

Evaluation Summary	
Age range	11-12
Number of pupils	419
Number of schools	6
Design	Determined suitable for younger students
Primary Outcome	Key Stage 2

Protocol for Evaluation of Rhythm for Reading

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

INTERVENTION

Rhythm for Reading is a ten-week course of weekly ten-minute lessons in which children take part in rhythm-based exercises while reading musical notation. The sessions are delivered in the first instance by specialists who later train classroom teachers to deliver the intervention; in this way it can be embedded sustainably in schools. The programme was designed by Marion Long prior to undertaking her doctoral studies at the Institute of Education, and first came to prominence when it was awarded third place in an IPPR competition to identify the best new policy ideas from academics in British universities. In a [series of small-scale RCTs](#) Rhythm for Reading has shown strong positive effects on reading comprehension, with a gain score effect size of 0.8 standard deviations in the largest of these (involving around 25 pupils). Moreover, early indications suggest it is more effective for weaker readers. The effect size is likely to be considerably lower when pre-test standard deviations (rather than gain standard deviations) are used, but the results still suggest that further investigation is warranted. The mechanism by which Rhythm for Reading works is not well understood, but Dr. Long has suggested that it may help children to detect the natural rhythm and phrasing of prose, which supports quicker and more intuitive comprehension.

RESEARCH PLAN

RESEARCH QUESTIONS

The primary research question is: what is the impact of Rhythm for Reading on reading behaviour?

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 run as a randomised controlled trial, with a minimum of 400 individual Year 7 pupils across 5 or more secondary schools randomly assigned to two groups. It will focus on the weakest readers. Children will be extracted from classes and taught in groups of 10. Children in the treatment group will receive the Rhythm for Reading intervention; children in the control group will experience their usual English teaching.

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. If necessary for sample size, low level 4 readers will also be included.

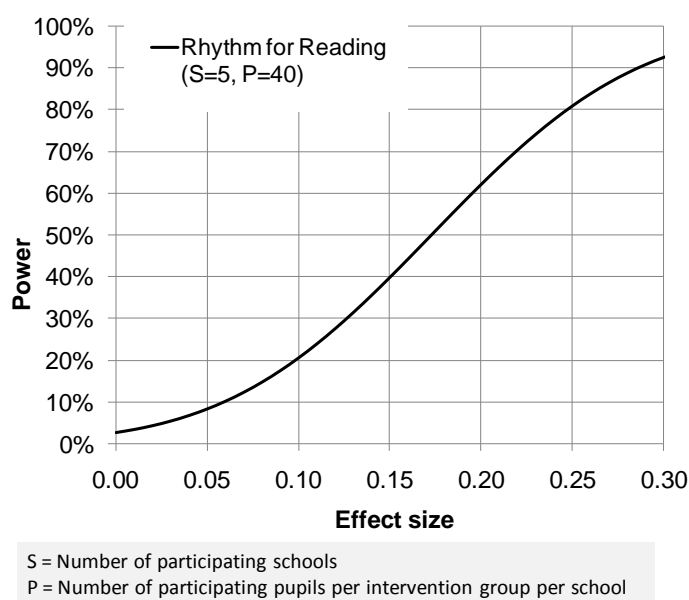
RANDOMISATION METHODS

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.

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



Randomisation will be conducted at a pupil level, and furthermore we will be controlling for variation in baseline scores in the final analysis. Intra-class correlation (ρ) 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 0.25. This could be considered moderate, equivalent to around 3 months of progress – quite reasonable for targeted interventions providing support to small groups of pupils.

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. 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 deliverers 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 exposure¹.

PROCESS EVALUATION

At the outset of the project, the process evaluation researchers will arrange a telephone interview with Marion Long which will inform the design of instruments. Researchers will observe two Rhythm for Reading intervention sessions: one delivered by Marion Long and the second by her colleague, Sally Cathcart. At the end of the intervention, other teachers will be trained in delivery. Researchers will attend one such training session. The evidence from these 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 engagement². These analyses will provide an indication of how accessible and usable the new methods are for schools and teachers.

At the end of each intervention, researchers will gather more in-depth information on these matters by means of telephone interviews with four teachers who will be trained to deliver Rhythm for Reading. The telephone interviews will follow a semi-structured interview schedule, reflecting the distinctive features of each intervention but also following a common design as far as possible. 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 any training and guidance materials and whether any improvements to these processes and documents would make a wider rollout more likely to succeed.

¹ 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.

² Only those delivering the programme need complete the log i.e. Marion Long and Sally Cathcart.

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 Dr Marion Long, a volunteer at Adamsrill School, Lewisham. 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:

Marion Long – Recruitment and retention of schools, delivery of intervention, supply of list of eligible pupils for randomisation, administration of tests (school teachers, i.e. not those who had delivered the intervention, will be used to deliver the tests). NFER Research Operations department will assist with the recruitment and retention of schools and with the administration of the tests as and when necessary.

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 2013: Recruit and consent schools and pupils

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

April-July 2013 Implementation of intervention programmes

July 2013 Post-testing and training of teachers

Aug-Sept 2013: Analysis and reporting

Oct 2013: Final report to EEF

Risks

Risk	Assessment	Countermeasures and contingencies
------	------------	-----------------------------------

School, teacher or pupil attrition	Likelihood: moderate Impact: 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	Likelihood: low Impact: 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	Likelihood: moderate Impact: 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.
Failure in recruiting pupils/schools	Likelihood: low Impact: high	Project team will make use of NFER's Research Operations Department to recruit more schools. Timescale could be revised
Poor completion of logs by teachers	Likelihood: moderate Impact: moderate	Set clear expectations at the start of the study what is required from those completing the logs. 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	Likelihood: moderate Impact: low	NFER has a large research department with numerous researchers experienced in evaluation who could be redeployed. Marion Long will train another delivery partner to stand in if required. Senior staff can stand in if necessary.
Project teams do not follow correct trial protocols	Likelihood: moderate Impact: high	Meetings with project teams at start of project. Provision of clear guidance describing protocols for distribution to all schools.