

Evaluation Summary	
Age range	Year 7
Number of pupils	577
Number of schools	20
Design	Efficacy Trial
Primary Outcome	Key Stage

## Protocol for Evaluation of Chatterbooks and Chatterbooks Plus

*Note: This protocol excludes aspects of the evaluation that are the sole responsibility of the University of Coventry and are not requirements of the EEF or NFER.*

### INTERVENTION

Chatterbooks is an extra-curricular reading initiative designed and delivered by the Reading Agency that aims to increase children's motivation to read. It consists of weekly small-group sessions, usually in a public or school library, where children read and discuss an age-appropriate, enjoyable book. The emphasis is on engaging children and encouraging creativity, rather than delivering instruction. Chatterbooks Plus will take the basic Chatterbooks programme and add a 15-minute period of dialogic reading, in which children read aloud and are offered explicit prompts on relevant vocabulary and situational knowledge in order to enhance comprehension. Dialogic reading is an [established approach](#) for much younger readers, but has not been adapted for or tested on an older group.

### RESEARCH PLAN

#### RESEARCH QUESTIONS

The primary research question is: what is the impact of Chatterbooks and Chatterbooks Plus on reading ability?

The secondary research question is: are improvements in attainment moderated by National Curriculum reading level or whether a pupil receives the pupil premium? Such interactions may not be causal.

### DESIGN

The project will be structured as a randomised controlled trial, with assignment carried out at the level of the individual pupil. A minimum of 450 pupils across around 20 secondary schools across Coventry, Solihull and Warwickshire will be involved. The trial will include three experimental groups: the first will receive the ordinary Chatterbooks programme; the second will receive the Chatterbooks Plus intervention; the third will act as a waitlist control group (they will ultimately receive whichever programme proves most effective). Children in the treatment groups will begin the programme in the summer 2013 term of Year 7. All children will be tested directly before and after intervention for reading ability. A follow-up test will be administered at the beginning of Year 8 to test whether either programme has 'inoculated' pupils against the usual effects of summer learning loss. Pre-testing will occur before pupils are randomised to avoid knowledge of the intervention affecting the pre-test results. After pre-testing, pupils will be randomised into the three groups. In order to avoid a

confounding teacher<sup>1</sup> effect, teachers will be randomised across the two treatment groups. Since teachers may prefer to visit schools in a particular area or with a certain ethos, they will be paired before randomisation<sup>2</sup>. At all stages, testing will be carried out by researchers who are blind to treatment allocation.

The trial will be designed, conducted and reported to CONSORT standards (<http://www.consort-statement.org/consort-statement/>).

### INCLUSION CRITERIA

Year 7 Pupils that were below National Curriculum level 4 in English and/or below level 4 in reading at the end of Key Stage 2, or pupils that are deemed to be 'vulnerable' Level 4 English achievers, as indicated by reading ages<sup>3</sup> below that of a 10 year old.

Children must have a reading age of at least that of a 7 year old to be included, in order to cope with the Chatterbooks reading requirement.

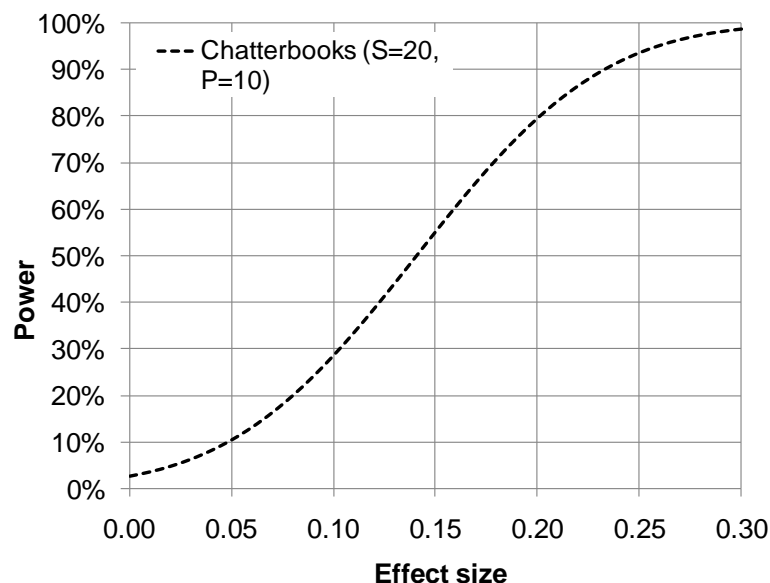
### RANDOMISATION METHODS

Both randomisations will be carried out by a statistician at NFER. Simple randomisation of pupils into three experimental groups of the same size will be carried out within each school. Stratified randomisation (or minimisation) of teachers will be performed.

### OUTCOME MEASURES

The digital version of the New Group Reading Test (NGRT; GL Assessment) will be used to measure reading ability. The NGRT has two subscales – ability and comprehension, which can be combined into a composite reading score. The composite score will be used as the primary outcome. The two subscales will be used as secondary outcomes.

### SAMPLE SIZE CALCULATIONS



S = Number of participating schools  
P = Number of participating pupils per intervention group per school

<sup>1</sup> Throughout this protocol, those tasked with delivering the programme are referred to as 'teachers' but in the case of Chatterbooks and Chatterbooks Plus they are actually university researchers.

<sup>2</sup> Whilst usually, pairing should be avoided before randomisation, this randomisation is just to guard against teacher effects so pairing is justified to ease practicalities.

<sup>3</sup> Using pre-existing data in schools.

Randomisation will be conducted at a pupil level, and furthermore we will be controlling for variation in baseline scores. Intra-class correlation ( $\rho$ ) is therefore likely to have a minimal impact on the effective sample size; we have conservatively assumed a value of  $\rho=0.02$  for the purposes of our calculations. The chart illustrates that the sample sizes will be sufficient to detect effect sizes at least of the order 0.20 – 0.25. This could be considered low-moderate, equivalent to around 3 months of progress – quite reasonable for targeted interventions providing support to small groups of pupils<sup>4</sup>.

## **ANALYSIS**

The primary outcome will be reading ability as assessed by the digital New Group Reading Test. Sub-group analysis on the primary outcome will be carried out on the following groups only: National Curriculum level and whether or not a pupil receives the pupil premium. The secondary outcomes will be the two NGRT subscales: reading ability and comprehension.

We will undertake basic descriptive analysis of baseline test data to provide a check that the randomisation process has been carried out successfully. Whilst we would not expect treatment and control groups to exhibit identical characteristics, we will carry out statistical tests to verify that any small differences that do arise are consistent with what one might expect assuming an unbiased randomisation.

We will then undertake our main analysis combining baseline and follow-up data. The definitive analysis will be ‘intention to treat’, reflecting the reality of how interventions are delivered in practice and avoiding attrition bias. We will use multi-level models to enable us to combine results across schools whilst accounting for clustering, and will include baseline data as a covariate in each of our models<sup>5</sup>. We will test hypotheses relating the impact of the interventions on pupils of differing abilities through the inclusion of interaction terms in the modelling.

The main analysis will be followed by an ‘on-treatment’ analysis where data from the teacher logs will be used to determine the extent of each pupil’s involvement with the interventions. We will also incorporate school-level variables into the analysis based on the questions addressing the extent to which teachers feel they maintained fidelity to the interventions, and any perceived contamination of the control groups of pupils. This analysis would enable us to estimate a ‘pure intervention effect’ (net of any fidelity issues, contamination, or non-completion). However, note that this analysis may be biased due to self-selection to differing levels of exposure<sup>6</sup>.

## **PROCESS EVALUATION**

At the outset of the project, the process evaluation researchers will arrange a telephone interview with the designer of each intervention which will inform the design of instruments. Researchers will obtain and analyse the training and guidance documents and attending a training session for both Chatterbooks and Chatterbooks Plus. Researchers will also observe one intervention session each of Chatterbooks and Chatterbooks Plus. The evidence from these document analyses and observations will inform the schedule for the later interviews and will directly contribute to the scalability evaluation.

The teacher log, which is proposed as a fidelity check for the interventions, will also contribute to the process evaluation. The record of whether and how the programme activities took place will give information on their practicability and manageability. The questions will provide data on teachers’ confidence and engagement. These analyses will provide an indication of how accessible and usable the new methods are for schools and teachers.

---

<sup>4</sup> Note that in the case of Chatterbooks, effect sizes are for paired comparisons between two of the three groups (e.g. Chatterbooks vs control or Chatterbooks Plus vs Chatterbooks). These differential effects will be smaller, and so are less likely to be detected for a given sample size.

<sup>5</sup> A second set of follow-up testing will be incorporated by extending the analysis to a repeated measures model.

<sup>6</sup> For example pupil motivation may be positively related to both levels of exposure to the intervention (through better attendance) and the amount of progress made between baseline and follow-up testing.

At the end of each intervention, researchers will gather more in-depth information on these matters by means of telephone interviews with two teachers delivering Chatterbooks and two who deliver Chatterbooks Plus. The telephone interviews will follow a semi-structured interview schedule, reflecting the distinctive features of each intervention but also following a common. We will look to gain a deeper understanding of teachers' perceptions of the intervention's impact and any barriers they perceive to exist for its wider rollout. Views would also be sought into the effectiveness of the training and guidance materials and whether any improvements to these processes and documents would make a wider rollout more likely to succeed.

Our report on the findings of the process evaluation will draw on these findings and make recommendations to ensure the sustainability and replicability of successful interventions when they are scaled up.

## **PERSONNEL**

The project will be led by Prof. Clare Wood, the founder and director of the Centre for Applied Research in Psychology at Coventry University. The design of the basic Chatterbooks programme will be the responsibility of the Reading Agency, a national charity that aims to promote reading and literacy at all ages, and the team at Coventry University will contribute the design of the dialogic elements for Chatterbooks Plus. The day to day coordination of the project will be overseen by a post-doctoral researcher to be appointed to the project. The impact evaluation will be led by Dr Ben Styles at NFER. The process evaluation will be led by Becky Clarkson at NFER. Camilla Neville will have overview of the evaluation at EEF and Emily Yeomans will oversee the grant.

## **ROLES AND RESPONSIBILITIES**

Each person will carry out their duties with the assistance of teams at their respective institutions:

Clare Wood – Recruitment and retention of schools, training and delivery of intervention, supply of list of eligible pupils for randomisation, administration of tests (different researchers will be used for intervention delivery and test administration)

Ben Styles – trial design, randomisation and analysis.

Becky Clarkson – process evaluation telephone interviews and visits.

## **DATA PROTECTION STATEMENT**

NFER's data protection policy is available at:

<http://www.nfer.ac.uk/nfer/about-nfer/code-of-practice/nfercop.pdf>

## **Timeline**

**Dec 2012:** meeting with partner organisations, write and register protocol

**Jan-Mar 2013:** Recruit and consent schools and pupils

**Mar 2013:** Pre-testing and random allocation of pupils

**April-June 2013** Implementation of intervention programmes

**June 2013** Post-testing (1)

**Aug 2013:** Analysis and interim results to EEF

**Sept 2013:** Post-testing (2)

**Oct 2013:** Analysis

**Nov 2013:** Report writing

**Dec 2013:** Final report to EEF

## Risks

Risk	Assessment	Countermeasures and contingencies
School, teacher or pupil attrition	<b>Likelihood:</b> moderate <b>Impact:</b> moderate	Clear information / initial meeting with schools explaining the principles of the trial and expectations. Both 'intention to treat' and 'on-treatment' analysis will be used. Attrition will be monitored and reported according to CONSORT guidelines.
Interventions are not implemented well	<b>Likelihood:</b> low <b>Impact:</b> moderate	Clear information / initial meeting with schools explaining the principles of the trial and expectations. Both 'intention to treat' and 'on-treatment' analysis will be used. Process evaluation will monitor this.
Control pupils exposed to elements of the interventions	<b>Likelihood:</b> moderate <b>Impact:</b> moderate	Clear information / initial meeting with schools explaining the principles of the trial and expectations. Both 'intention to treat' and 'on-treatment' analysis will be used.
Delays in training of teachers and commencing interventions	<b>Likelihood:</b> moderate <b>Impact:</b> low	Agree a clear timetable with project teams up front Revise timetable for pre and post testing periods
Failure in recruiting pupils/schools	<b>Likelihood:</b> low <b>Impact:</b> high	Project teams could make use of NFER's Research Operations Department to recruit more schools (at additional cost) Timescale could be revised
Poor completion of logs by teachers	<b>Likelihood:</b> moderate <b>Impact:</b> moderate	Set clear expectations at the start of the study what is required from participating teachers/schools Clear, simple design, and pre-population of logs with pupil names ensure log is straightforward to complete.
Researchers lost to project due to sickness or absence	<b>Likelihood:</b> moderate <b>Impact:</b> low	NFER has a large research department with numerous researchers experienced in evaluation who could be redeployed. Senior staff can stand in if necessary.
Project teams do not follow correct trial protocols	<b>Likelihood:</b> moderate <b>Impact:</b> high	Meetings with project teams at start of project. Provision of clear guidance describing protocols for distribution to all schools.