



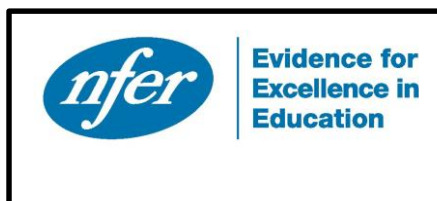
Education  
Endowment  
Foundation

# Rhythm for Reading

Evaluation Report and Executive Summary

May 2014

Independent evaluators:



Dr Ben Styles

Rebecca Clarkson

Katherine Fowler

# The Education Endowment Foundation (EEF)

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

We aim to raise the attainment of children facing disadvantage by:

- Identifying promising educational innovations that address the needs of disadvantaged children in primary and secondary schools in England
- Evaluating these innovations to extend and secure the evidence on what works and can be made to work at scale
- Encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

Founded by the education charity the Sutton Trust, as lead charity in partnership with Impetus Trust, the EEF is funded by an initial £125m grant from the Department for Education. With investment and fundraising income, the EEF intends to award as much as £200m by 2026.



## Literacy Catch-up Round

This study was funded by the Education Endowment Foundation as one of 23 projects funded through a themed funding round on literacy catch-up at the primary-secondary transition. Projects funded within this round aimed to identify effective ways to support pupils who do not achieve Level 4 in English by the end of Key Stage 2. It was one of four programmes with a particular focus on comprehension.

**For more information please contact:**

**James Richardson**  
Senior Analyst  
Education Endowment Foundation  
9th Floor, Millbank Tower  
21-24 Millbank  
SW1P 4QP



Education  
Endowment  
Foundation

**p:** 020 7802 1924  
**e:** james.richardson@eefoundation.org.uk  
**w:** www.educationendowmentfoundation.org.uk

## About the evaluator

The programme was independently evaluated by a team from the National Foundation for Educational Research. The lead evaluator was Dr Ben Styles. For the impact evaluation he was assisted by Sally Bradshaw and Alix Godfrey and for the process evaluation by Rebecca Clarkson and Katherine Fowler.

### Contact details:

#### Dr Ben Styles

The Mere  
Upton Park  
Slough  
Berkshire  
SL1 2DQ

**p:** 01753 637386

**e:** [b.styles@nfer.ac.uk](mailto:b.styles@nfer.ac.uk)

# Contents

**Executive summary ..... 2**

**Introduction..... 5**

**Methodology ..... 7**

**Impact evaluation ..... 12**

**Process evaluation ..... 19**

**Conclusion ..... 24**

**References ..... 26**

**Acknowledgements..... 26**

**Appendices ..... 27**

## Executive summary

### The project


Rhythm for Reading is a programme which aims to improve children's reading ability by taking part in rhythm-based exercises such as stamping, clapping and chanting, while reading musical notation. The intervention builds on the evidence of a link between the natural rhythm and phrasing of prose and intuitive reading comprehension.

Rhythm for Reading was originally developed as an intervention for primary school pupils. In this evaluation, the intervention was delivered to Year 7 pupils who had not reached a secure Level 4 in English at the end of Key Stage 2. They received weekly ten-minute sessions over a period of ten weeks between April and July 2013. Pupils were randomised to receive the intervention in Year 7 or to a waitlist control group. The intervention was delivered in schools by the developer.

The evaluation was funded by the Education Endowment Foundation as one of 23 projects focused on literacy catch-up at the primary-secondary transition. It was one of four programmes funded with a particular focus on reading comprehension.

### What impact did it have?

Rhythm for Reading demonstrated an effect size for all pupils of 0.03 and an effect size for children on free school meals (FSM) of 0.11. Effect sizes of this magnitude can be translated into approximately one month of additional progress for the pupils who received the intervention compared to those who did not. However, neither of these findings is statistically significant, suggesting that the difference in outcomes between the control and the intervention group occurred by chance. As a result, it is not clear that Rhythm for Reading is an effective intervention to use with Year 7 pupils who had not achieved the expected level of attainment in literacy at the end of primary school.

Group	Number of pupils	Effect size (95% confidence interval)*	Estimated months' progress	Is this finding statistically significant?*	Evidence strength**
Rhythm for Reading	175	<b>0.03</b> (-0.13, 0.18)	+1 <sup>1</sup>	No	
Rhythm for Reading (FSM pupils)	66	<b>0.11</b> (-0.16, 0.37)	+2	No	Security not provided for sub-group analysis
* Effect sizes with confidence intervals that pass through 0 are not 'statistically significant', suggesting that the difference occurred by chance.					
** For more information about evidence ratings, see Appendix C.					

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

There is some evidence of pupils responding differently to the intervention depending upon their pre-test score. The relationship suggests that children of lower ability are likely to perform worse after the intervention than children of similar ability in the control group. For children of higher ability, they are likely to perform better after the intervention than similar children in the control group. It is unclear what accounts for this difference across the ability range. More research is needed to understand this interaction more fully and therefore this finding should not be used to justify Rhythm for Reading's use with a particular ability group without further work.

The programme was delivered in each school by the developer of Rhythm for Reading so it is unclear how effective it would be in a wider context. From the observations and interviews with staff and pupils in the process evaluation, a lack of engagement and poor behaviour spoiled the delivery of some of the sessions. This may have contributed to the lack of impact in the intervention group.

As this evaluation had a waitlist design, the control group received the intervention in the first term of Year 8. These sessions were delivered by teaching assistants who had received training from the developer. Follow-up interviews suggest that the sessions were popular with pupils although some of the pupils were sceptical as to how useful they would be in improving their reading and this may have contributed to a lack of engagement. As the intervention was designed for younger pupils, it is likely that by Year 8 the pupils were too old to engage with the Rhythm for Reading activities.

Combining the findings from the randomised controlled trial (RCT) and the process evaluation suggests that Rhythm for Reading might be more suited for evaluation with younger children.

### **How secure is this finding?**

There is some evidence from previous research of a link between the processing of rhythm and how this might impact on reading behaviour. The Rhythm for Reading programme has, to date, been evaluated using small scale non-randomised studies that showed promising results. This report is the first independent evaluation of the programme, and the first one to use randomised controlled trial methodology.

The trial was run as a small scale efficacy trial, randomly allocating 419 pupils in six schools to an intervention group or a waitlist control group. Efficacy trials seek to test evaluations in the best possible conditions to see if they hold promise. They do not indicate the extent to which the intervention will be effective in all schools since the participating schools are selected from one area, and the programme is delivered by the developers.

The primary outcome was reading ability as assessed by scores from the GL Assessment New Group Reading Test (NGRT). The secondary outcomes were the two NGRT subscales: sentence completion and passage comprehension. The tests were administered by the schools and therefore not blind, but the deliverer had no role in their administration.

Analysis was completed on an 'intention to treat' basis, reflecting the reality of how interventions are delivered in practice, followed by an 'on-treatment' analysis where data from teacher logs was used to determine the extent of each pupil's involvement with the intervention. On-treatment analysis allows for an estimate of 'pure intervention effect'.<sup>2</sup>

85% of the pupils were included in the final analysis. However, both control and intervention group suffered similar levels of attrition and further analysis shows there was no evidence of bias in pupil characteristics in either group.

---

<sup>2</sup> For more details on 'intention to treat' and 'on-treatment' analysis, see the Evaluation Glossary on EEF website: <https://educationendowmentfoundation.org.uk/our-work/the-eeef-approach-to-evaluation/glossary/>

This was a successful trial and the outcomes are secure.

### How much does it cost?

Rhythm for Reading training costs £5,000 per school. Schools need to renew their licence to continue accessing the online resources; this costs £500 per year. As the cost is per school and not per pupil, per pupil cost depends on the number of pupils who experience the intervention, delivered by either a specialist or a trained teacher. If we assume a total of 90 Year 7 pupils receive the intervention in the first year (half the size of an average secondary school cohort), this is a per pupil initial cost of £56, dropping to £6 per pupil in subsequent years providing staff capacity to deliver is maintained.

### Key conclusions

1. Rhythm for Reading had no significant impact on reading ability overall.
2. There was some evidence that Rhythm for Reading worked differently across the ability range. Analysis suggests that the intervention caused lower reading scores among the lower attaining pupils and higher scores among higher attaining pupils within the sample.
3. Previous evaluation work in combination with the process evaluation findings indicates that the intervention may be more suitable to be trialled and evaluated with more able younger pupils (ie below Year 7).

# Introduction

## Intervention

Rhythm for Reading is a programme which aims to improve children's reading ability by taking part in rhythm-based exercises such as stamping, clapping and chanting, while reading musical notation. In this evaluation, the intervention was delivered to Year 7 pupils who had not reached a secure Level 4 in English and/or reading at the end of Key Stage 2. They received weekly ten-minute sessions over a period of ten weeks between April and July 2013.

## Background evidence

A growing body of evidence supports the argument that pre-school children experience difficulty with processing rhythm and that this impacts negatively on reading behaviour (Anvari et al., 2002; Overy, 2008). There have been three pilot studies of Rhythm for Reading. The first was conducted with Year 6 pupils at Adamsrill Primary School in Lewisham between February and May 2012. A second took place at a secondary school in Kent with Year 7 pupils between November 2012 and February 2013. The third pilot took place in a secondary school in Bermondsey with Year 9 pupils in 2013. All three studies had a positive effect on reading behaviour in underachieving Year 4–6 pupils (Long, 2007; Long, 2008; Long and Hallam, 2013). A further evaluation of the rhythm-based approach was conducted with Year 4 pupils by Long, the developer, in 2009 and the findings showed substantial gains in reading comprehension for participants in the rhythm-based reading group but not for the comparison pupils who were based in another school. Long worked as a volunteer in all three schools. Since the pilots, these schools and others have purchased the programme, except for Adamsrill Primary School where Long continues to volunteer as a friend of the school, piloting the programme with Year 2 pupils as an early intervention strategy. The intervention is now being purchased for use by primary and secondary schools.

As the intervention has previously been evaluated by the developer using non-randomised allocation, the present study represents the first independent evaluation of Rhythm for Reading. Since it was a localised intervention delivered under ideal conditions, this stage of evaluation can be seen as an efficacy trial.

Rhythm for Reading does not link directly to existing government policy. However, it is advocated by the developer as a cost-effective method of teaching reading that is particularly successful with disadvantaged pupils. In April 2013, the Department for Education published a policy entitled 'Raising the achievement of disadvantaged children' and the programme aims to sit squarely within this agenda.

This evaluation is one of 24 projects funded from a £10 million grant dedicated to improving literacy levels of pupils who do not achieve Level 4 in English at the end of Key Stage 2. The grant was given by the Department for Education in 2012 and administered by the Education Endowment Foundation.

## Evaluation objectives

The rationale for conducting this trial was to determine whether Rhythm for Reading offers a cost-effective intervention to improve reading in pupils who have started secondary school with limited literacy skills. The impact evaluation sought to determine the effect of Rhythm for Reading on reading ability. Furthermore, it aimed to determine whether any improvements in attainment were influenced by the



National Curriculum reading level of the pupil, or whether a pupil receives the pupil premium.<sup>3</sup> The process evaluation assessed Rhythm for Reading in terms of fidelity to the programme intentions and its scalability.

### **Project team**

The programme was designed and delivered in its entirety by Dr Marion Long. For the purposes of this report, we refer to Dr Long as ‘the developer’. The evaluation team at NFER was led by Dr Ben Styles. Sally Bradshaw and Alix Godfrey assisted with the impact evaluation. Rebecca Clarkson and Katherine Fowler carried out the process evaluation.

### **Ethical review**

Headteachers consented to the trial being carried out within their schools. This consent was followed up by a letter to parents allowing opt-out consent. The pattern of headteacher consent followed by parental opt-out consent, as adopted for other EEF transitions trials run at NFER, was approved by NFER’s Code of Practice Committee on 23 January 2013.

### **Trial registration**

This trial has been registered: <http://www.controlled-trials.com/ISRCTN10404460/>

---

<sup>3</sup> As data on eligibility for free school meals is more easily obtainable, this measure was used as a proxy for receipt of pupil premium.

# Methodology

## Design

The project was run as a randomised controlled trial, with 419 individual Year 7 pupils across six secondary schools randomised at pupil level to two groups: intervention and a waitlist control. Pupils in the intervention group were to receive the Rhythm for Reading intervention for ten minutes a week, for ten weeks; pupils in the control group experienced their usual English teaching. This design was chosen since previous evaluations of Rhythm for Reading were developer led and not independent. Pupils were tested for reading ability both before and after the intervention.

## Eligibility

Year 7 pupils from an area of south-east London local to the developer, with a National Curriculum level of 4 or below in English and/or reading at the end of Key Stage 2, were invited to take part in this trial. Opt-out consent was sought from parents of pupils who met the eligibility criteria and had been selected for participation prior to randomisation.

## Intervention

Rhythm for Reading is a ten-week course of weekly ten-minute sessions in which children take part in rhythm-based exercises while reading musical notation. Up to ten children are taken out of lessons to participate in the exercises, which appear at first hand unrelated to reading. The sessions are delivered in the first instance by specialists who later train classroom teachers to deliver the intervention. The specialists are employed by the Rhythm for Reading organisation and trained by the developer. They have a professional musical background through both performing and teaching. The programme was designed by the developer prior to undertaking 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. The mechanism by which Rhythm for Reading works is not well understood, but the developer has suggested that it may help children to detect the natural rhythm and phrasing of prose, which supports quicker and more intuitive comprehension.

The intervention was delivered to six schools in south-east London that had not previously experienced it. The developer visited each school weekly to deliver the sessions between April 2013 and July 2013 when the pupils were in their final term of Year 7. While the intended delivery model is that school staff shadow specialists during the ten-week demonstration of the programme, for this evaluation the developer delivered all the sessions and then, together with members of her team, trained teaching assistants after the impact evaluation was complete. The control pupils experienced their normal curriculum.

## Outcomes

The New Group Reading Test (NGRT; GL Assessment<sup>4</sup>) was used to measure reading ability. At baseline, the digital version of the test was used, while at follow-up the paper version was used. The digital NGRT requires the use of headphones; however, schools sometimes lack a full set of functioning

---

<sup>4</sup> <http://www.gl-assessment.co.uk/products/new-group-reading-test>

headphones so are unable to administer the test properly. For this trial, all schools used headphones for the baseline test and the developer reported no feedback from schools experiencing technical difficulties. However, because of technical problems in other trials and some reservations held by the evaluator about the validity of the digital test, the switch was made to paper for follow-up. The digital test is adaptive – the items that each pupil sees are dependent on whether they have got previous items right or wrong. It thus covers a greater age and ability range than individual paper tests. In particular, if a pupil is struggling with the sentence completion items, the test defaults to phonics items that are not present in the equivalent paper version of the test. When this feature of the test became apparent, concern was raised over whether the follow-up test would be measuring what was originally intended: a reading score made up of sentence completion and passage comprehension items. Along with all other transition trials run by NFER, the decision was made to move to paper testing in order to be confident in the measurement of outcomes at the end of the trial.

The NGRT has parallel forms; form A was used at baseline and form B at follow-up. The NGRT has two subscales, sentence completion and passage comprehension, which can be combined into a composite reading score.

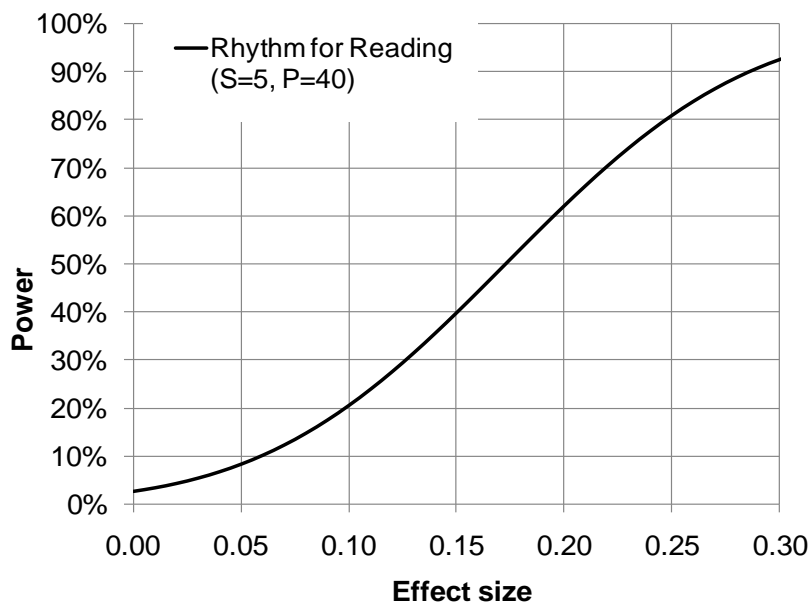
The composite reading score was used as the primary outcome. The two subscales were used as secondary outcomes. These outcomes were chosen since one aim of the intervention was to improve reading ability and the NGRT is a reliable test that has been standardised for the age group in question.

In order to minimise bias, the developer was not involved in data collection other than collating feedback from schools on how the process went, on behalf of the evaluator. At baseline, school staff invigilated while pupils took the digital tests. Results were calculated using GL Assessment's online system, and were accessed through GL Assessment's online platform following testing. School staff took responsibility for organising the paper testing, with completed scripts being sent to GL Assessment for blind marking.

While the developer was not involved in data collection, complete blinding at test administration cannot be assured as this was the responsibility of school staff, who were likely to have been aware of the intervention taking place and the allocation of pupils to the intervention group.

## Sample size

Figure 1: Power curve



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

Randomisation was conducted at a pupil level, and variation in baseline scores was controlled for in the final analysis. Intra-class correlation ( $\rho$ ) was therefore likely to have a minimal impact on the effective sample size, and so we conservatively assumed a value of  $\rho=0.02$  for the purposes of our calculations. We also assumed a correlation of 0.75 between baseline and follow-up scores on the basis of previous work with reading tests. The power curve in Figure 1 illustrates that a sample size of 400 pupils should be sufficient to detect effect sizes of the order 0.25. This could be considered moderate, equivalent to around 3 months of progress and quite reasonable for targeted interventions providing support to small groups of pupils.

During recruitment, the developer was able to recruit an extra school and a total of 419 pupils were randomised. This allowed for 5% attrition which was deemed adequate to account for pupil absence.

## Minimum detectable effect size (MDES)

Once all the data from the trial was available, the assumed parameters from the above calculations were compared to the actual parameters and included in a calculation of MDES.

Randomisation was carried out at the pupil level thus cancelling out the effect of clustering when estimating internally valid uncertainty around the effect.  $\rho$  can hence be regarded as zero. The adjusted R-squared for the primary outcome model without the intervention term was 0.46, implying a value of 0.68 would have been more appropriate for the correlation between baseline and follow-up scores. Using the actual number randomised, this yields an MDES of 0.2 at 80% power.

## Randomisation

The developer and the individual schools involved in the trial were responsible for pupil recruitment. Randomisations were carried out by a statistician at NFER using a full syntax audit trail within SPSS. Randomisation was stratified by school: simple randomisation of pupils into two groups of the same size was carried out within each school. This was necessary to help timetabling of the sessions within schools.

Schools required the results of the randomisation prior to baseline testing for timetabling reasons, and so these were released directly to them through NFER's school portal. The developer was not provided with the results of the randomisation until after baseline testing had occurred.

## Analysis

The primary outcome was reading ability as assessed by raw scores from the NGRT. Raw scores were used in preference to age-standardised scores because of potential ceiling or floor effects in the latter. Sub-group analysis on the primary outcome was carried out on the following groups only: pre-test score, National Curriculum level and whether or not a pupil was eligible for free school meals (FSM). The secondary outcomes were the two NGRT subscales: sentence completion and passage comprehension. All outcomes and subgroup analyses were pre-specified at the protocol stage.

The definitive analysis was 'intention to treat', reflecting the reality of how interventions are delivered in practice. It was necessary to take school context into account in the analysis owing to the restricted nature of the randomisation (Kahan and Morris, 2012). Five dummy variables were included in the regression model to represent school; one school was the default category. The definitive primary outcome analysis regressed post-test raw score on pre-test score, school, randomised group, sex, FSM and age in months. The pre-test was the digital version of NGRT, which is adaptive. A scaled score was hence used to represent reading ability at pre-test. This takes into account both the difficulty of the items attempted and whether they were answered correctly. Subgroup analysis was carried out using a separate regression model on FSM pupils and by exploring the interaction between randomised group and pre-test score. Secondary outcomes were analysed using raw scores in the relevant domains in place of overall reading scores.

The main analysis was followed by an 'on-treatment' analysis where data from intervention logs was used to determine the extent of each pupil's involvement with the intervention. For each school's intervention group, the total delivery time was divided by the total number of intervention pupils to give a measure of dosage. The dosage measure was then used in place of the dichotomous group allocation. There was no evidence of contamination so control pupils were allocated a dosage of zero. This analysis allows for an estimate of '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>5</sup>

## Process evaluation methodology

The process evaluation covered the intervention from session observations in May 2013 to training at the end of November 2013.

Information was collected from the following sources: observations of intervention sessions in situ, reviewing the qualitative parts of the 'intervention logs' completed by the sole deliverer, a review of the training materials and observations of two training sessions offered to teaching assistants in one school,

---

<sup>5</sup> For example, pupil motivation may be positively related both to levels of exposure to the intervention (through better attendance) and to the amount of progress made between baseline and follow-up testing.

and a group interview with two of the three<sup>6</sup> teaching assistants who had attended the training. Training had to be postponed due to other school commitments and some schools wanted to delay training until trial results were known. The first school that was able and willing to timetable training was visited. Further information was also gathered in informal discussions with the delivery partner. These methods were chosen to allow for coverage across different parts of the intervention.

A team of NFER researchers conducted the process evaluation. Where more than one researcher was observing training and sessions, and undertaking interviews, frequent team meetings were held to share information and plan next steps. All researchers have contributed to the report writing to ensure full coverage of information gathered.

A detailed observation schedule was developed to ensure that comparative information was gathered from each of the sessions observed. Part of the 'intervention log' was also used to capture information about what was happening in each of the sessions. This included a space to record a brief outline of each session and a section to detail whether any deviation from the programme had occurred.

---

<sup>6</sup> The third teaching assistant was ill on the day.  
Education Endowment Foundation

# Impact evaluation

## Timeline

**Table 1: Timeline**

Month (2013)	Activity
January – February	Recruit and gain consent from schools and pupils
March	Random allocation of pupils and pre-testing (20–27 March) <sup>7</sup>
April	Implementation of intervention programmes
May	Implementation of intervention programmes
June	Implementation of intervention programmes Post-testing (26 June – 9 July)
November	Training of teaching assistants; control group pupils begin to experience intervention

## Participants

All schools involved in this project were in an area of south-east London classified as urban. The schools (and therefore the location) were selected because they were secondary destinations for Adamsrill Primary School pupils, consistent with the theme of Literacy in Primary to Secondary Transition. The developer is a volunteer at Adamsrill, her local primary school. Further details are provided in Tables 2–4. Schools were recruited to the project by the developer and were required to sign a memorandum of understanding (see Appendix). The developer visited local schools after booking meetings with key school staff. Recruitment, and the acquisition of headteacher and parental consent, was undertaken in January 2013. A timeline including recruitment, testing and intervention implementation can be found in Table 1.

1209 pupils joined Year 7 in the six schools in the school year 2012/13.<sup>6</sup> Of these, 419 pupils were deemed eligible for the study on the basis of Key Stage 2 results and were randomised to the intervention or control groups.

**Table 2: Ofsted ratings for schools**

Ofsted rating of overall effectiveness of the school	Number of schools
Good	3
Requires Improvement	2
Inadequate	1

<sup>7</sup> All schools tested in March, apart from one which was delayed until after the Easter holidays due to staff illness. Schools needed to know the results of the randomisation before pre-testing for timetabling reasons. The results were not released to Dr Long until after pre-testing.

**Table 3: School type**

School type	Number of schools
Comprehensive 11–16	2
Comprehensive 11–18	4

**Table 4: Pupils eligible for FSM**

Pupils eligible for FSM	Number of schools
Highest quintile	4
Second highest quintile	2

**Table 5: School attainment**

GCSE performance band (2012)	Number of schools
Second lowest quintile	1
Lowest quintile	5

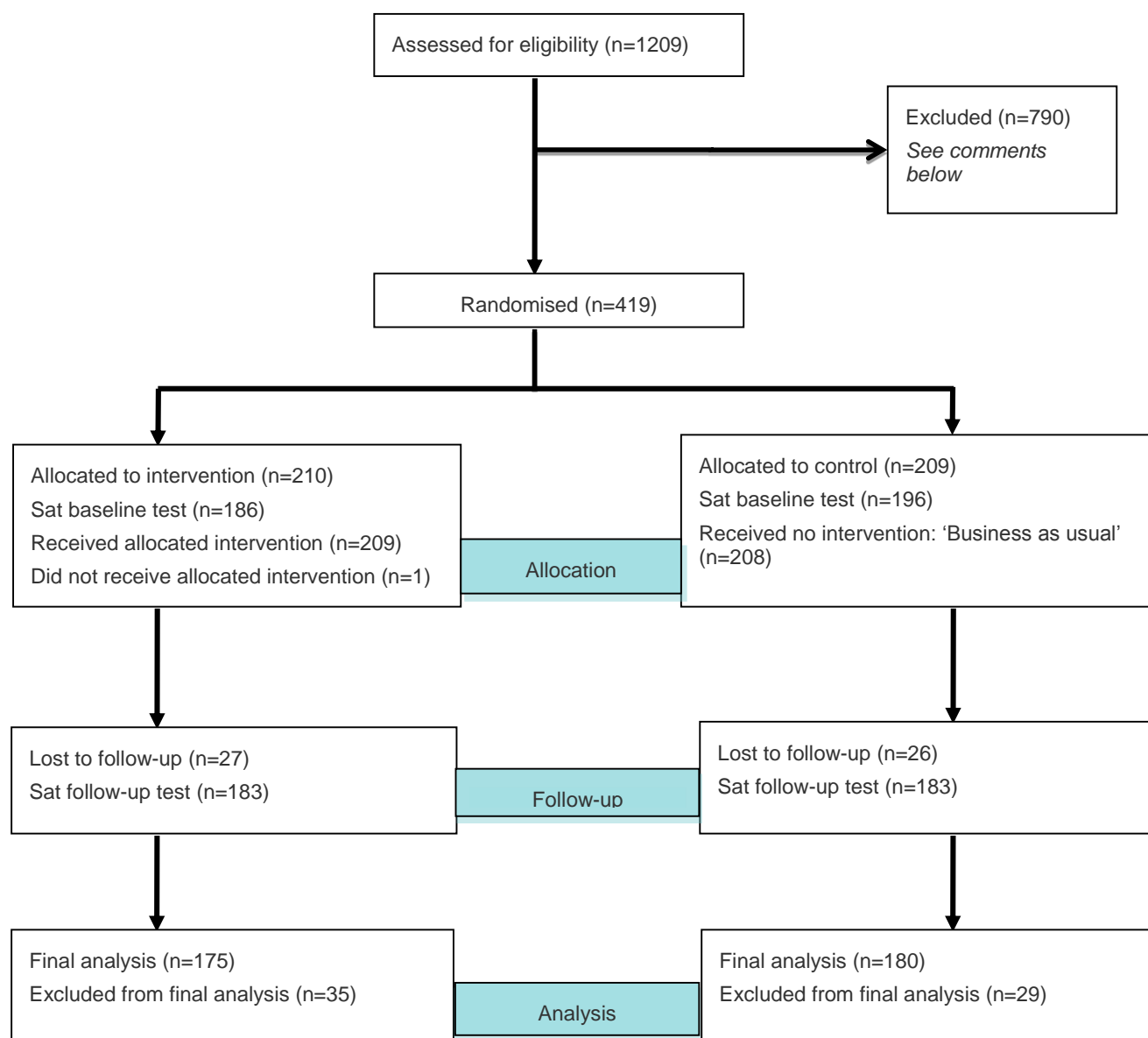
All schools involved were urban secondary schools in south London. Their Ofsted ratings vary: three received *Good* ratings, two received *Requires Improvement* and one was rated *Inadequate*. Two of the schools were single-sex schools (one girls' school and one boys' school), while the other four were mixed. The two single-sex schools had a joint, mixed-sex, sixth form. All schools were ethnically diverse, with the number of ethnicity categories ranging from 16 to 18 (out of a maximum of 18). Four schools were in the highest quintile of schools for FSM eligibility; two were in the second highest quintile. Five of the six schools were in the lowest quintile for school attainment, based on their GCSE performance band. The sixth was in the second lowest quintile.<sup>8</sup>

---

<sup>8</sup> All information taken from 2011–12 data  
Education Endowment Foundation



Figure 2: Participant flow diagram



Based on information provided by the developer, two pupils withdrew completely from the study, one for religious reasons and another left the school. These are considered as 'not receiving allocated intervention' in the flow diagram in Figure 2. The remaining pupils were put forward for baseline and follow up testing regardless of their cooperation with the interventions. .

Individual schools themselves screened for eligibility on the basis of Key Stage 2 results and only provided data on those pupils who consented and were randomised. Schools did not provide the numbers of parents who opted out. Therefore, although we can be confident that the vast majority of excluded pupils were ineligible, we do not know how many did not take part through parental opt-out. Pupils who did not sit both pre- and post-tests were excluded from the final analysis, owing to inability to compare their results at the two time periods.

## Pupil characteristics of analysed groups

While we expect no systematic bias to have arisen from randomisation, bias may have occurred due to attrition. Chi-squared tests on all three background factors presented in this section revealed no significant differences between groups for the data after attrition.

**Table 6: National Curriculum level in reading at baseline ( $\chi^2= 7.6$ ,  $df=4$ ,  $p=0.11$ )**

National Curriculum level (source: baseline NGRT)	Intervention group		Control group	
	Frequency	Percentage	Frequency	Percentage
1 or below	10	6	21	12
2	45	26	38	21
3	58	33	64	36
4	60	34	51	28
5 or above	2	1	6	3
Missing	0	0	0	0
Total	175	100	180	100

Table 6 shows that the vast majority of pupils satisfied the eligibility criteria as applied by schools. It should be noted that schools were using Key Stage 2 results to determine eligibility criteria so some improvement may have occurred since then.

**Table 7: FSM eligibility ( $\chi^2= 0.18$ ,  $df=1$ ,  $p=0.67$ )**

Pupil eligible for FSM (source: NPD)	Intervention group		Control group	
	Frequency	Percentage	Frequency	Percentage
Yes	66	38	60	33
No	103	59	103	57
Missing	6	3	17	9
Total	175	100	180	100

**Table 8: Sex ( $\chi^2= 0.07$ ,  $df=1$ ,  $p=0.79$ )**

Sex of pupil (source: schools via GL Assessment)	Intervention group		Control group	
	Frequency	Percentage	Frequency	Percentage
Male	119	68	120	67
Female	56	32	60	33
Missing	0	0	0	0
Total	175	100	180	100

Further to pupil background measures, it was also important to test whether significant imbalance at pre-test had ensued as a result of attrition. The baseline effect size was 0.14 (-0.07, 0.35) and was not significant ( $p=0.12$ ). It was less than the effect size that was generated by randomisation alone before attrition.

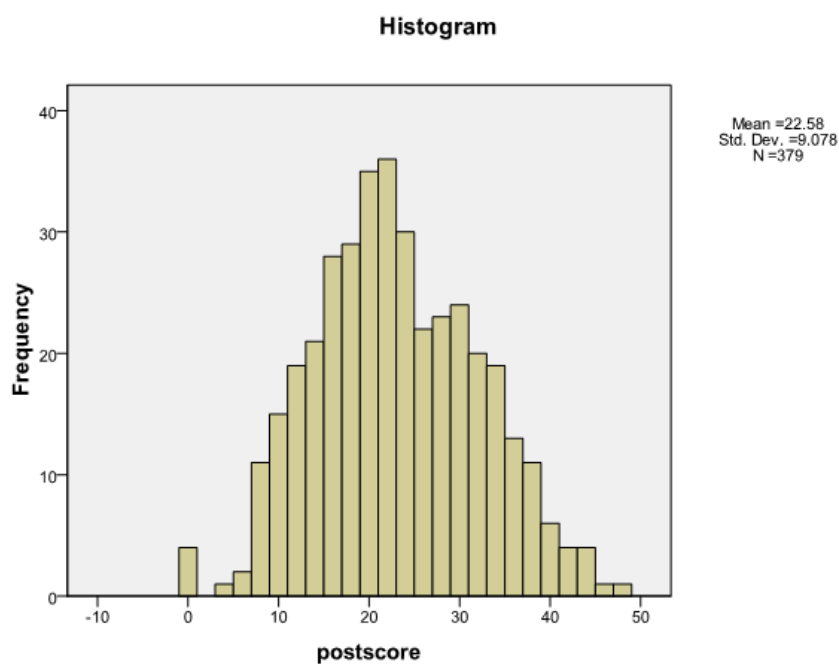
## Outcomes and analysis

Table 9: Effect size

Outcome description	Outcome measure	Effect size (Hedges' g)	95% confidence interval (lower)	95% confidence interval (upper)	p	Number of intervention pupils in model	Number of control pupils in model
Primary	Reading score (NGRT)	0.03	-0.13	0.18	0.745	175	180
Primary (FSM)	Reading score (NGRT)	0.11	-0.16	0.37	0.429	66	60
Secondary	Sentence completion score	-0.09	-0.25	0.08	0.288	175	180
Secondary	Passage comprehension score	0.10	-0.08	0.27	0.282	173	172

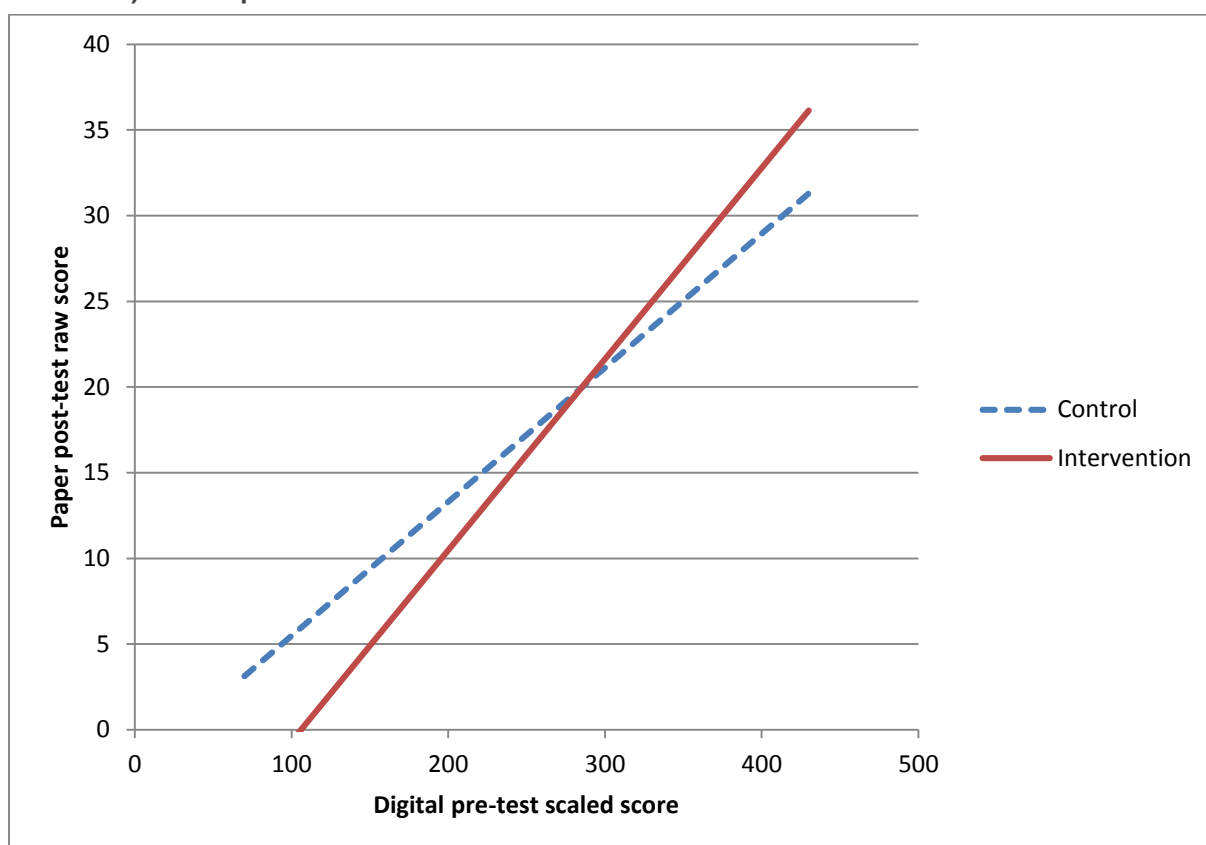
An ANOVA of post-test reading score by randomised group showed no significant impact of the intervention ( $F=0.56$ ,  $p=0.45$ ,  $n=366$ ).

Figure 3: Distribution of scores on the post-test



All outcomes analysed were pre-specified in the protocol. All sub-group analyses were pre-specified in the protocol apart from the use of FSM as a proxy for pupil premium; a separate FSM analysis is a requirement of all EEF evaluations. Background data on pupils was obtained both from schools through the standard GL Assessment data form and from the National Pupil Database (NPD). Where the same variable was obtained, the latter source was used in preference. The main primary outcome analysis consisted of a regression model containing the following variables: pre-test score, intervention group and school. FSM, age in months and sex were included in an initial run of the model but none was significant so all were excluded. The use of a linear regression model represents a slight deviation from the protocol where it was specified that a multi-level model would be used. Because randomisation was restricted by school, it was necessary to account for school in the model but this was done using dummy variables rather than in a multi-level model. This method was used in preference since a multi-level model would not have estimated variance adequately with only six schools. As this was a pupil-randomised trial, clustering for the intervention effect is not an issue in terms of internal validity. Model results are presented in the appendix.

**Figure 4: Interaction between intervention and pre-test score; post-test score (as predicted by the model) versus pre-test score**



The protocol specifies an analysis by both National Curriculum level and pre-test score. Both were carried out and a significant positive interaction was observed with pre-test score ( $p < 0.01$ ). As expected, analysing the data by the more crude measure of National Curriculum level did not reveal any significant effects. Figure 4 is a model showing the expected post-test performance based on pre-test scores for the default school. The model suggests that children of lower ability (below 320 on the pre-test) are likely to perform worse in the intervention group than children of similar ability in the control group. For children of higher ability (above 320 on the pre-test) in the intervention group, the model suggests that they are likely to perform better than children of similar ability in the control group. Figure 4 is a representation of how the intervention might impact on children of varying ability. The lines cross at a point just below the pre-test score corresponding to the Level 3/Level 4 boundary (320). This is a theoretical pre-test score at which the intervention might begin to have an impact, but in reality we cannot say for certain that Rhythm for Reading was detrimental for a certain ability group and effective for another. The important thing to note is that the gradients of the lines are significantly different.

All the above analysis was ‘intention to treat’. The ‘on-treatment’ analysis used a measure of intervention experienced by each pupil in terms of equivalent minutes of one-to-one tuition (see ‘Analysis’ section on page 11). The intended dosage was 10 and this was achieved in three schools. In the remaining three schools, the dosage was 1, 17 and 19 reflecting varying amounts of time available within the school day for intervention delivery. The ‘on-treatment’ analysis revealed no significant effect of dosage ( $p = 0.669$ ).

## Cost

Schools that wish to purchase Rhythm for Reading would need to pay £5,000. In the first term, this would include a ten-week demonstration led by a Rhythm for Reading specialist for up to 60 pupils. Teachers from the school who would like to be trained attend a session each week.<sup>9</sup> Schools would be able to select 12 pupils for measurement of the effectiveness of the programme (using the Neale Analysis of Reading Ability assessment) and would receive a report at the end. In term one, teachers would experience two training events and receive a pocket-sized handbook with key messages about the programme, together with login details for the Rhythm for Reading teacher forum and Rhythm for Reading online resources. During term two, teachers may request face-to-face refresher sessions, but the forum provides the main source of support for delivery. Schools may request that reading measurements be taken again and a report confirming that the training has been successful be issued. At the end of term three, schools would need to renew their licence to continue accessing the online resources, at a cost of £500. Subsequent annual renewal also costs £500 a year.

A school would need to display the resources for each session on an interactive whiteboard or projector screen. Staff attend demonstration sessions for a minimum of ten minutes per week for ten weeks. The two training events are 30–40 minutes in length and could be delivered as twilight sessions. A small amount of planning time for each session also required (a maximum of half an hour, according to the developer) as all resources are available online and organised into session delivery.

The schools that took part in the evaluation were not required to pay a fee. There was no supply cover needed as the sessions were delivered by the developer for the duration of the trial. Teaching assistants subsequently received training during their free periods.

As the cost is per school and not per pupil, per pupil cost depends on the number of pupils who experience the intervention, delivered by either a specialist or a trained teacher. If we assume a total of 90 Year 7 pupils receive the intervention in the first year (half the size of an average secondary school cohort), this is an initial per pupil cost of £56, dropping to £6 per pupil in subsequent years providing staff capacity to deliver is maintained.

---

<sup>9</sup> Note it is the intention that teachers are trained rather than teaching assistants

## Process evaluation

This programme followed an unconventional path as it did not start with training sessions. The entire programme was delivered by the developer in each school. Initially she had expected the help of a colleague, whom she had trained, to deliver some of the programme sessions; however, the training sessions organised by the school did not coincide with the colleague's availability. Teachers and teaching assistants interested in the programme (through observing the sessions, being the contact within the school, or through word of mouth from other colleagues or the pupils taking part) were invited to take part in training after the programme was completed. They were then encouraged to use the knowledge from the training and the accompanying resources to continue to offer Rhythm for Reading sessions in their schools.

### Session observations and intervention logs

Eight Rhythm for Reading sessions were observed in two secondary schools. These sessions were observed on the same day (21 May 2013), with each session taking approximately ten minutes to deliver. One session lasted only five minutes; it was stopped early due to disengagement and poor behaviour of the pupils in the group. To a large extent, Key Stage 3 pupils cooperated and engaged with Rhythm for Reading when it took place during lesson times. When sessions were scheduled during early morning tutor time, pupils were far more restless and easily distracted. Some pupils resented being involved in the intervention, sensing that they were being targeted. For others, spending time with their tutor group was important and they simply did not want to miss out.

There was concern among some teachers that the post-test paper version was too difficult for the lowest ability children to access. Unlike the adaptive digital version, the paper test has a specified ability range. To allay concerns of floor effects operating, a distribution of post-test scores was produced (Figure 3). The distribution of scores shows a normal distribution, suggesting that no floor effects were operating, and the paper test catered for children across the ability range.

The delivery of the sessions was fast-paced and concentrated, which increased the engagement and motivation of most of the pupils in the groups. A maximum of ten pupils attended each session. The session invariably happened in a music room or an ICT suite, ie a room that was quite large but unused at the time the sessions were taking place. Pupils were invited to sit on chairs in a semi-circle, where they could see an interactive whiteboard. The developer used the whiteboard to display musical notation and played an audio file, which the pupils were encouraged to chant and stamp along to. All the audio files were pre-recorded and the musical notation pre-written, with a number of various compositions available for each session. These are downloaded from the Rhythm for Reading website.

All the observed sessions followed a similar schedule and consisted largely of the same tunes, suggesting that each group was at the same stage in the programme, and therefore the intervention, as observed, was being adhered to faithfully.

Information from the intervention log suggests that the intervention was faithfully delivered in three of the participating schools, with ten pupils in a group receiving ten sessions of ten minutes each. In one school, pupils did not receive a faithful delivery of the intervention owing to timetable constraints imposed upon the deliverer by the school. As certain teachers did not want pupils removed from their lessons, this meant delivery had to happen during a restricted time interval. In this school, group sizes ranged between 15 and 20 pupils and only nine sessions were delivered.

### Training observations

Two training sessions were observed. The first training session consisted of a one-hour session held at one of the secondary schools taking part in the trial, in October 2013. The session was conducted by the developer. The training was attended by three teaching assistants employed by the school where the training took place. The three teaching assistants were given a small, ring-bound booklet that acted as an agenda for the training session and could be used as an aide-memoire in the future.

The concept and core values of Rhythm for Reading were introduced, followed by a description of how it was developed and the research evidence that had fed into this development. As the trial had already taken place, the teaching assistants were told about this and that they would be working with the control group, who were on a waitlist. The developer demonstrated what a session would consist of.

The teaching assistants were fully engaged throughout the hour's delivery. The intimate and informal atmosphere of the training allowed them to ask many questions and seek further clarification. All seemed enthused after the session and were looking forward to participating in the delivery for the waitlist control pupils.

Following this short, overview training session, the three teaching assistants attended ten Rhythm for Reading intervention sessions. First, they observed how a session is delivered, then, in tandem with the Rhythm for Reading instructor, they delivered some of the sessions. After the ten sessions they were expected to deliver Rhythm for Reading sessions on their own.

The eighth training session of the series of ten was observed in November 2013. Three intervention sessions were happening concurrently, each having a different deliverer from the developer's team and one of the teaching assistants as trainee, ie one-to-one, on the job, training. One of the teaching assistants was absent for this training session. The session observed was being delivered to one of the Year 8 groups who had been on the waitlist. As this session was number eight out of ten, it was expected, as in the training plan outlined above, that the teaching assistant trainees would be delivering alongside the deliverer. In essence, the teaching assistant trainees observed and spent some time at the beginning of the session encouraging the pupils to settle down. Considerable time was spent encouraging one boy to join in. Throughout the ten-minute session much of the teaching assistant trainee's time was spent on behaviour management.

After the ten training sessions there is an expectation that the developer will go through the website with the trainees, so they know where to access resources. In discussion with the developer after the training session, it was made clear that the trainees would be able to have further tuition for their first few sessions of solo delivery, should they require it. The trainee teaching assistants mentioned that they would benefit from an evaluation follow-up session at the end of the training to assess how the training had gone. It was not clear that this was something habitually offered to trainees.

This is a fairly long training process when compared to other schemes; however, as the sessions will be delivered by a Rhythm for Reading expert to a group of children who the school would like to receive the intervention, and the school will have staff trained at the same time, this does not seem a great impediment. In a wider roll-out, schools would be able to follow this model of having the scheme delivered with staff training happening concurrently.

### **Group interview**

Evaluation staff from NFER interviewed the two teaching assistants who attended the second observed training session (the third teaching assistant was absent). The interview was conducted as a small group, rather than individually, due to time constraints of the interviewees. It took approximately one hour to complete.

The interview covered questions and topics on the training, potential delivery of the intervention, resources needed, issues of cross-contamination, and other issues such as perceptions of scalability. The responses have been used to inform the detail involved in the sections describing the training and intervention sessions in schools, and further information from the interviews has been used to inform sections below.

## **Implementation**



A key aspect of the success of Rhythm for Reading appeared to be the brevity of the sessions, which enabled most pupils to maintain motivation. The teaching assistants also mentioned that having a musical background was also beneficial for deliverers.

The main barriers to implementing Rhythm for Reading were behavioural problems within the sessions and the teaching assistants' perceived lack of musical knowledge. Time spent managing behaviour meant that some pupils did not get the full benefit of the activities. The teaching assistants also felt they would need extra practice before they would feel comfortable delivering sessions.

## Fidelity

As the developer delivered all the intervention sessions in all the schools, it seems there was a high level of quality control. As mentioned in the 'Session observations' section (page 20), delivery was variable in the different schools, although the developer was able to tailor each session to the needs of the pupils due to her intimate knowledge of the programme. The website, which outlines the structure of the ten sessions, was clearly laid out, assuming that future deliverers would follow this structure.

When interviewed, the teaching assistants felt that they did not understand the structure of the intervention and how it all fitted together. It was not always clear why, in a session, they would move from one song to another. They wondered how the developer and the other training deliverers knew how to choose the next song, as the structure of the ten sessions on the website was not always closely followed. It transpired that the trainees had not seen the website at the time and had watched the deliverers use resources on a USB which, unlike the online resources, does not have a menu and weekly programme.

It appeared that delivery could be group dependent, with groups doing different songs at the same time point. This would indicate a mismatch between the online resources, which have a structured week-by-week approach, the expectation that this is what the teaching assistants would deliver and the expertise of the training deliverers. Both teaching assistants could foresee that they would need to spend some time doing extra preparation to feel comfortable delivering the sessions.

## Outcomes

The teaching assistants were asked whether they had seen any benefits from pupils taking part in the sessions. It should be reiterated at this point that the training sessions were being conducted in the term after intervention delivery with Year 8 pupils, who were part of the waitlist control, rather than Year 7. One teaching assistant was not convinced that Rhythm for Reading was a worthwhile experience, but mentioned that they would like to see it in operation with younger children. The teaching assistant reported that they were failing to see how it was working. The pupils did not seem to trust that the programme would work and did not really understand why it would work, which the teaching assistant thought would be a less problematic issue for younger children.

Both teaching assistants thought it was hard to tell if pupils were making progress within the sessions, eg are they looking at the right note or are they just repeating what the deliverer says without knowing? Teaching assistants had not observed any assessment in the sessions, although one of the trainers mentioned being silent for a few bars of music to see what the pupils would do/whether they knew what to do (this was in an unobserved training session). They did not know whether the intervention had had an impact on the literacy ability of the pupils involved.

No unintended consequences or negative effects of the intervention were mentioned by the interviewees, nor were they observed in the sessions.

## Scalability

On the whole, though, the teaching assistants were positive about the intervention and felt it would be popular. They both had interest shown by other members of staff in the school; one teacher was particularly enthusiastic and wanted to sign up for training.

Implementing the intervention, post-training, would be fairly simple. As the sessions are only ten minutes long they could be fitted into spaces in the school day. The sessions could be carried out with a larger group or with a whole class, and would be suitable as a starter or plenary activity.

Few resources are needed for the implementation of Rhythm for Reading and those that are needed are easily downloadable from the website. Deliverers did mention that on occasion there had been technical issues in some of the sessions, but these were primarily to do with school computer systems and a weak Wifi signal, not with Rhythm for Reading online resources.

One teaching assistant mentioned that having no musical knowledge was an impediment. They felt that they needed to learn all the notes and structure before being confident in delivery, and that this would take extra time outside the training. It may be easier for someone with a basic knowledge of music to learn the delivery.

## Control group activity

The teaching assistants were probed about whether there had been any cross-contamination between the intervention group and the control group. Both teaching assistants believed that this had not happened. They felt it was unlikely that Year 7 or 8 pupils would share this experience outside of the sessions, but felt it may be more likely for primary pupils.

## Recommendations for improvements

### Address behaviour

During the session observations, it was clear that poor behaviour was common among the pupils taking part. The deliverers had not been given any form of behaviour management training. If an external agency were to be used, they would need thorough training in behaviour management. Alternatively, it is likely that behavioural issues would not cause such a disturbance if the sessions were led by members of school staff.

### Target appropriate age group

Comments from the teaching assistants and the deliverers indicated that Rhythm for Reading, in its current format, would be suitable for pupils in Year 7 and below, with primary-age pupils benefitting the most. The Year 8 pupils in the waitlist control session found the intervention '*babyish*'.

### Length of session

Children need to be at the session promptly for the session to begin on time. Many sessions were hindered by pupils arriving late, settling themselves, etc. This could easily be managed by adding a couple of minutes to the beginning of the session, ie making a session 15 rather than ten minutes long, with the expectation that Rhythm for Reading would then be able to have concentrated delivery for ten minutes.

## Conclusion

### Limitations

One major limitation of this study was that the developer delivered all of the sessions so there was no exploration possible of how effective the intervention would be when delivered by a variety of practitioners. The training of teaching assistants that took place after the evaluation was largely well executed and this could happen on a larger scale if needed.

85% of pupils that were randomised were included in the final model. For a pupil-randomised trial an attrition level of 15% is average. The final analysis required pupils to have sat both baseline and second follow-up tests and the highest dropout observed at a single testing sweep was 13%. A completers analysis was performed and there is a possibility this was open to bias due to the extent of missing data. However, multiple imputation and sensitivity analyses (Torgerson and Torgerson, 2013) were beyond the scope of this evaluation. In an attempt to reduce attrition, a further analysis was carried out replacing pre-test score with raw score in Key Stage 2 reading. Unfortunately, due to the imperfect match with NPD this resulted in fewer cases being used in the analysis. Furthermore, the correlation between Key Stage 2 reading and post-test was only  $r=0.36$  as compared to  $r=0.62$  for the pre-test to post-test correlation.

While internally valid, this trial has little external validity from a strict statistical perspective. The schools recruited into the trial were local to the developer and therefore were not randomly selected. Consequently the sample cannot be said to be representative of any meaningful population of schools.

There was no evidence of contamination and, while we cannot be sure tests were administered completely blind, administrators had no role in intervention delivery.

The change from digital to paper testing between baseline and follow-up could be regarded as a limitation of the study. When analysing the results of an RCT with a baseline measure, the most unbiased approach is to treat the baseline as a covariate in either ANCOVA or regression (Senn, 2006). The role of the baseline in such a model is predominantly to explain outcome variance and hence increase power. As EEF's paper on pre-testing discusses, the measure need not be the same as, or a parallel test to, the post-test. As long as it correlates well with the post-test, it serves its function. In this case, the digital pre-test can almost be regarded as a parallel test to the post-test since it contains largely the same items as the paper form A of the test. The correlation between the digital pre-test and paper post-test was  $r=0.62$ ; thus the pre-test explained approximately 38% of the post-test variance. This was sufficient for the power of the analysis to be retained at the level intended (see 'Minimum detectable effect size' section on page 9).

### Interpretation

Whether a more sophisticated treatment of missing data would have changed the results is doubtful since the baseline characteristics of analysed groups did not differ significantly. This is indicative of an unbiased attrition. Furthermore, it seems unlikely that a trial run in a different area of the country would reveal anything strikingly different in terms of impact since there is no region-specific aspect to the intervention. The above limitations should therefore be seen in the context of fairly conclusive impact results. Rhythm for Reading did not have a significant impact on reading ability overall. Furthermore, taking into account the large differences between some schools in the extent to which the developer was able to deliver the programme, there was also no significant impact.

The significant interaction between intervention and pre-test score is difficult to interpret. It is not the intention of the programme to work in this way and the result certainly cannot be used to justify Rhythm for Reading use in a particular ability group without further research. The Hawthorne effect cannot be

excluded here. The trial was originally conceived to have three arms, with one containing some group activity for an active control. Unfortunately, this design was not possible within the budget available and a two-arm trial was used. Rather than the intervention itself leading to the different relationship between pre-test and post-test, pupils of different abilities could have been reacting differently simply to being involved. Equally, the interaction could represent a genuine effect of the intervention.

Since delivery was exclusively by the developer, administering the intervention on a larger scale would rely on others being able to deliver to the same standard. Whether a similar interaction between intervention and pre-test would be observed in an effectiveness trial is open to debate since the conditions of the trial would be quite different.

This intervention was built on evidence that experiencing difficulty with processing rhythm impacts negatively on reading behaviour (Anvari et al., 2002; Overly, 2008). As such, evidence can only be of a correlative nature. This evaluation attempted to demonstrate a causal impact. This trial has not been able to conclude that the Rhythm for Reading programme improves the processing of rhythm and reading, despite it being the most robust evaluation design employed to date.

## **Future research and publications**

The process evaluation findings suggest that this intervention may be more suited to younger pupils. Furthermore, much of the earlier evaluation work was done on younger pupils. The significant positive interaction between pre-test score and intervention in combination with the process evaluation findings suggests a further trial on younger, more able pupils may be warranted.

## References

**Anvari, S.H., Trainor, L.J., Woodside, J. and Levy, B.A. (2002).** Relations among musical skills, phonological processing and early reading ability in preschool children, *Journal of Experimental Child Psychology*, 83, 111–130.

**Kahan, B.C. and Morris, T.P. (2012)** Reporting and analysis of trials using stratified randomisation in leading medical journals: review and reanalysis, *BMJ* 2012, 345, e5840.

**Long, M. (2007).** The effect of a music intervention on the temporal organisation of reading skills. Unpublished PhD, Institute of Education, University of London.

**Long, M. (2008).** A rhythmic view of reading: From word recognition to reading comprehension, [www.ippr.org/index.php?option=com\\_ippr&view=publications&old\\_id=631](http://www.ippr.org/index.php?option=com_ippr&view=publications&old_id=631)

**Long, M. (2014).** ‘I can read further and there’s more meaning while I read’: An exploratory study investigating the impact of a rhythm-based music intervention on children’s reading, *Research Studies in Music Education*, Volume 36 Issue 1, pp. 99 - 116.

**Long, M. and Hallam, S. (2012)** Rhythm for Reading: A rhythm-based approach to reading intervention, Proceedings of Music Paedeia: From Ancient Greek Philosophers Toward Global Music Communities, pp. 221-232, 30<sup>th</sup> ISME World Conference on Music Education, 15-20 July, 2012, Thessaloniki, Greece.

**Overy, K. (2008).** Classroom rhythm games for literacy support. In T. Miles, J. Westcombe and D. Ditchfield (eds), *Music and dyslexia: A positive approach*, 26–44. Chichester, England: John Wiley and Sons.

**Senn, S. (2006)** Change from baseline and analysis of covariance revisited, *Statistics in Medicine*, 25, 4334–4344

## Acknowledgements

The evaluation team is sincerely grateful to Dr Marion Long who recruited schools, managed communication with schools, kept NFER up to date with progress of the trial and provided material for the ‘Background evidence’ section (page 5). At NFER, the statistical analysis was undertaken with the help of Sally Bradshaw and Alix Godfrey. Simon Rutt commented on a final draft of the report. We are grateful for the comments and additions provided by Camilla Nevill and James Richardson at EEF, two anonymous referees and Dr Long.

## Appendix A: Model Results

Results of main effect model:

		Coefficients <sup>a</sup>				
		Unstandardised Coefficients		Standardised Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	-5.403	2.133		-2.534	.012
	pre-test score	.090	.006	.591	14.913	.000
	intervention	.229	.704	.013	.325	.745
	school1	-2.942	1.927	-.072	-1.527	.128
	school2	1.724	1.304	.088	1.322	.187
	school3	-.897	1.609	-.029	-.557	.578
	school4	5.506	1.302	.281	4.227	.000
	school5	-.371	1.435	-.015	-.258	.796

a. Dependent Variable: post-test score

Neither female, FSM nor age in months was significant so these were excluded from the model.

Results of FSM model:

**Coefficients<sup>a</sup>**

Model		Unstandardised Coefficients		Standardised Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-4.923	3.727		-1.321	.189
	pre-test score	.087	.010	.570	8.447	.000
	intervention	.935	1.177	.053	.794	.429
	school1	-2.007	3.234	-.053	-.621	.536
	school2	1.184	2.341	.062	.506	.614
	school3	-.815	2.651	-.031	-.307	.759
	school4	6.655	2.311	.359	2.879	.005
	school5	-2.391	2.718	-.086	-.880	.381

a. Dependent Variable: post-test score

Results of prior attainment interaction model:

**Coefficients<sup>a</sup>**

Model		Unstandardised Coefficients		Standardised Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-2.358	2.404		-.981	.327
	pre-test score	.078	.007	.513	10.505	.000
	intervention	-9.476	3.714	-.529	-2.552	.011
	school1	-2.694	1.913	-.066	-1.409	.160
	school2	2.190	1.304	.112	1.679	.094
	school3	-.507	1.602	-.017	-.316	.752
	school4	5.837	1.297	.298	4.500	.000
	school5	-.130	1.425	-.005	-.091	.927
	intervention*pre-test score	.033	.013	.563	2.661	.008

a. Dependent Variable: post-test score

## Appendix B: Contract for Schools

Please sign both copies, retaining one and returning the second copy to \*\*\*.

### Agreement to participate in the Evaluation of Literacy in Transition Project

**School Name:** \*\*\*

#### Aims of the Evaluation

The aim of this project is to evaluate the impact of a rhythm-based intervention on pupils' reading attainment. At the end of the ten-week period, school staff will be trained to deliver the programme so that it can be embedded sustainably in the future work of the school.

The results of the research will contribute to our understanding of what works in raising the pupils' attainment and will be widely disseminated to schools in England. Ultimately we hope that the evaluation will equip school staff with additional knowledge and skills with which to better support children with reading.

#### The Project

The Literacy in Transition Project consists of a main intervention and a control group. Participating pupils will have attained a 'fragile' Level 4 or lower in literacy at the end of Key Stage 2. Taking place in the summer term of 2013, the project supplies a ten-week course of ten-minute weekly sessions delivered to groups of ten pupils by Marion Long or her associates: Sally Cathcart and Lorna Tyack. Group One will receive the rhythm-based reading intervention – an embodied approach to reading, whereas Group Two will remain in their usual school routine. The impact of the intervention will be measured by Marion Long's team who will assess the pupils using the Neale Analysis of Reading Ability prior to and after the intervention period. The NARA provides very rich data as it tests oral reading of connected text. It is administered individually and requires 20 minutes per pupil. Impact will also be measured by an independent evaluator, NFER, who will use the GL Assessment New Group Reading Test (NGRT).

#### Structure of the Evaluation

Baseline tests using NARA and NGRT are planned for the end of March and final tests in early July. The evaluation team requires all the pupils involved in the project to undertake the digital New Group Reading Test (GL Assessment) before and after the ten-week intervention period. Following baseline testing, all eligible pupils will be assigned to the control or intervention group by the evaluation team. One of the evaluation team will observe Marion Long and Sally Cathcart deliver two intervention sessions and also a teacher-training session at the end of July. Following their training, teachers are welcome to deliver the rhythm-based intervention to pupils in the control group. The evaluation team will also interview four teachers who have taken part in the training session.

The evaluation is being conducted by Ben Styles and Rebecca Clarkson from NFER. Pupils who are selected and whose parents agree<sup>10</sup> for them to take part are randomly allocated to either the intervention group or a control group. Random allocation is essential to the evaluation as it is the only way that we can say for sure what the effect of the intervention is on children's attainment. It is important that schools understand and consent to this process.

The pupils in intervention Group One receive rhythm-based exercises (ten minutes per week) for ten weeks. The pupils in the control group receive their usual school experiences.

#### The New Group Reading Test (NGRT) Digital Version

As part of the evaluation it is required that all the pupils selected take the digital version of the New Group Reading Test before the intervention commences and after the intervention has finished. For both of these sessions the following will apply:

The school will be given a 'testing window' and it is essential that ALL pupils sit the test within this time period.

---

<sup>10</sup>Via a passive consent letter.  
Education Endowment Foundation



*Process*

1. Marion Long will collect the contact details of the staff member/s who will be responsible for organising the test delivery in your school. These contact details will include the staff members' names, school email addresses and telephone numbers.
2. These contact details will be shared with NFER, as NFER have purchased the tests for all the pupils to use.
3. NFER will liaise with the test provider to ensure delivery of the tests. The test provider in this case is GL Assessment (GLA). NFER will share the contact details with GLA.
4. GLA will send information to the nominated staff members about how to access the test. This will include login details.
5. The staff member will need to log in and then upload the specified pupil information from your school management system. This will need to be done in advance of the testing session/s. GLA will provide technical support.

*Testing*

On the day of the testing session pupils will log in individually and the test will link to their preloaded data automatically. GLA will provide technical support.

**You will need to ensure that there is adequate ICT provision and support from staff members timetabled in to cover the testing session/s.**

*Results*

As the test is digital and adaptive it is marked almost immediately. Detailed reports are downloaded from GLA and can be exported to your school management system. This can be carried out at your own convenience. NFER will also have access to these results for the purposes of the evaluation. NFER will access these results from the GLA server, so you do not need to send any data. Pupils' test responses and any other pupil data will be treated with the strictest confidence. The responses will be collected online by GL Assessment and accessed by NFER. Named data will be matched with the National Pupil Database and shared with EEF. No individual school or pupil will be identified in any report arising from the research.

**RESPONSIBILITIES****THE LIT PROJECT TEAM WILL:**

- DELIVER 10 INTERVENTION SESSIONS AND TWO TRAINING SESSIONS AND SUPPLY NECESSARY RESOURCES
- BE THE FIRST POINT OF CONTACT FOR ANY QUESTIONS ABOUT THE EVALUATION
- ENSURE ALL STAFF IN THE TEAM ARE TRAINED AND HAVE RECEIVED CRB CLEARANCE
- PROVIDE ON-GOING SUPPORT TO THE SCHOOL
- SEND OUT REGULAR UPDATES ON THE PROGRESS OF THE PROJECT THROUGH A NEWSLETTER
- TEST CHILDREN ON READING ATTAINMENT USING NARA IN MARCH AND JULY 2013
- SHARE OUTCOMES WITH THE SCHOOL
- DISSEMINATE RESEARCH FINDINGS (THE SCHOOL AND PUPILS' IDENTITIES WILL NOT BE PUBLISHED).

**THE NFER EVALUATING TEAM WILL:**

- CONDUCT THE RANDOM ALLOCATION
- COLLECT AND ANALYSE DATA FROM THE PROJECT COLLECTED FROM THE ONLINE ADMINISTRATION OF NGRT, THE TEACHER LOG AND THE PROCESS EVALUATION
- ENSURE ALL STAFF CARRYING OUT OBSERVATIONS AND WORKING WITH PUPIL DATA ARE TRAINED AND HAVE RECEIVED CRB CLEARANCE
- DISSEMINATE RESEARCH FINDINGS (THE SCHOOL AND PUPILS' IDENTITIES WILL NOT BE PUBLISHED).

The School will:

- Consent to random allocation and commit to the outcome (whether treatment or control)
- Allow time for each testing phase and liaise with the evaluation team to find appropriate dates and times for testing to take place
- Assign a member of staff to liaise with GL Assessment and NFER if required
- Ensure the shared understanding and support of all school staff for the project and personnel involved
- Be a point of contact for parents / carers seeking more information on the project.

**We commit to the Evaluation of the Literacy in Transition project as detailed above**

**Signatures**

ON BEHALF OF THE LITERACY IN TRANSITION TEAM

PROJECT LEADERS: MICHELLE LOUGHREY AND MARION LONG

DATE: \_\_\_\_\_

ON BEHALF OF NFER:

LEAD EVALUATOR BEN STYLES:



DATE: 5/2/13

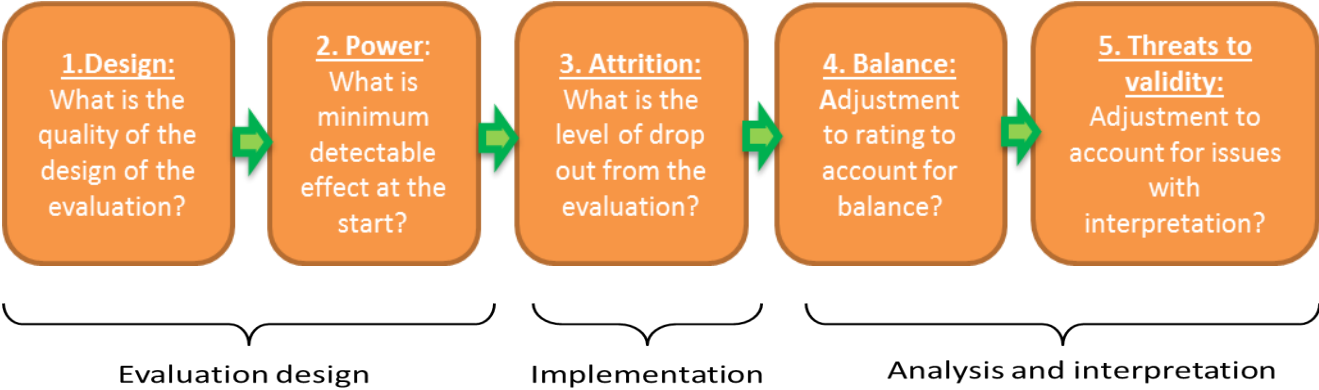
ON BEHALF OF THE SCHOOL:

HEAD TEACHER [NAME]: \_\_\_\_\_

OTHER RELEVANT SCHOOLS STAFF [NAMES]: \_\_\_\_\_

DATE: \_\_\_\_\_

### Appendix C: Security rating summary – Rhythm for Reading



Rating	1. Design	2. Power (MDES)	3. Attrition	4. Balance	5. Threats to validity
🔒🔒🔒🔒🔒	Fair and clear experimental design (RCT)	< 0.2	< 10%	Well-balanced on observables	No threats to validity
🔒🔒🔒🔒🔒	Fair and clear experimental design (RCT, RDD)	< 0.3	< 20%	↓	One threat
🔒🔒🔒🔒🔒	Well-matched comparison (quasi-experiment)	< 0.4	< 30%	0.14 on imbalance on prior attainment but regression used	↓
🔒🔒🔒🔒🔒	Matched comparison (quasi-experiment)	< 0.5	< 40%	↓	↓
🔒🔒🔒🔒🔒	Comparison group with poor or no matching	< 0.6	< 50%	↓	↓
🔒🔒🔒🔒🔒	No comparator	> 0.6	> 50%	Imbalanced on observables	Significant threats

The final security rating for this trial is 3🔒. This means the estimates are reasonably secure.

The trial was designed as a high-powered, individually randomised, independently-run, efficacy trial and could have achieved a maximum of 5🔒. However it was run a short time period in only six schools, which is a concern for the generalisability of the findings.

The trial was well-implemented with only 15% attrition, which is average. There was some imbalance on prior attainment observed at the baseline (an effect size of 0.14) but this was not significant and regression was used to control for it. There was some threat to the validity of the study because the tests were delivered by the schools, rather than blind by the evaluator. However, there is no evidence that they did not deliver these under exam conditions as required.

Additional detail about the EEF’s security rating system can be found at:

[www.educationendowmentfoundation.org.uk/evaluation](http://www.educationendowmentfoundation.org.uk/evaluation).

You may re-use this document/publication (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence v2.0.

To view this licence, visit [www.nationalarchives.gov.uk/doc/open-government-licence/version/2](http://www.nationalarchives.gov.uk/doc/open-government-licence/version/2) or email: [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk)



Education  
Endowment  
Foundation

The Education Endowment Foundation  
9th Floor, Millbank Tower  
21–24 Millbank  
London  
SW1P 4QP  
[www.educationendowmentfoundation.org.uk](http://www.educationendowmentfoundation.org.uk)

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned. The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

This document is available for download at [www.educationendowmentfoundation.org.uk](http://www.educationendowmentfoundation.org.uk)



Education  
Endowment  
Foundation

The Education Endowment Foundation  
9th Floor, Millbank Tower  
21–24 Millbank  
London  
SW1P 4QP  
[www.educationendowmentfoundation.org.uk](http://www.educationendowmentfoundation.org.uk)