach (on following page)

Figure 2: Logic model for the reading and questioning STOP! approach (on following page)





Feasibility of measurement and impact evaluation

Research questions

The primary aim of the trial is to explore the feasibility of the Teacher Choices approach (see Table 1, RQs 1 and 2). It will contribute to the overall research aims of the Teacher Choices study. In order to do this, we will first explore the feasibility of the text-aligned assessment and subsequently conduct an impact evaluation (see Table 1, RQs 3, 4 and 5), which would enable us to estimate the impact of the two reading aloud choices- GO! approach and STOP! approach on a number of attainment and non-attainment pupil outcomes.

Within the impact evaluation, the primary research question for this trial is:

What is the difference in pupils' comprehension skills who received the GO! approach compared to those who received the STOP! approach? This measure will comprise the reading and listening comprehensions of the text selected for the intervention and therefore it is a text-aligned comprehension skill.

Within the impact evaluation, the secondary research questions for this trial are:

1. What is the difference in pupils' wider reading comprehension skills (of an unseen text) who received the GO! approach compared to those who received the STOP! approach?
2. What is the difference in pupils' engagement in reading lessons who received the GO! approach compared to those who received the STOP! approach?
3. What is the difference in pupils' confidence in reading who received the GO! approach compared to those who received the STOP! approach?
4. What is the difference in pupils' liking of reading who received the GO! approach compared to those who received the STOP! approach?

Design

The story time trial will be a cluster RCT with randomisation to take place at school level. We considered random allocation of year groups, i.e. Year 4 and Year 5 within a school to be in different randomised groups, but this will increase the ICC and therefore require a much larger sample.

Table 2: Trial design

Trial design, including number of arms		Two-armed RCT
Unit of randomisation		School
Stratification variables (if applicable)		N/A
Primary outcome	variable	Text-aligned comprehension score
	measure (instrument, scale, source)	Reading and listening comprehensions of the text selected for the intervention, assessment developed by the NFER
variable(s)		Pupils' wider reading comprehension skills

Secondary outcome(s)	measure(s) (instrument, scale, source)	Pupils' engagement in reading lessons, confidence in reading and liking of reading
		Reading comprehension of an unseen text Pupil survey using existing scales from the PIRLS questionnaire
Baseline for primary outcome	variable measure (instrument, scale, source)	There is no baseline measure for the primary analysis
Baseline for secondary outcome	variable measure (instrument, scale, source)	There is no baseline measure for the primary analysis

Randomisation

After eligible schools sign the Memorandum of Understanding (MoU) and send the teacher and pupil baseline data, we will randomly allocate each participating school to the GO! or the STOP! approach. This means, providing baseline teacher and pupil data will be a condition to randomisation. We will recruit at least 100 primary schools and the trial will be a standard parallel group design (1:1), in which all Year 4 and Year 5 teachers from a school will use the allocated approach with their participating classes over a period of approximately three weeks. This will be a simple school-level randomisation with no stratification.

Participants

Schools: All state-funded primary schools in England with Year 4 and Year 5 pupils are eligible to take part. In most schools, we will require at least one Year 4 and one Year 5 class to participate. Beyond this, schools can nominate as many Year 4 and Year 5 classes as they wish to participate. If the schools do not nominate all Year 4 or all Year 5 classes, the nominated classes need to be mixed ability rather than streamed or set classes. If a school has a Year 4 or Year 5 class as a mixed year group with Year 3 or Year 6, it will not be eligible to participate. However, if Year 4 and Year 5 are taught together, this class can take part in the trial.

Teachers: All Year 4 and Year 5 teachers with a classroom teaching responsibility for the nominated classes are eligible to participate. Teaching assistants and school leaders are not eligible to take part.

Pupils: All Year 4 and Year 5 pupils from the participating classes are eligible to participate.

Sample size calculations

Data from National Curriculum tests (DfES, 2004) indicates annual progress of about 0.8 of a standard deviation (SD) at age 7 (at the end of Key Stage 1), falling to 0.7 at 11 (at the end of Key Stage 2) and only 0.4 at age 14 (end of Key Stage 3). If we assume an increase of 0.8 SD per year, we might expect 0.1 per month given children are at school for roughly three-quarter of the year; so 'meaningful progress' might be defined as 0.1 SDs over a four-

week period. However, the analysis will focus on the difference between the two groups that might have made around 0.1 SDs of progress each.

Given that the main aim of the trial is about the feasibility of running such a trial in primary schools, the target of 100 primary schools is decided to make the trial as big as feasible within the available timeframe. Table 3 presents expected minimum detectable effect size with the following assumptions: 100 primary schools with 45 pupils in each year group and 0.12 ICC.

Table 3: Sample size calculations

		OVERALL
Minimum Detectable Effect Size (MDES)		0.20
Intraclass correlation (ICC)		0.12
Alpha		0.05
Power		0.8
One-sided or two-sided?		Two-sided
Average cluster size		90
Number of schools	Intervention	50
	Control	50
	Total	100
Number of pupils	Intervention	4,500
	Control	4,500
	Total	9,000

Outcome measures

Baseline measures

In the original trial, KS1 English reading ARE was intended to be the baseline measure. However, this is not available for pupils who are in Year 4 in 2021-22 as these assessments were cancelled in May 2020 due to covid-related school closures. The revised trial design does not include a baseline measure for primary or secondary analysis models.

In the absence of KS1 as a baseline measure, it was decided that NFER will collect English or reading assessment data for schools that used commercial tests for their pupils. At the time of writing the protocol, it is not known how many trial schools used such tests in the spring or summer terms 2021 but it is assumed that this would be the case for approximately half the trial schools. NFER will collect this data directly from the assessment providers and will use them as baseline measure in additional analyses. Further details on the methods will be included in the Statistical Analysis Plan (SAP).

Primary outcome

NFER's Centre for Assessment will create an end-point assessment to measure the proximal outcomes based on the selected book- *The Iron Woman* by Ted Hughes. The

assessment will also include reading comprehension questions using an unseen text that measures pupils' transferable skills in reading comprehension (distal outcome). The assessment team will also generate the questions for the STOP! approach and will ensure that the end-point assessment does not repeat any questions covered in the STOP! condition questions. All the questions (or items) will be created during workshop sessions where groups of researchers will work together to write items for a particular text. This will result in a wealth of items, which will be further developed and refined by a member of the project team. Once a collection of items is devised, an experienced member of the assessment team will review them prior to preparation for 'qualitative research' or assessment trialling. We envisage to write approximately 50-60 questions, which will be divided between the STOP! approach and the outcome measurement.

The end-point assessment will comprise of four parts. Table 4 that provides an overview of the end-point assessment.

Table 4: Overview of the end-point assessment

Assessment section/outcome measures	Number of independent items	Allotted time (minutes)	Total marks
Global questions on the selected text	4	5	4
Reading comprehension on the selected text	20	20-25	20
Listening comprehension on the selected text	11	10-15	11
Wider reading comprehension based on an unseen text	16	15-20	16

In order to ensure ease of administration and marking, the vast majority of items for the primary outcome measure will be closed response, but we will produce a variety of different item types, including but not limited to multiple choice, to give pupils a more interesting and rounded experience. For example, other items will include matching, sequencing, true/false or table completion questions.

For the assessment, each pupil will have a reading stimulus booklet (containing the passage from *The Iron Woman* and a passage for the unseen text). There will be a separate answer booklet, which contains all the questions. Class teachers will administer these assessments after the three-week intervention period by following the administration instructions provided by us.

In total, the assessment will be worth 54 marks and it will take about an hour to administer. We will recommend to schools that the administration is split across two sessions, with a break between them – e.g. before and after playtime/lunchtime or on separate days of the week – so that children complete the first three parts, containing the items focused on *The Iron Woman*, in one sitting and the last part in a second sitting.

Whilst the assessment has undergone informal trialling, only at the end of the RCT will it be possible to carry out formal psychometric testing on the functioning of items and the reliability of individual scales. More detail on the method will follow in the SAP.

Secondary outcomes

Attainment outcome

The wider reading comprehension score will be used as a secondary outcome of attainment. We will use a pre-existing assessment of reading comprehension, although the questions may need to be revised so as to create a measure that is age-appropriate for both year groups. See Table 4 for further details on the number of questions and total marks for this outcome.

Non-cognitive outcomes

NFER will use three scales from the PIRLS to measure pupils' engagement in reading lessons, their confidence in reading and their liking of reading. These scales come from the student questionnaire used in the PIRLS 2016. The questionnaire is freely available to administer and it is validated in England. The PIRLS student questionnaire items were developed to be combined into scales to measure a single underlying latent construct. These scales were constructed using the item response theory (IRT), specifically the Rasch partial credit model (Martin *et al.*, 2017).

The Students Engaged in Reading Lessons (ERL) scale was created based on students' degree of agreement with nine statements. The Cronbach's Alpha reliability coefficient of the scale for England in PIRLS 2016 was 0.84. The percentage of variance among the scale items accounted for by the principal component of the scale in England was 44% (Martin *et al.*, 2017, p. 14.89). The Students Confident in Reading (SCR) scale was created based on students' degree of agreement with six statements. The Cronbach's Alpha reliability coefficient of the scale for England in PIRLS 2016 was 0.82. The percentage of variance among the scale items accounted for by the principal component of the scale in England was 54% (Martin *et al.*, 2017, p. 14.85). The Students Like Reading (SLR) scale was created based on students' responses to ten items. The Cronbach's Alpha reliability coefficient of the scale for England in PIRLS 2016 was 0.87. The percentage of variance among the scale items accounted for by the principal component of the scale in England was 47% (Martin *et al.*, 2017, p. 14.93).

We will submit a permission request to IEA (International Association for the Evaluation of Educational Achievement) in order to reuse selected items from the student questionnaires. Once this request is approved, we will create a pupil survey using these items. The survey will be administered after the three-week intervention period but before the end-point assessment takes place.

Compliance

Compliance will be measured using information collected via the online teacher feedback survey. We will send a printed daily recording sheet, which teachers can use on a daily basis to record brief information about the implementation of the sessions. The use of the daily recording sheet is optional as an aid to help teachers to record compliance information in the survey at the end of the intervention – hence, we will not request teachers to return completed daily recording sheets.

To comply with the intervention, teachers must:

- read the guidance booklet of their reading aloud approach
- make a credible attempt to comply with the Do's and Don'ts of the approach
- plan a suitable and regular time to deliver the daily reading aloud sessions
- keep the daily session close to 15 minutes
- use the specified book
- if in the STOP! condition, use the prompt questions provided

Teachers assigned to the STOP! approach will avoid contamination by not having too long periods of continuous reading without questions. Teachers assigned to the GO! approach will avoid contamination between arms by not asking questions to pupils.

The dosage measure for compliance will be total amount of time (in minutes) that the teacher used with an allocated reading aloud approach. This will include time spent in reading aloud whilst completing the selected book and time spent while reading aloud any other book using the allocated approach for any remaining sessions within the intervention period.

Analysis

As one of the research questions is around feasibility of using a text-specific test in future, we will start by exploring the reliability of the proposed outcomes. Please see above section on primary outcome measures for further details of how this will be explored. Detailed methods will be provided in the SAP.

The primary analysis will be intention-to-treat. Due to the cluster-randomised design, analysis will use a multilevel model with two levels: schools and pupils. The total score for text-aligned reading and listening comprehension questions will be regressed on three covariates: whether the pupil belongs to the GO! versus STOP! trial arm and their year group.

Each of the four secondary outcome models will also be constructed similar to the primary model. There will be four separate models- one for the wider reading comprehension outcome (using the total score from the reading comprehension of the unseen text) and three models for each of the PIRLS scales. We will use the same covariates as used in the primary model: whether the pupil belongs to the GO! versus STOP! trial arm and their year group.

We will also derive a baseline measure using the English or reading scores from the spring/summer terms prior to the trial. This will be done using the raw test scores that we receive from the assessment providers. We will link the tests via their z-scores (calculated using the means and standard deviations reported by the providers). This standardised baseline measure will then be used as an additional covariate in the primary outcome model. Results from these models will be compared for sensitivity analysis.

We will also present a set of sample size parameters which will be important for future teacher choices trials. These will include school-level intra-cluster correlation, pre-post correction and effect size.

Any other additional exploratory analyses will be detailed in the SAP.

Implementation and process evaluation

IPE research focus

The implementation and process evaluation (IPE) will focus on Research questions 1 and 2 of the whole study and will serve two purposes:

- **Feasibility of the RCT methodology:** To understand how the trial methodology is followed by teachers, including adherence to randomisation condition, administering the tests and returning data as well as their views on the amount/ nature of 'trial information' provided.
- **Implementation of the choice/approach:** To understand how the reading aloud approach is implemented by teachers, including views on how much/the nature of guidance and support that is needed for the 'choice' compliance, the extent of implementation fidelity, how the 'choice' compares to the usual practice, and if adaptations are required and why.

Research methods

The IPE will be informed by the EEF's latest IPE guidance (EEF, 2019), with a focus on compliance, fidelity and usual practice data collection, and collaboratively developing a logic model. It will also be guided by the other dimensions and actors in EEF's IPE guidance (Humphrey et al., 2017), focusing additionally on: responsiveness, perceived quality, implementation adaptation and on implementation support factors. These are the dimensions and factors that will help us explore the extent to which teachers are able to implement the choice, whether they adapt it to their context; and the extent to which this approach and design is feasible in the classroom.

The IPE for 'The Story Time trial' will involve:

1. A pre-trial consultation on the amount/nature and content of the 'choice' guidance document, to ascertain if it is suitable to trial and to act on any changes/improvements suggested by teachers prior to it being issued in the trial (this will include consultation feedback from schools in a Research School network)
2. An IDEA workshop to refine the logic model(s), confirm compliance requirements, and create brief TIDIER frameworks for each of the reading aloud approaches
3. Observation visits³ to see the reading aloud approaches in action, including brief interviews with teachers and head teachers, and discussions with a small number of pupils.
4. A short feedback survey to gather
 - teachers' views on the 'choice' guidance (e.g. what works well, what needs improving);

³ Visits to schools will depend on COVID-19 regulations guidance at the time of the fieldwork, and will be carried out only if we can guarantee safeguarding of pupils, teachers and staff. If school visits are advised against, NFER will find alternative ways to collect this information.

- information about compliance/implementation/dosage of the approach (e.g. engagement with the guidance; use of questioning; use of the selected text; whether or not teachers excluded any pupils from the Story Time sessions; whether or not the book was completed during the trial period; total minutes used by the approach; time of day sessions typically delivered);
- teacher outcomes (e.g. skills at whole class reading aloud; understanding of best practice for whole class reading aloud);
- pupil outcomes (e.g. attention in class; reading comprehension skills); and
- usual practice.

Teachers will receive a printed daily recording sheet that they can use at their discretion to record daily activities. The aim of these daily records is to inform and facilitate the response to the compliance and dosage questions on the feedback survey at the end of the intervention.

5. Telephone interviews with a purposive sub-sample of eight teachers (two STOP! and two GO! where, according to survey responses⁴, the choice and trial method are going well; and two STOP! and two GO! where there are challenges/suggestions for improvements), to get more in-depth views on the above, as well as on the implementation of the reading aloud approaches.

Table 5: IPE methods overview

Research methods	Data collection methods	Participants/ data sources (type, number)	Data analysis methods	Research questions addressed	Implementation/ logic model relevance
IDEA workshop	TIDIER framework; logic model completion	EEF team, NFER team	Descriptive analysis	RQ1; RQ2	Context
Pre-trial consultation on guidance	Consultation form/email/calls, with non-trial teachers	Up to three heads of literacy; up to three Y4/5 teachers	Perspective analysis; formative feedback	RQ1; RQ2	Quality; implementation support
Classroom observations and interviews (subject to COVID-19 restrictions guidance)	Semi-structured observations and interviews	8 sites (4 GO!, 4 STOP!); observe and interview head/SLT, up to 2 teachers per site; and small group interview with pupils	Within/cross-site triangulation and variation	RQ1; RQ2; RQ6 RQ7	Fidelity Implementation and support quality Usual practice Responsiveness Adaptation
Short feedback survey	Online teacher survey. (Teachers will be provided with a daily recording sheet to be filled out voluntarily. This will help inform their survey responses.)	Sent to all teachers involved (up to 200 teachers)	Basic frequencies; thematic analysis; formative views	RQ1; RQ2; RQ6	Compliance; fidelity; responsiveness; adaptation; usual practice; dosage; <i>perspectives on trial methodology; adherence to trial methodology, and any contamination</i>

⁴ These will be purposive, if time allows for sampling based on survey responses. If not, they will be sampled by key school characteristics.

Telephone interviews	Purposive sub-sample from survey responses, semi-structured interviews	8 teachers; 2 Stop and 2 Go where choice and trial method are going well; 2 Stop and 2 Go where challenges	Thematic analysis; formative analysis	RQ1; RQ2; RQ6; RQ7	Implementation; quality; fidelity; compliance; responsiveness; adaptation; usual practice; <i>perspectives on trial methodology</i>
----------------------	--	--	---------------------------------------	--------------------	---

Process analysis

NFER will analyse the data collected through the teacher survey to explore if teachers implemented the allocated story time approach for 15 minutes in the majority of lessons and whether teachers used another book in the same approach in any remaining sessions in the 3-week window.

We will explore the nature of fidelity, implementation, any adaptations, use of the STOP! prompts and any additional questions used by teachers in the STOP! condition, and any further reading done in the same approach; through classroom observations (subject to COVID-19 restrictions' guidance), interviews and surveys. Qualitative analysis will involve perspectives analysis (detailed analysis of responses gathered from different viewpoints, such as teachers, senior leaders and pupils), and observational analysis (analysis of observed behaviours in classrooms using the two reading aloud approaches). Taken together, these analyses will provide an in-depth, comprehensive understanding of the practicalities of implementation, including challenges encountered and addressed. Responses and behaviours will be coded into themes that emerge from the data sources (thematic analysis) structured with reference to relevant research questions (bearing IPE dimension and supporting factors in mind as appropriate to each RQ). All IPE data sources will be triangulated to draw out the range and prevalence of themes that will shape the reporting narrative.

To aid the formative aspect of the overall study and learning for future Teacher Choices trials, NFER will also:

- Draw on learning from Trial 1 to inform Trial 2 (such as the recruitment strategy, trial design issues, data collection approaches and communication with schools)
- Facilitate a discussion workshop with EEF – based on formative feedback/findings so far from the trials (e.g. at the end of Trial 1 and whilst Trial 2 is underway) – to inform reporting.

Cost information

The intervention will not bring about direct monetary costs to schools. In terms of other resources, there might be costs associated with teacher planning time since they will have read the text prior to the intervention to familiarise with it, and those in the STOP! condition will need time to familiarise with the prompts provided. We will use the case studies and telephone interviews to ask teachers about the time involved in preparation for the intervention. We will report the cost of the book that is being provided through the trial.

Ethics and registration

The trial will be designed, conducted and reported to CONSORT standards (<http://www.consort-statement.org/>). This trial is set up within a feasibility study to understand and explore the best designs and methods to evaluate teacher classroom practices, hence it will not be registered. The evaluation will be conducted in accordance with NFER's Code of Practice, available at: <https://www.nfer.ac.uk/media/1166/codeofpractice.pdf>

The school headteacher will provide the ethical agreement for participation in the trial when signing the MoU. A nominated teacher from each school will act as a key contact person for the trial. NFER will provide full details about the evaluation to Year 4 and Year 5 teachers and parents. Parents will have the opportunity to withdraw their child from data processing if they have objections to this.

Data protection

All data gathered during the trial will be held in accordance with the data protection framework created by the Data Protection Act 2018 and the General Data Protection Regulation 2016/679. It will be treated in the strictest confidence by the NFER. NFER will share names and contact details of all school staff involved in the trial with EEF. EEF will use this data to provide support on the reading approaches.

Explicit permission will be gained from each school for NFER to approach assessment providers to collect assessment data for their pupils. For this purpose, NFER will share school names with assessment providers so that they are able to provide NFER with the outcomes of pupils' English or reading assessments from the spring/summer term 2021.

The pupil personal data collected by the NFER will not be shared with any other organisation.

NFER is the sole data controller for this trial (The Story Time Trial). EEF is the data processor.

The legal basis for processing personal data is covered by GDPR Article 6 (1) (f): *'processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of the personal data'.*

We have carried out a legitimate interest assessment, which demonstrates that the evaluation fulfils one of NFER's core business purposes (undertaking research, evaluation and information activities) and it has broader societal benefits. It is therefore in our legitimate interest for to process and analyse personal data for the administration of the. We have considered and balanced any potential impact on the data subjects' rights and find that our activities will not do the data subject any unwarranted harm.

The headteacher of participating schools will sign a MoU, which collects the details of the nominated teacher who becomes the key contact for the trial. During the trial, schools will also provide details of Year 4 and Year 5 teachers that agreed to take part in the study, and details of their pupils (where parents have not objected to their child's data being processed).

NFER will collect the following personal data for participating teachers: names, contact details, job role, years of experience, specialism (if any), their activity records, and their attitudes and views about the choice and the trial and how it has been implemented.

For each student of participating classes, the following personal data will be collected: names, date of birth, gender, Free School Meal eligibility (FSM), English as an Additional Language (EAL), assessment data for English or reading from assessment providers, score for the end-point assessment and their attitudes towards reading via a survey (and interviews in a small number of cases, where possible).

All the personal data will be transferred via NFER secure data portal. Teacher surveys will be completed online via Questback. Pupil assessments and surveys (completed on paper) will be returned to NFER using a secure courier.

All personal data will be deleted within one year of publication of the final report; currently expected to be published in summer 2022. The privacy notice intended for research participants (school staff and parents) can be found at https://www.nfer.ac.uk/media/3870/eetc_tc2_privacy_notice.pdf

Personnel

EEF will be responsible for developing the teacher guidance and will act as a delivery provider. NFER will be the independent evaluator for the Teacher Choices trials.

Table 6: Personnel

Name	Institution	Role
Megan Dixon (in her previous role)	EEF	Intervention design and producing guidance for teachers
Emily Yeomans, Joe Collin (in his previous role) and Igraine Rhodes	EEF	Leading on learning for EEF and programmes te
Rob Coe	EEF	Advising on evaluation design, statistical analysis and testing
Ben Weidmann	EEF	Advising on evaluation design and statistical ana
Florentina Taylor and Arnaud Vaganay (in their previous roles); and Faizaan Sami	EEF	Leading on EEF evaluation side
Ben Styles	NFER	Trial director
Palak Roy	NFER	Trial manager
Joana Andrade	NFER	Trial statistician
David Sims	NFER	Process evaluation director
Jennie Harland	NFER	Process evaluation researcher
Caroline Stanford	NFER	Process evaluation researcher
Maria Galvis	NFER	Process evaluation researcher
Kathryn Hurd and Kathryn Gallop	NFER	Operations lead
Guvi Chohan	NFER	Operations researcher
Neal Hepworth	NFER	Assessment development

Risks

Table 7: Risk assessment

Risk	Likelihood	Impact	Mitigation actions
Insufficient schools or teachers recruited to the study due to local/national COVID-19 lockdowns/school closures	MAY HAPPEN	SIGNIFICANT	Discuss the possibility of drawing a top-up sample; communicate with the schools that took part in the original trial and liaise with EEF and NFER's Communications team to tweet about the trial. Also, discuss the possibility of shifting the timetable for the whole trial in case sufficient schools are not recruited.
Local/National COVID-19 lockdowns leading to school closures	MAY HAPPEN	SIGNIFICANT	A local lockdown will have different effects depending on when it occurs. Recruitment and baseline data collection: Schools may be unable to take part in the trial and supply baseline data. Approaching schools who were recruited for the original trial will ensure we have some existing schools willing to take part and aware of the trial requirements. Submission of baseline data will be a pre-requisite of randomisation. Choice implementation: Severe disruption in a subset of schools. NFER's ongoing communication with schools will ensure if Story Time sessions are disrupted due to COVID-19, compliance records reflect this.

Risk	Likelihood	Impact	Mitigation actions
			Dropout from or insufficient IPE data: Tools for interviews and focus groups will be designed so that they can be adapted to phone data collection. If insufficient schools want to take part in telephone interviews, alternative options of data collection mechanisms will be considered or the school sample will be expanded. Follow-up measurements: Severe disruption in a subset of schools. We would want to delay measurement until they re-open if possible. COVID-19 disruption is unlikely to introduce bias.
Tight timetable for recruitment, baseline and randomisation	LIKELY	SIGNIFICANT	Discuss strategies to increase the rate at which each activity takes place including more NFER staff, regular internal catch-ups and communicate to EEF
Tight analysis and reporting timetable	MAY HAPPEN	SIGNIFICANT	Discuss strategies to increase the rate at which each activity takes place including more NFER staff, regular internal catch-ups and communicate to EEF
Insufficient schools and/or teachers recruited to the study	MAY HAPPEN	SIGNIFICANT	Discuss the possibility of drawing a top-up sample; communicate with early EOI schools, speak to EEF and NFER's Communications team to tweet about the trial
Insufficient baseline completions ahead of randomisation	MAY HAPPEN	SIGNIFICANT	Keep one to two week window between baseline data collection and randomisation so timely reminders can be sent and data can be received before randomisation.
School drop out from the intervention. That is, schools do not implement the assigned Teacher Choice	MAY HAPPEN	MODERATE	NFER to contact schools to inquire school status and gain feedback.
Teacher drops out from the intervention. That is, they do not implement the assigned Teacher Choice	MAY HAPPEN	MODERATE	NFER to contact schools to inquire school status and gain feedback.
Attrition from assessment data collection.	MAY HAPPEN	MODERATE	NFER to contact schools to inquire school status and gain feedback.
Key contact attrition (e.g. key contact becomes unresponsive or leaves the school during the course of the trial)	UNLIKELY	VERY LOW	Establish contact with the Head teacher who signed the MOU
Intervention (Teacher Choice) is not well implemented	MAY HAPPEN	MODERATE	EEF and NFER put the processes in place in terms of teacher guidance and NFER attends to email queries received
Teachers in a given trial arm adopts aspects of the other choice (contamination)	MAY HAPPEN	MODERATE	EEF and NFER put the processes in place in terms of teacher guidance and NFER attends to email queries received
Schools do not complete the data on implementation and	MAY HAPPEN	MODERATE	NFER to communicate to schools the importance of such data, and the importance of honesty and openness in responses

Risk	Likelihood	Impact	Mitigation actions
fidelity; or provide responses influenced by the evaluation, the choice or by social responsibility bias			(however much or little they used the approach, their experience and reasons are important).
Drop out from or insufficient implementation and process evaluation data	MAY HAPPEN	MODERATE	NFER to communicate to schools the importance of such data
Confidential test material sent via the courier is lost.	UNLIKELY	VERY SEVERE	Full Standards and Testing Agency (STA) security procedures and enhanced process established and in use by NFER.
Loss of taped interview recordings.	UNLIKELY	VERY SEVERE	Only NFER equipment is used. Recordings deleted once notes are written up.
Recruitment criteria affect schools'/teachers' capacity to take part	LIKELY	SEVERE	Review eligibility criteria and response to initial recruitment

Timeline

Table 8: Timeline for 'The Story time trial'

Dates	Activity
May/June 2021	• Trial delay discussion (budget revision/addendum)
June- September 2021	• School recruitment ◦ Schools sign MoU ◦ Schools supply pupil data • Protocol revision • Revise teacher guidance (EEF)
October 2021	• Randomisation • NFER to communicate randomisation results and send the Teacher Guidance to participating schools • EEF's support emails and ad hoc calls
November – December 2021 (second half of autumn term 2)	• Trial SAP • Teachers implement the Story Time Trial choices • Teachers keep a record of Story Time sessions on the log • Case-study visits and lesson observations • Teachers complete the pupil survey and end-point assessment with participating pupils • Teachers complete a short online survey
January – February 2022	• Short teacher telephone interviews • NFER marks the tests received from schools
March – April 2022	• Trial analysis and report writing
May 2022	• NFER sends the first draft report to EEF • NFER conducts focus group discussion(s) about reporting formats (1 focus group)
June – August 2022	• Report revisions
September 2022	• Report publication • NFER sends summary report to schools

References

- Department for Education and Skills, 2004. *Statistics of education: variation in pupil progress 2003*. [online] Available at: <<https://webarchive.nationalarchives.gov.uk/20130321095824/http://www.education.gov.uk/researchandstatistics/statistics/a00194385/>> [14 January 2021].
- Education Endowment Foundation, 2019. *Implementation and process evaluation guidance for EEF evaluations*. [pdf] Available at: https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/IPE_guidance.pdf [14 January 2021].
- Humphrey, N., Lendrum, A., Ashworthy, E., Frearson, K., Buck, R. and Kerr, K., (2016). *Implementation and process evaluation (IPE) for interventions in education settings: An introductory handbook*. [pdf] Available at: <https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/IPE_Handbook.pdf> [14 January 2021].
- Education Endowment Foundation (2020). *Teacher choices*. [online] Available at: <<https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/teacher-choices/>> [14 January 2021].
- Martin, M. O., Mullis, I. V. S., and Hooper, M. eds., 2017. *Methods and procedures in PIRLS 2016*. [pdf] Available at: <https://timssandpirls.bc.edu/publications/pirls/2016-methods/P16_MP_Chap14_Context_Questionnaire_Scales.pdf> [14 January 2021].