Education
Endowment
Foundation

# SPOKES

## Evaluation report and executive summary

July 2016


**Independent evaluators:**

Louise Tracey, Bette Chambers, Tracey Bywater and Louise Elliott


Institute for
**Effective Education**

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.

The EEF aims 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; and
- encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus - Private Equity Foundation) and received a founding £125m grant from the Department for Education.

Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.

**For more information about the EEF or this report please contact:**

**Danielle Mason**
Head of Research and Publications
Education Endowment Foundation
9th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP

**p**: 020 7802 1679
**e**: danielle.mason@eefoundation.org.uk
**w**: www.educationendowmentfoundation.org.uk

# About the evaluator

The project was independently evaluated by a team from the Institute for Effective Education (IEE), University of York: Louise Tracey, Bette Chambers, Tracey Bywater and Louise Elliott, with statistical analysis conducted by Chris Whitaker of Whitaker Research Ltd.

The lead evaluator was Louise Tracey.

Contact details:

**Louise Tracey**
**Institute for Effective Education**
University of York
Heslington, UK
YO10 5DD

Tel: 01904 328160
Email: louise.tracey@york.ac.uk

# Contents

# Executive Summary

## The project

The SPOKES (Supporting Parents on Kids Education in Schools) programme is a ten-week intervention for parents designed to help struggling readers in Year 1. The programme teaches parents strategies to support their children's reading such as listening to children read, pausing to let them work out words, and praising them when they concentrate and problem-solve.

The programme was based primarily in Plymouth and was conducted with six cohorts of children and parents, one cohort each term from Spring term 2012 to Autumn term 2014. This evaluation was designed to assess the impact of the programme on children's reading outcomes. It was a randomised controlled trial involving the parents of 808 children from 68 primary schools. Parents of Year 1 children identified as 'struggling readers' were recruited through their child's school to participate in the project. Parents of the intervention children participated in ten weekly SPOKES sessions over one term and parents of children in a comparison group received books and newsletters. The impact evaluation measured the impact of SPOKES—on children's literacy (letter identification, word identification, and phonetic awareness) and on a range of social and emotional outcomes—at the end of the programme and at six- and twelve-month follow up points. The process evaluation was designed to collect parents' views and experiences of SPOKES, to help to explain the findings from of the impact evaluation, and to provide feedback to inform the future design and delivery of the programme.

The programme was developed as part of a Department of Health grant agreement by Professors Stephen Scott and Kathy Sylva. This project was funded by the Education Endowment Foundation.

| Key Conclusions |
|---|
| 1. There is no evidence that SPOKES had an effect on the standardised reading outcomes or the social and emotional outcomes of the participating children, overall. |
| 2. However, the follow up results for attainment suggest that SPOKES had a positive impact on some aspects of reading for boys over the longer term, equivalent to between three and seven months of additional progress. |
| 3. Parents who participated in the SPOKES programme were very positive about it and believed that their children's literacy was enhanced from their participation. |
| 4. Factors that may have contributed to the lack of overall impact include a low attendance rate and a sample that included children of higher ability than the programme was designed for. |
| 5. In addition to the pre-agreed analysis, the evaluation team also assessed the impact of SPOKES on Key Stage 1 scores for the subgroup of pupils who had reached KS1. This analysis is promising and the positive results warrant future research and further analysis of all the participants at the end KS1. |

## How secure are the findings?

***Security rating awarded as part of the EEF peer review***

These findings have moderate security. The study was a large, well-designed, two-armed randomised controlled trial with randomisation at the pupil level. The trial was classified as an efficacy trial, meaning that it sought to test whether the intervention can work under ideal or developer-led conditions. For the pupils included in the analysis, those whose parents were allocated to receive the intervention were observed to be similar to the pupils in the comparison group. However, 22 % of the pupils were not included in the final analysis because they did not complete all the tests at the beginning and end of the

trial. This is considered to be a relatively high level of attrition and reduces our ability to be sure that the testing is an accurate reflection of all the pupils in the sample.

## What are the findings?

In this efficacy trial there was no evidence, for the intervention group as a whole, that SPOKES had an impact on the outcomes specified in the trial protocol at post-test or at six- and twelve-month follow-up. A positive effect for SPOKES was found on KS1 teacher-based literacy measures; these, however, were not pre-specified as outcome measures.

Analysis of the results for boys only shows that, compared to other boys, those whose parents took part in SPOKES experienced about three months of additional progress on word identification and phonetic awareness at the six-month follow-up, and about seven months of additional progress on phonetic awareness at the twelve-month follow-up. In each case, the positive result would be unlikely to have occurred by chance. The positive results are not conclusive because follow-up results were not available for all boys in the trial, and because positive results were not found for all of the three outcome measures. However, given that SPOKES is intended to have an impact on pupil attainment by changing parental behaviour, some positive results at both follow-up points suggests that SPOKES may have improved aspects of reading for boys over the longer term. For children eligible for free school meals there was no evidence of an impact on literacy or social and emotional outcomes.

The average attendance rate indicated that only 57% of the parents participated in five or more of the ten sessions, which was judged by the development team to be the level of attendance required to be considered a 'completer' of the SPOKES programme. This low attendance rate might have contributed to the fact that there was no evidence of impact on the overall group.

The process evaluation indicated that most of the intervention parents who were interviewed reported liking and benefitting from their participation in the SPOKES programme. At both immediate post-test and subsequent follow-ups they believed that it improved their children's reading. The table below gives the average effect size and estimated months' progress at post-test across the three primary outcome measures compared to the control group.

## How much does it cost?

The cost for the programme implementation in this study was approximately £804 per place, which could be considered per child, per parent, or per family. These costs are predominantly staff costs to implement the weekly SPOKES sessions.

| Outcome | No. of pupils | Effect size (95% confidence interval) | Estimated months' progress | Security rating | Cost rating |
|---|---|---|---|---|---|
| Letter ID | 679 | 0.08 (-0.07, 0.23) | +1 | 🔒🔒🔒 | ££££ |
| Word ID | 659 | 0.05 (-0.10, 0.20) | +1 | 🔒🔒🔒 | ££££ |
| Word Attack | 660 | 0.03 (-0.13, 0.18) | +1[1] | 🔒🔒🔒 | ££££ |

[1] *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.*

# Introduction

## Intervention

SPOKES—Supporting Parents On Kids' Education in Schools—is a series of ten weekly sessions to teach parents strategies for supporting their children's reading. It is designed for the parents of struggling readers in Year 1. Examples of strategies parents are taught include listening to children read, pausing to let them work out words, and praising them when they concentrate and problem-solve. It also encourages discussion of story content and plot.

The SPOKES programme evaluated in the current study is based on an earlier version of the SPOKES programme which was a combination of support for parents in improving their children's social behaviour (drawn from the Incredible Years programme) and helping parents to support their children's learning to read. The SPOKES programme implemented in the current evaluation was different to that previously evaluated: it was shorter than the original SPOKES programme and focused primarily on teaching parents to support their children's oral language and reading. The SPOKES programme evaluated in this EEF study aimed to provide parents with the specific skills they need to support their young children's learning to read words and comprehend story meaning.

The parenting sessions were implemented by the Plymouth Parent Partnership—a social service agency located in Plymouth—and delivered by pairs of group leaders consisting of an educational psychologist and a Parenting Programme Facilitator. The group leaders were trained to implement the programme by the programme developers and attended weekly supervision sessions on filmed groups (see 'Implementation' section in the process evaluation).

## Background evidence

There is strong evidence that parenting skill is one of the strongest predictors of positive outcomes for children, both behavioural and cognitive (Gardner, 2013). However, the evidence about how to improve parental support for children's academic success, and what kind of involvement improves children's attainment, is much less clear.

The early version of SPOKES, cited above, was evaluated in a randomised study of 108 five- and six-year-old children selected from eight primary schools in inner-city London. The selected children were identified as being at risk of social exclusion through antisocial behaviour. The study focused on improving children's behaviour and literacy and was found to be effective in improving the word reading (decoding), writing skills, and behaviour of these emotionally dysregulated children (Scott, *et al.*, 2012; Sylva *et al.*, 2008). The intervention was delivered between 1999 and 2001 before widespread phonics teaching was rolled out. The early version was longer than the version of SPOKES evaluated here and, through the inclusion of dimensions of the Incredible Years parenting programme, focused more on improving the children's social behaviour and literacy rather than focusing primarily on literacy attainment like the updated SPOKES programme implemented for this evaluation.

The programme evaluated in this current study is referred to as SPOKES throughout the report. It focuses on improving children's literacy skills, although children's social behaviour and the parenting behaviours of parents or carers were also measured as secondary outcomes. The literacy strategies of SPOKES are based on the Pause Prompt Praise (PPP) approach initially developed in New Zealand in the 1970s to train parents as tutors (McNaughton, Glynn and Robinson, 1987) and on broader 'whole language' approaches developed by Clay (1991).

The current version of the SPOKES programme has not been widely implemented. However, the previous study, and wider research indicating the strong influence of parenting skill on children's academic achievement (Gardner, 2013), suggested that an efficacy trial of the SPOKES programme—a trial of the programme under ideal and controlled conditions as opposed to a real-world

implementation—was warranted. If proved effective, it would contribute to policy change to help struggling readers.

## Evaluation objectives

The impact evaluation was designed to determine whether the participation of parents in the SPOKES programme had an impact on:

1. children's early literacy skills;
2. children's social behaviours; and
3. parents' parenting behaviours.

A 'light touch' process evaluation was conducted alongside the impact study. This was primarily to determine if there were any discernible factors that might explain the outcome findings of the study, such as the level of participation of the parents and their implementation of the programme. The process evaluation was also designed to provide feedback to the development and implementation teams to help guide refinements of the programme and its delivery in the future. In addition, a random sample of one third of the parents in the intervention group was interviewed by researchers by phone with the aim of understanding parental perceptions of SPOKES.

## Project team

**Louise Tracey, PhD. Lead Evaluator**. Dr Louise Tracey is a Research Fellow in the Institute for Effective Education at the University of York. She conducts evaluations of literacy and mathematics programmes in primary schools using experimental and quasi-experimental methodologies. Dr Tracey recently conducted an evaluation of the Success for All programme, a longitudinal matched study involving 40 primary schools over four years and literacy assessments in Reception classes as part of the Born in Bradford cohort study.

**Bette Chambers, PhD. Co-investigator**. Professor Bette Chambers is Director of the Institute for Effective Education at the University of York. Professor Chambers develops and evaluates effective practices in early childhood and literacy education and promotes the use of evidence-based practices. She has authored or co-authored numerous articles, books, and practical guides for teachers.

**Tracey Bywater, PhD. Co-Investigator**. Dr Tracey Bywater is a Professor in the Institute for Effective Education at the University of York, focusing on enhancing parental input in supporting children's success and long-term positive outcomes. She is also an Honorary Senior Research Fellow in the School of Psychology, Bangor University, Wales. Dr Bywater has led on many randomised controlled trials (including Incredible Years) to improve behavioural, cognitive, and health outcomes in children and their parents or carers.

**Louise Elliott.** Louise Elliott is the Data Manager at the Institute for Effective Education where she manages all database organisation, data entry, cleaning, and descriptive statistical analyses conducted in the research work.

## Ethical review

The evaluation team obtained ethical approval from the Department of Education, University of York Ethics Committee in January 2013. Head teachers signed an agreement outlining the main commitments of the three parties in the study: the school, the project developers, and the evaluators. The evaluation team provided information and consent forms for parents/guardians (see Appendix A).

The assessments were conducted in primary schools by fully trained, Disclosure and Barring Service-checked assessors hired, trained, and supervised by the lead evaluator.

Data were managed in accordance with the Data Protection Act (1998). The trial database is securely held and maintained on the University of York's research data protection server, with non-identifiable data. Confidentiality is maintained and no one outside the trial team has access to the database. Data were checked for missing values and/or double entries. All outputs were anonymised so that no schools or students could be identified in any report or dissemination of results.

## Trial Registration

This trial was registered at http://www.isrctn.com/ISRCTN13234022.

# Methodology

## Trial Design

The trial was a two-armed, within-school, pupil-randomised efficacy trial. It was designed to assess the impact of the programme on reading outcomes of struggling readers in Year 1. The programme was to be run for six cohorts of parents, one cohort each term, starting in the Spring term 2013. Parents of struggling readers were to be identified and recruited through their child's school. The Plymouth Parent Partnership was responsible for recruiting the schools and the schools were expected to take part in recruiting parents for two cohorts of the study. Five cohorts were recruited solely from the Plymouth area and one cohort extended the geographical area to include the surrounding region. This was to ensure that recruitment was not taking place from the same pool of children between cohorts—that struggling readers from Year 1 in each school were only identified once per academic year.

In each of the six cohorts, it was anticipated that approximately eight pupils in Year 1 per school would be designated as struggling readers. Where siblings were identified it was intended that only one would be selected for the outcome measures. Forty-five primary schools were anticipated to take part—each school participating in two cohorts, and identifying eight eligible pupils for each—making a total of 720 pupils. Pupils were to be randomly assigned within schools to the SPOKES programme or to the control group (approximately 360 in each). The parents of those assigned to the intervention condition were offered the ten-week SPOKES programme and those in the control condition received children's books and newsletters. The control group was anticipated to provide an ideal counterfactual for the intervention as being within the same school, control group children would receive the same education beyond the SPOKES programme; furthermore, children within schools tend to share more characteristics than children from different schools. The Institute for Effective Education (IEE) conducted the random assignment.

The children (control and intervention) had their vocabulary assessed in school prior to the programme commencing, and their literacy assessed (letter identification, word reading, and phonetic awareness) at the end of the programme, at six months, and at twelve months for the cohorts that had reached those milestones. Additional child and parent behavioural measures and parents' perceptions were also collected at those time points.

## Eligibility

The 'struggling reader' definition was intended to include children who had been identified as having difficulties in their literacy learning using the previous year's teacher-completed Early Years Foundation Profile 5+ (EYFSP). This measure was chosen as the EYFSP is a statutory requirement at the end of the Reception year and therefore would have been documented for all children. However, as schools often felt that the EYFSP was not a good reflection of the child's current ability due to being completed at the end of Reception year, this was changed so that a 'struggling reader' was defined as a child identified by the teacher as not being where they 'would want or expect the child to be in literacy' based on their knowledge of the child. Therefore, both the EYFSP and the daily experiences of the Year 1 teacher were used to assess eligibility for the programme. This means that, as the schools were responsible for identifying participants, the identification process and criteria may have been subject to differences between schools.

In addition, in order to be eligible for the study, a child's parents had to agree to (and be able to) participate in the SPOKES programme *if their child was subsequently selected to be in the intervention group*. While this may have resulted in some selection bias this would have held true for both the intervention and control conditions.

## Intervention

The SPOKES programme aims to give parents the skills they need to help their children learn to read, particularly children who are struggling readers. It consists of a series of ten group sessions led by two trained group leaders (an educational psychologist and a Parenting Programme Facilitator) following a programme manual to ensure uniformity of delivery.

During the programme, parents learn strategies to support their children's literacy learning, these include Pause Prompt Praise and 'whole language' strategies. In the PPP approach parents are shown how to respond when their child does not recognise a word. Instead of intervening immediately, which can encourage dependence, parents are taught to wait for five seconds to allow the child time to grapple with the problem. If the child did not produce the right answer, parents are encouraged first to offer a prompt before revealing the correct answer. Parents are also taught how to give effective feedback by praising a specific reading behaviour. In the whole-language approach parents are encouraged to discuss books with their child, to relate the story to their everyday experience, and to support their child to anticipate what might happen next.

Supporting these core components, most sessions consist of practical activities (for example, making puppets for re-telling stories, word games, making up novel story endings), group discussion, role play, homework, and demonstrations of reading with children. The programme promotes collaborative learning to help motivate parents to learn new strategies. In addition, a library visit and a one-to-one visit (usually in the home and with both parent and child present) are included in the SPOKES schedule (Sylva *et al*, 2013). Books and games are made available for parents to borrow and share with their family between sessions. Parents are also encouraged to make use of print available in their environment to further support their child's reading skills.

For this study SPOKES was delivered by the Plymouth Parent Partnership ('the Partnership'), an organisation that helps support parents, especially those whose children have learning difficulties. The Parenting Programme Facilitators were provided by the Partnership, and the educational psychologists were provided by Plymouth Community Psychology Service. All SPOKES group leaders were provided with a three-day training programme that focused on techniques for training parents in the skills they need to support reading at home. An additional three-day training course in the Incredible Years parenting programme was also provided prior to this to all group leaders who were not already trained in the Incredible Years. They also attended weekly supervision sessions conducted by the developers and a Partnership disseminator. These sessions were designed to help group leaders enhance their knowledge and skills to deliver the SPOKES programme and maintain fidelity.

Parents of the control group children received children's books to share with their children and newsletters containing information about activities to do with their children, based around the relevant free book given out at the time. The Partnership contact telephone number for the SPOKES programme was also featured in each of the newsletters. These resources were sent home in children's book bags over the ten-week period and were an added incentive to both schools and parents to participate in the study.

## Outcomes

The primary tests for the impact analyses were administered by assessors, hired and trained by evaluators from the Institute for Effective Education. They all had Disclosure and Barring Service (DBS) checks, received training and coaching from the IEE, and were blind to the condition that each child was in. Blinding to condition was easy to maintain as it was the parents who participated in the treatment, not the children. Results of these tests were supplemented by the results of national routinely collected assessments recorded on the National Pupil Database (NPD). Additional measures were completed by the teacher or the parent (facilitated by an IEE assessor).

Measures were as follows:

1) Completed by the child:

a) Woodcock Reading Mastery Tests (revised, WRMT-R, Woodcock, 1998). The co-primary outcome measures were the Letter Identification, Word Identification, and Word Attack sub-tests of the WRMT-R. These were administered at post-intervention; for cohorts one to five at six months post-baseline, and for cohorts one and two at 12 months post-baseline. They are standardised, individually administered tests of reading designed to assess letter recognition, word recognition, and phonemic awareness. These capture reading skills that beginning readers are expected to learn. Reliability coefficients for these sub-tests range from 0.94 and 0.98 for the target age range.

b) British Picture Vocabulary Scale (second edition, BPVS II, Dunn *et al*., 1997). The BPVS was administered at baseline only and formed the pre-test for each participating child. The BPVS II is an individually-administered, norm-referenced test of receptive vocabulary for Standard English. It is a standardised test showing the extent of English vocabulary acquisition. The BPVS was chosen for the pre-test because children's oral language correlates highly with children's IQ and subsequent reading ability. Cronbach's Alpha, which is a measure of internal consistency, is high: for the Year 1 age group it is 0.96.

c) Year 1 phonics screening check. This is a compulsory check administered in schools in the summer term to assess Year 1 pupils' phonics decoding skills. This data was collected from the NPD for all participating pupils (cohorts one to five). Data was not available for cohort six at the time the analysis was done. This is a measure of importance to parents and schools as it is seen as a marker of phonic competence and was chosen to assess children's phonemic awareness at a specific time point in their school career.

2) Completed by the teacher about the child:

The Strengths and Difficulties Questionnaire (SDQ, Goodman, 1997), with impact supplement. Additional detail on the children's behaviour and emotional wellbeing was collected using the SDQ. This was completed by teachers at post-test and subsequent follow-up phases. It is a 25-item behavioural screening questionnaire covering five dimensions of children's behaviour (age 4–16): conduct problems, inattention-hyperactivity, emotional symptoms, peer problems, and pro-social behaviour. The impact supplement asks teachers about any difficulties they feel the child may have. The different dimensions were to be looked at separately to assist in identifying whether the programme had any impacts on children's social behaviour, and for which children and behaviours the SPOKES programme was most effective.

3) Completed with the parent:

The PARYC (McEachern *et al.*, 2012) parenting questionnaire. This is a relatively new measure with three subscales: supporting positive behaviour, proactive positive parenting, and setting limits. This measure was included to assess the parenting skills that the SPOKES programme is designed to affect. The first two subscales were used as they more accurately reflected the underlying principles of the SPOKES programme and were piloted in autumn 2013. The PARYC items are rated via a seven-point Likert scale. The PARYC was normed specifically with the parents of children aged five and six. Cronbach's alpha for Supporting Positive Behaviour is 0.78 and for Proactive Parenting is 0.85.

In the study, this assessment was to be administered in person at baseline by the Partnership home visiting team during pre-programme home visits, with a post-test assessment administered with all parents (one main carer per child) by telephone interview at post-test. The Partnership unfortunately did not have the staff resources to conduct these initial home visits and so the PARYC was collected from parents during a 15-minute telephone interview by the IEE evaluation team at post-test and subsequent follow-up phases.

In addition, although not specified in the protocol, the Key Stage 1 assessment literacy results were available for cohorts one and two at the end of the study. These are national formal assessments, administered and scored by the children's teachers and taken by all children at the end of Year 2. It was decided to include these measures in this report even though they were not specified in the protocol, as they are regarded nationally as a key indicator of children's literacy development. The results of the reading and writing elements of the assessments for these children were obtained from the NPD.

## Sample size

Using the specifications of the sample size outlined in the Trial Design section above (approximately 16 pupils per school from 45 schools across six cohorts), the IEE estimated the statistical power of the planned analyses using Optimal Design software. Assumptions were as follows:

- Pre-post correlation (squared): 0.49
- Intra-class correlation: 0.15
- Criterion for statistical significance: $p < 0.05$

Based on these assumptions it was determined that a minimum detectable effect size (MDES) of 0.20 could be detected with a probability of 0.80.

Due to the small numbers of parents recruited in some schools and concerns relating to retention, the Plymouth Parent Partnership over-recruited to the study, and 68 schools in total took part over the lifetime of the project. This resulted in 808 participants—a number that the team felt would allow for any attrition. As outcomes were linked to the individual child rather than the school (children could be retained in the programme if they moved schools within the Plymouth area) attrition was anticipated to be low.

## Randomisation

Using a random number generator in Excel, the IEE data manager conducted the randomisation of the subjects who had been recruited by Plymouth Parent Partnership. For each cohort of participants, pupils were randomised to intervention or control conditions once parental consent had been received. Where possible, pupils were randomised within school. Because within-school randomisation was not possible where there was only one child per school, in schools with odd numbers of eligible children, the highest possible even number of children was randomly selected and randomised within school and the randomly selected 'extra' child was added to a pool and randomised a second time. Where siblings were identified, only one was randomised to be included in the outcome measures. The Oxford development team intended to conduct more in-depth research with one third of the sample of the children and parents involved in the project. The IEE team randomly assigned an additional third of the parents in the intervention group to be interviewed. Table 1 shows the allocation to condition of participants per cohort and the proportions assigned to participate in the York and Oxford administered measures. Only the York measures are reported in this report.

**Table 1: Allocation of participants for data collection by York or Oxford**

| Sample | | | | | |
|---|---|---|---|---|---|
| York testing: intervention (½ participants) | | | York testing: control (½ participants) | | |
| York interviewing 1/3 of intervention parents | | Oxford testing 1/3 of intervention | Oxford testing 1/3 of control | | |

## Analysis

The evaluation was intended to be a within-school design (among 45 schools) comparing outcomes between intervention and control, with 16 children per school (8 in each of the two cohorts each school was expected to participate in) with poor reading ability allocated randomly to the two conditions. The primary analysis was an 'intention to treat' design, meaning that even if parents dropped out or ceased attending sessions for any reason, associated outcomes would be retained in the main analyses. Analysis of covariance (ANCOVA) was specified as the main statistical analysis in the protocol, with baseline BPVS scores as the covariate. The model took into account the pre-test (BPVS), cohort, age, gender, and Free School Meals status, with the effect sizes calculated using Hedge's g. The effect size was calculated using the Centre for Evaluation and Monitoring, University of Durham effect size calculator[2] using the observed standard deviation for each group. The means were estimated after fitting the model. School was not included in the model because of the lower than anticipated number of pupils per school recruited and retained in the study, and the lack of a relationship between the group a parent attended and the school their child attended.

Because children participated in the study at different times of the year depending on the cohort they were in, standardised scores that take into account children's age were used in the analyses of the primary outcomes (WRMT-R scores). Consequently, age was not included in the model for these measures but was for the social and behavioural secondary outcomes and the Year 1 Phonics assessment and Key Stage 1 reading and reading/writing results.

Further analysis was conducted to explore whether the programme was more or less effective for FSM pupils compared to their non-FSM peers, and for girls or boys. Data on attendance at the SPOKES session for parents was collected by the Partnership, and FSM data was obtained from the NPD to inform this analysis.

Additional analyses were also performed to include parental participation levels (dosage). The development team specified that attendance at five or more sessions was regarded as sufficient to be regarded as a completing the programme. In accordance with this, the intervention group was split into those who attended four or fewer sessions and those who attended five or more. The three-level treatment factor used contrasts to compare the controls with those attending four or fewer sessions, and comparing those attending four or fewer with those who attended five or more.

Both the treatment by pre-test (BPVS) and the treatment by cohort interactions were included in the model. There was no evidence of a treatment by pre-test interaction being significant. The treatment by cohort interaction had occasional significant effects but with no consistent findings for the time point of outcome measure. So for ease of interpretation we decided to exclude both interaction terms.

[2] www.cem.org/effect-size-calculator

## Process evaluation methodology

A 'light touch' process evaluation was conducted alongside the impact study. This was primarily to determine if there were any discernible factors that might explain the outcome findings of the study, such as the level of participation of the parents and their implementation of the programme. The process evaluation was also designed to provide feedback to the development and implementation teams to help guide refinements of the programme and its delivery in the future.

Data for this process evaluation was collected by the Partnership and included attendance data, parent-completed weekly evaluation forms collected at the end of each session, and an overall evaluation form for parents to complete at the end of the intervention.

In addition, a random sample of one third of the parents in the intervention group was interviewed by researchers by phone. The goals of the interviews were to gain the parents' perceptions of:

- how many sessions they missed and what made their attendance difficult;
- what would make it easier or more attractive for them to attend;
- what specific features of the SPOKES programme the parents liked the most;
- what they liked least and their suggestions for changes;
- what SPOKES strategies or activities they implemented with their child and those they did not; and
- what was their overall view of SPOKES.

The interview questions are presented in Appendix B. For those participating in one of the first five cohorts, the interviews were repeated at six months, and for those in the first two cohorts, at 12 months to assess any longitudinal benefits.

**Process Evaluation Analysis**

*Evaluation Forms*

The responses to the Likert-scale weekly and end-of-programme evaluation forms were input into an Access database and then transferred to SPSS for analysis. A researcher read through the open responses to the questions on the end-of-programme evaluation forms which were completed by the parents. She identified common themes, and created a coding frame. Key quotes for each code were noted.

*Telephone Interviews*

Responses to the telephone interview questions were input directly onto the computer using a specially designed Access database. This was then exported to Excel in preparation for a thematic analysis. A researcher read through the responses to the questions, identified common themes, and created a coding frame. Answers for each question were grouped and noted and a short summary of the key themes was created.

# Impact evaluation

## Timeline

SPOKES is a ten-week programme, which means that it lasts for approximately a full school term. The intervention was to occur over three academic years. The programme was run each term, with one cohort per term, for six terms starting in January 2013. Parents in the treatment group received the intervention for one term during the academic year when their child was in Year 1. There were six cohorts (see Table 2).

**Table 2: Intervention timeline**

| Cohort | School Term (and academic year) of intervention |
|---|---|
| Cohort 1 | Spring Term (2012/2013) |
| Cohort 2 | Summer Term (2012/2013) |
| Cohort 3 | Autumn Term (2013/2014) |
| Cohort 4 | Spring Term (2013/2014) |
| Cohort 5 | Summer Term (2013/2014) |
| Cohort 6 | Autumn Term (2014/2015) |

Recruitment took place in the preceding term for each cohort. This meant that there could be a long time lapse between recruitment to the trial and the start of the programme. Pre- and post-testing was undertaken as close to the beginning and end of each run of the programme as possible, although the timing did vary depending on the school term dates for the two academic years (for example, some post-test assessments for the Summer programme took place at the beginning of the following Autumn Term). Post-test assessments were intended to be scheduled within two weeks of completion of the programme. Where this was not possible the mean delay at follow-up was approximately four to five weeks. Six-month and twelve-month follow-ups were scheduled depending on when the post-test was administered. Again, some leeway was necessary to take into account the school term dates although this was the same for both intervention and control groups. This meant that there was a rolling programme of assessments in the schools over the time period January 2013 to March 2015 (see Table 3).

**Table 3: Timeline for data collection by phase and cohort**

| | 2013 | | | 2014 | | | | | | 2015 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Jan | May | Sep | Nov | Jan | March | May | July | Sept | Nov | Jan | March |
| **Pre-test*** | C1 | C2 | C3 | | C4 | C5 | | | C6 | | | |
| **Post-test#** | | C1 | C2 | | C3 | C4 | | | C5 | | C6 | |
| **6-month follow-up#** | | | | C1 | | C2 | | C3 | | C4 | | C5 |
| **12-month follow-up#** | | | | | | | C1 | | C2 | | | |

*Pre-test: BPVS

# Post-test, six-month and twelve-month follow-up: WRMT sub-tests, SDQ, PARYC

C = Cohort

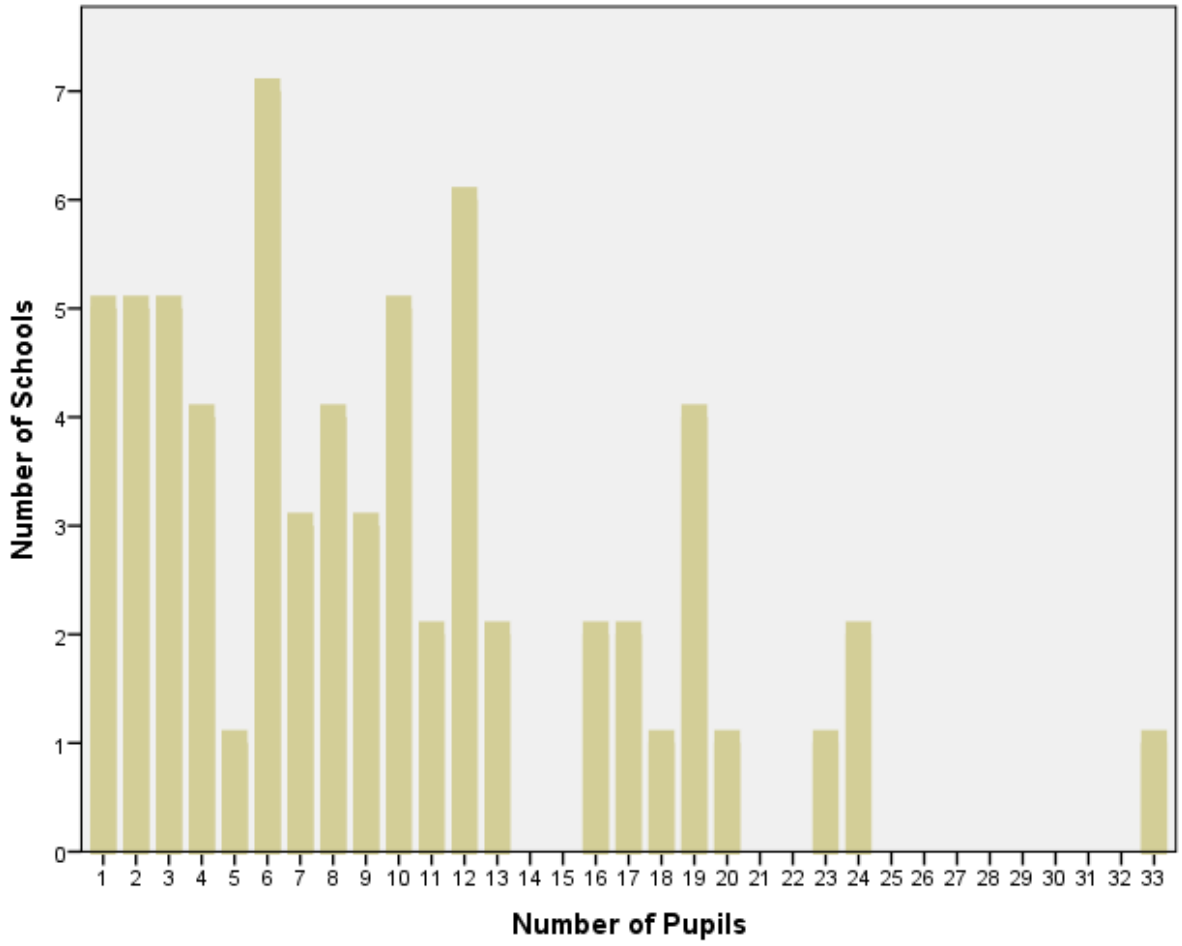
## Participants

Recruitment was conducted through schools. As indicated above, school personnel, primarily Year 1 teachers or the Year 1 literacy lead, were asked to identify struggling readers as defined in the 'eligibility' section above. Those children's parents were then invited to coffee mornings or afternoon tea sessions at the school where more information was made available about the study and they were invited to sign up for the research. Any parents unable to attend were offered home visits where they could be provided with the information. Parents were required to agree to participate in the programme if they were selected, and to take part in the evaluation regardless of randomisation outcome.

To ensure that pupils were not recruited from the same pool across rounds, schools could only take part in one cohort during the academic year 2012/2013 (cohorts one and two), and one cohort in the academic year 2013/2014 (cohorts three, four and five). All schools were eligible for cohort six as this was the only term in the academic year 2014/2015 that the programme was run and therefore the parents of Year 1 pupils would not be approached for recruitment to the programme at any other time in that school year. Plymouth Parent Partnership was responsible for recruitment. Following completion of parental consent, child and parent information were passed to the evaluation team's data manager for random assignment to intervention or control groups.

It was originally estimated that the parents of 16 pupils would be recruited per school across 45 schools, providing a total of 720 parent/pupil participants. Because there were fewer children per school than anticipated, more schools were recruited to participate (68 rather than 45) and more children (808 rather than 720). There were between 1 and 21 children recruited per school within each cohort. As indicated in Figure 1, in the final (analysed) sample (n = 627, 66 schools) there were between 1 and 33 pupils per school across the six cohorts.

**Figure 1: Number of pupils per school in the final (analysed, post-test) sample**




Tables 4–7 provide the characteristics of the participating schools. As can be seen, 4% (3 schools) were in areas with the highest band of deprivation as determined by the Income Deprivation Affecting Children Index (IDACI), and the same number were rated 'inadequate' by Ofsted. Sixty per cent of schools were in the highest two quintiles for proportion of children in receipt of Free School Meals (having18% of pupils or more in receipt of FSM). The Key Stage 2 reading achievement outcomes were fairly evenly spread across the reading attainment quintiles. Four schools were infant schools with the remaining being primary schools, and the majority were in urban locations, with only two being classified as rural.

**Table 4: IDACI rankings of recruitment schools**

| | Lowest band | | | | | | Highest band |
|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| Number of schools (%) | 32 (47%) | 8 (12%) | 8 (12%) | 6 (9%) | 9 (13%) | 2 (3%) | 3 (4%) |

**Table 5: Ofsted ratings of recruitment schools**

| | Outstanding | Good | Requires Improvement | Inadequate |
|---|---|---|---|---|
| Number of schools (%) | 10 (15%) | 46 (68%) | 9 (13%) | 3 (4%) |

**Table 6: FSM quintiles of recruitment schools**

| | Lowest quintile | | | | Highest quintile |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| Number of schools (%) | 6 (9%) | 11 (16%) | 10 (15%) | 21 (31%) | 20 (29%) |

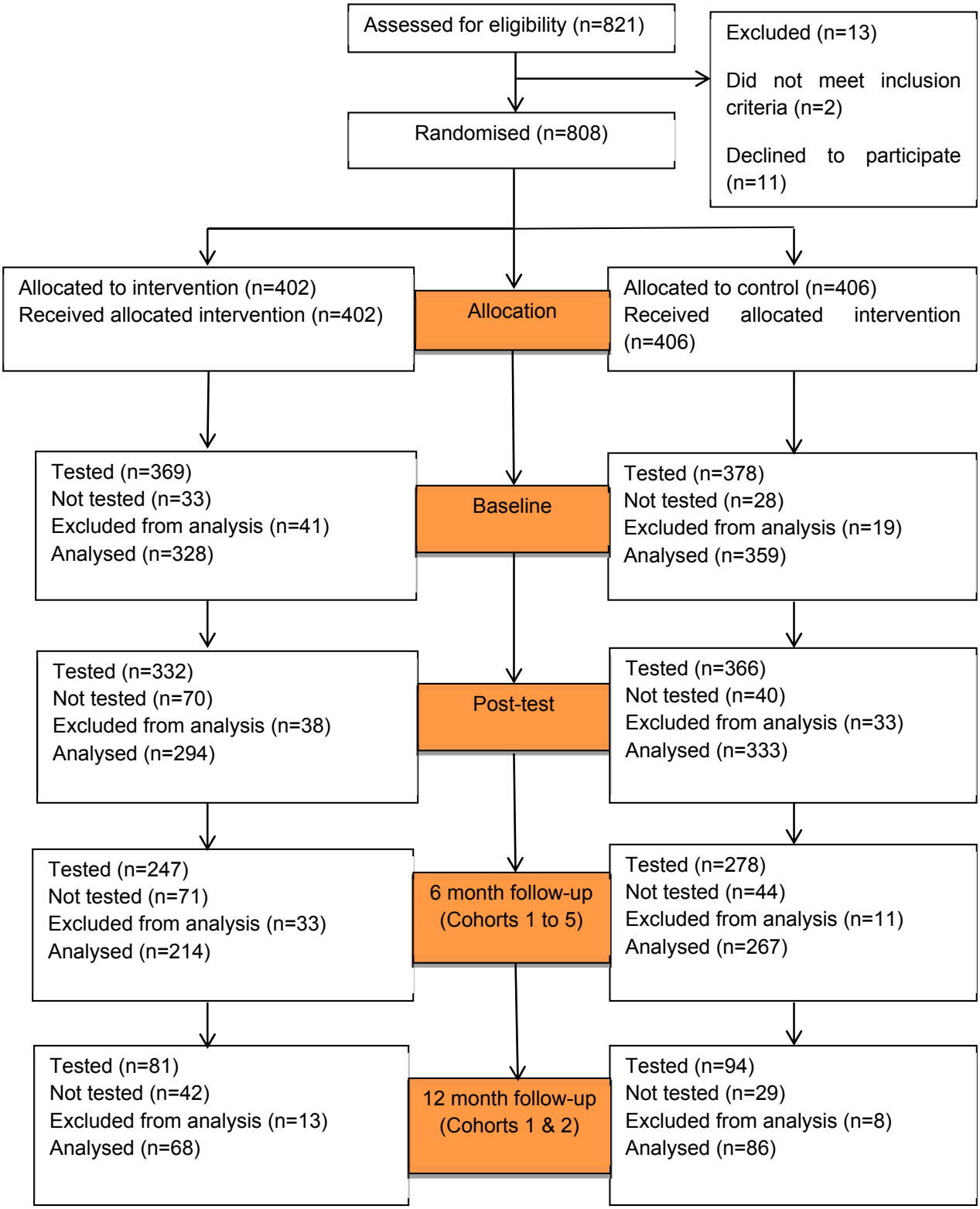
**Table 7: Key Stage 2 reading attainment quintiles of recruitment schools**

| | Lowest quintile | | | | Highest quintile |
|---|---|---|---|---|---|
| | 5 | 4 | 3 | 2 | 1 |
| Number of schools (%)* | 15 (24%) | 10 (16%) | 15 (24%) | 10 (16%) | 12 (19%) |

** 6 of the 68 schools did not have data available.*

Of the 808 participating families, 734 were included in the analyses. This represents less than 10% attrition between pre- and post-test. Figure 2 provides the overall participant flow chart (consort diagrams for each individual cohort are provided in Appendix C). The numbers in the participant flow chart are based on the number of children who completed the WRMT-R Letter Identification subtests at each follow-up. On occasion, a child might have been absent for one of the test administrations but been present for other data collection, hence the numbers do not add up to the total. If children were not available for the administration of the baseline assessment they were excluded from further data collection. Cases were excluded if children did not complete any post-test or follow-up assessment. Children who moved school were followed within area.

**Figure 2: Consort Diagram—whole sample**



## Pupil characteristics

Table 8 summarises the pupil characteristics of the sample, comparing the recruited sample pre- and post-test profiles in terms of gender, Free School Meals status, and pre-test attainment. There was little difference between the standardised pre-test scores in the two groups. In addition, there was little difference between the standardised pre-test scores for those who also gave post-test scores. The control group had a higher proportion of boys than girls, whereas the intervention group was fairly evenly split. The proportion of pupils ever eligible for FSM (having ever claimed FSM) was similar in the two groups. The profile for pupils completing post-tests was very similar to the pre-test, with the exception of those in the post-test group where those with lower baseline attainment were lost from the control condition due to withdrawal from the study. In contrast, for the intervention group pupils with higher baseline attainment were lost to follow up.

**Table 8: Pupil characteristics of the sample**

| | All pupils cohort 1 to 6 (as randomised) | | | All pupils with post-test scores* (as analysed) | | |
|---|---|---|---|---|---|---|
| | Intervention | Control | Difference | Intervention | Control | Difference |
| **Pre-test standardised score** | 100.92** | 99.97** | 0.95 | 100.60# | 100.28## | 0.32 |
| **Gender** | | | | | | |
| Male | 50.0% | 60.1% | -10.1% | 50.3% | 62.8% | -12.5% |
| Female | 49.5% | 39.7% | 9.8% | 49.7% | 37.2% | 12.5% |
| Unknown | 0.5% | 0.2% | -0.3% | - | - | - |
| **FSM (Ever FSM)** | | | | | | |
| No | 65.4% | 68.5% | -3.1% | 71.8% | 71.5% | 0.3% |
| Yes | 24.9% | 26.6% | -1.7 | 27.2% | 27.3% | -0.1% |
| Unknown | 9.7% | 4.9% | 4.8% | 1.0% | 1.2% | -0.2% |
| **Number of pupils (all)** | 402 | 402 | - | 294 | 333 | - |

** All pupils with a valid pre-test and a valid post-test (Letter Identification) standardised score.*

*** All pupils with valid pre-test standardised score.*

*# ES = 0.031 (for the intervention group at baseline compared to the intervention group included in the analysis, i.e. those for whom a post-test score was established)*

*## ES = 0.002 (for the control group at baseline compared to the control group included in the analysis, i.e. those for whom a post-test score was established).*

## Outcomes and analysis

As seen in Figure 1 (above) the number of analysed pupils involved in each school varied between 1 and 33 with more than half of schools having 8 participating pupils or fewer. Only 14 out of 66 schools had 16 or more pupils included in the final analysis. In addition, although participants were recruited through schools, parents did not attend a programme according to the school their child attended. As a result of the lower number of pupils per school recruited, and the lack of a relationship between the group a parent attended and the school their child attended, we did not include school as a factor in the model. We initially checked whether there were any interactions between treatment and BPVS (the pre-treatment assessment of ability) or with cohort. Having found no evidence that interactions were present, the final model included BPVS, gender, FSM, cohort, and treatment with age included for the non-age standardised measures—SDQ, PARYC, Year 1 Phonics and Key Stage 1 reading outcomes.

Given that school was not included in the model, and that the total number of pupils recruited was higher than originally estimated, a post-hoc analysis was completed. Optimal Design software was used to estimate the statistical power of the analyses using the completed sample size for a single-level trial. The calculated pre-post correlation showed a lower than expected correlation between the pre-test and the primary outcome ($r^2$= 0.9). Using this revised figure, and 0.05 as the criterion for statistical significance, it was determined that a minimum detectable effect size (MDES) of 0.216 could be detected with a power of 0.80.

As specified in the protocol, an intention-to-treat analysis was conducted on all eligible pupils, alongside sub-analyses of whether or not the programme was more or less effective when considering only (1) FSM pupils (children categorised as 'ever FSM' in the NPD), (2) girls, and (3) boys. The protocol also specified analysis according to parental level of participation in the programme. Low levels of attendance meant that, in discussion with the developers, the two categories of 'completers' and 'non-completers' were created, the former being deemed to have completed the programme if they attended five or more of the ten sessions. Consequently, outcomes for the 'completers' were compared with those for the control condition in this sub-analysis.

Given that this intervention is aimed at struggling readers, an additional post-hoc sub-analysis was conducted for those participants with lower than average attainment on the pre-test. Only children whose BPVS score was more than 0.5 standard deviations below 100 (that is, who had received a standardised score of 95% or less) were included in this analysis.

There was evidence that those children lost to follow-up were different from those retained (with generally lower baseline scores). Due to the imprecise nature of extrapolating their data they were not included in the analysis at subsequent time points.

As seen above, for the study as a whole there were no statistically significant differences at the conventional level at pre-test between treatment and control. In addition, assumptions of normality and equal variability held for the primary literacy outcomes and Key Stage 1 reading and reading/writing assessment outcomes. For Year 1 Phonics outcomes the residuals (not reported here) tended to be negatively skewed. A square transformation was used and results in the assumptions holding although it made little difference to the conclusions drawn. Un-standardised means are given in the tables.

The results of the overall analysis, and sub-analyses, are reported below and in Appendices D–I. Where an effect size was found to be statistically significant at the conventional level this is indicated in the table by an asterisk (*).

**Intention-to-treat analysis**

*Outcomes at post-test*

As seen in Table 9, there were no statistically significant differences between the results for the control and intervention conditions at immediate post-test for the primary (literacy) measures for the whole

sample (all eligible pupils) participating in the evaluation. While the effect sizes were positive in favour of the SPOKES programme for the primary measures they were small (representing approximately one months' progress) and were likely to have occurred by chance rather than as a result of participating in the intervention.

**Table 9: Primary outcomes at post-test (all eligible pupils)**

| Outcome | Standardised scores | | | | | | Effect size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| Letter ID | 323 (32) | 107.95 (107.02, 108.87) | 8.32 | 356 (27) | 107.24 (106.38, 108.09) | 8.07 | 620 (291, 329) | 0.08 (-0.07, 0.23) |
| Word ID | 315 (29) | 121.42 (119.90, 122.93) | 13.44 | 344 (26) | 119.93 (118.54, 121.32) | 12.89 | 604 (286, 318) | 0.05 (-0.10, 0.20) |
| Word Attack | 313 (30) | 118.98 (117.99, 119.97) | 8.77 | 347 (26) | 118.36 (117.41, 119.32) | 8.88 | 604 (283, 321) | 0.03 (-0.13, 0.18) |

In addition, there were no statistically significant differences between conditions for the secondary (social and behavioural) measures at immediate post-test (Table 10). The results were also inconsistent with both positive and negative effect sizes occurring ranging from -0.08 to +0.18.

**Table 10: Secondary Outcomes – social and behavioural measures at post-test (all eligible pupils)**

| Outcome | Raw scores | | | | | | Effect size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **SDQ** | | | | | | | | |
| Prosocial Score | 271 (27) | 7.20 (6.89, 7.50) | 2.49 | 301 (19) | 6.93 (6.64, 7.22) | 2.53 | 526 (244, 282) | 0.08 (-0.09, 0.24) |
| Total Difficulties Score | 271 (27) | 7.36 (6.68, 8.04) | 5.61 | 302 (19) | 7.70 (7.03, 8.38) | 5.85 | 527 (244, 283) | -0.02 (-0.19, 0.14) |
| Externalising Score | 271 (27) | 4.49 (4.02, 4.97) | 3.94 | 302 (19) | 4.65 (4.17, 5.12) | 4.11 | 527 (244, 283) | -0.01 (-0.17, 0.16) |
| Internalising Score | 271 (27) | 2.86 (2.51, 3.22) | 2.94 | 302 (19) | 3.06 (2.69, 3.42) | 3.15 | 527 (244, 283) | -0.04 (-0.20, 0.13) |
| Impact Score | 120 (17) | 1.13 (0.88, 1.38) | 1.37 | 134 (9) | 1.10 (0.86, 1.33) | 1.35 | 228 (103, 125) | 0.18 (-0.06, 0.43) |
| **PARYC** | | | | | | | | |
| Supporting Frequency | 170 (15) | 41.08 (40.28, 41.88) | 5.21 | 220 (18) | 41.10 (40.28, 41.93) | 6.09 | 357 (155, 202) | -0.02 (-0.22, 0.18) |
| Supporting Problematic | 170 (15) | 1.11 (0.92, 1.29) | 1.22 | 220 (18) | 1.14 (0.97, 1.31) | 1.25 | 357 (155, 202) | -0.01 (-0.21, 0.19) |
| Proactive Frequency | 169 (15) | 38.38 (37.32, 39.44) | 6.90 | 220 (18) | 39.48 (38.54, 40.41) | 6.93 | 356 (154, 202) | -0.08 (-0.28, 0.12) |
| Proactive Problematic | 169 (15) | 1.16 (0.92, 1.40) | 1.53 | 220 (18) | 1.10 (0.91, 1.29) | 1.44 | 356 (154, 202) | 0.03 (-0.17, 0.23) |
| Total Frequency | 167 (15) | 79.46 (77.87, 81.04) | 10.26 | 217 (18) | 80.68 (79.12, 82.23) | 11.45 | 351 (152, 199) | -0.07 (-0.27, 0.13) |
| Total Problematic | 167 (15) | 2.29 (1.94, 2.65) | 2.30 | 217 (18) | 2.26 (1.94, 2.57) | 2.31 | 351 (152, 199) | 0.01 (-0.19, 0.22) |

*Outcomes at six-month and twelve-month follow-up*

Tables 11 and 12 show that at the six- and twelve-month follow-up there were no statistically significant differences between control and intervention conditions for the WRMT-R sub-tests (Letter Identification, Word Identification, and Word Attack). The effect sizes do increase between the immediate post-test and the six-month follow-up and the twelve-month follow-up although these remain non statistically significant and may be a related to the smaller numbers analysed at successive time points.

**Table 11: Primary outcomes at six-month follow-up (all eligible pupils)**

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| Letter ID | 241 (22) | 104.17 (103.32, 105.03) | 6.67 | 271 (21) | 103.07 (102.11, 104.03) | 7.90 | 469 (219, 250) | 0.15 (-0.03, 0.32) |
| Word ID | 219 (22) | 120.43 (118.62, 122.24) | 13.39 | 248 (22) | 118.59 (116.81, 120.38) | 14.05 | 423 (197, 226) | 0.11 (-0.07, 0.29) |
| Word Attack | 239 (21) | 118.94 (117.73, 120.14) | 9.30 | 270 (22) | 117.49 (116.29, 118.70) | 9.90 | 466 (218, 248) | 0.11 (-0.06, 0.28) |

**Table 12: Primary outcomes at twelve-month follow-up (all eligible pupils)**

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| Letter ID | 79 (15) | 97.48 (95.81, 99.15) | 7.42 | 91 (12) | 95.96 (94.09, 97.83) | 8.92 | 143 (64, 79) | 0.13 (-0.17, 0.43) |
| Word ID | 62 (14) | 111.45 (107.80, 115.11) | 14.39 | 69 (11) | 108.45 (104.71, 112.19) | 15.54 | 106 (48, 58) | 0.27 (-0.07, 0.62) |
| Word Attack | 78 (15) | 114.64 (111.90, 117.38) | 12.09 | 89 (12) | 111.42 (108.25, 114.58) | 14.93 | 140 (63, 77) | 0.25 (-0.05, 0.56) |

Statistically significant differences at the conventional level and a positive effect in favour of the intervention condition were found for two sub-tests of the Strengths and Difficulties Questionnaire (SDQ) at the six-month follow-up (Total Difficulties score: Effect Size = 0.217; Internal Difficulties score: Effect Size = 0.21; $p \leq 0.05$). Results for the other sub-tests of the SDQ and the PARYC were not statistically significant. The effects do not appear at the twelve-month follow-up point (Tables 13 and 14).

**Table 13: Secondary outcomes—social and behavioural measures at six-month follow-up (all eligible pupils)**

| Outcome | Raw scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **SDQ** | | | | | | | | |
| Prosocial Score | 187 (22) | 7.35 (6.95, 7.74) | 2.69 | 214 (20) | 7.32 (6.98, 7.66) | 2.49 | 359 (165, 194) | 0.05 (-0.15, 0.25) |
| Total Difficulties Score | 187 (22) | 8.57 (7.55, 9.59) | 6.97 | 214 (20) | 7.48 (6.66, 8.30) | 6.02 | 359 (165, 194) | 0.22 (0.02, 0.41)* |
| Externalising Score | 187 (22) | 4.98 (4.31, 5.66) | 4.60 | 214 (20) | 4.38 (3.83, 4.93) | 4.03 | 359 (165, 194) | 0.16 (-0.03, 0.36) |
| Internalising Score | 187 (22) | 3.58 (3.07, 4.09) | 3.48 | 214 (20) | 3.10 (2.66, 3.54) | 3.23 | 359 (165, 194) | 0.21 (0.02, 0.41)* |
| Impact Score | 93 (12) | 1.31 (0.99, 1.64) | 1.57 | 102 (10) | 1.30 (1.03, 1.58) | 1.38 | 173 (81, 92) | 0.13 (-0.15, 0.41) |
| **PARYC** | | | | | | | | |
| Supporting Frequency | 115 (14) | 42.23 (41.20, 43.27) | 5.52 | 138 (13) | 41.62 (40.71, 42.52) | 5.29 | 226 (101, 125) | 0.09 (-0.16, 0.34) |
| Supporting Problematic | 115 (14) | 0.90 (0.66, 1.13) | 1.27 | 138 (13) | 1.05 (0.83, 1.27) | 1.28 | 226 (101, 125) | -0.13(-0.38, 0.12) |
| Proactive Frequency | 112 (13) | 39.76 (38.61, 40.91) | 6.08 | 138 (13) | 40.33 (39.31, 41.35) | 5.99 | 224 (99, 125) | -0.11 (-0.36, 0.14) |
| Proactive Problematic | 112 (13) | 0.92 (0.66, 1.18) | 1.36 | 138 (13) | 0.99 (0.72, 1.25) | 1.57 | 224 (99, 125) | -0.07 (-0.31, 0.19) |
| Total Frequency | 110 (13) | 81.86 (79.95, 83.78) | 10.04 | 136 (13) | 81.98 (80.28, 83.67) | 9.88 | 220 (97, 123) | -0.04 (-0.29, 0.21) |
| Total Problematic | 110 (13) | 1.80 (1.38, 2.22) | 2.20 | 136 (13) | 2.04 (1.62, 2.47) | 2.50 | 220 (97, 123) | -0.10 (-0.36, 0.15) |

** $p \leq 0.05$*

**Table 14: Secondary outcomes—social and behavioural measures at twelve-month follow-up (all eligible pupils)**

| Outcome | Raw scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **SDQ** | | | | | | | | |
| Prosocial Score | 71 (14) | 7.44 (6.76, 8.11) | 2.83 | 73 (7) | 7.16 (6.61, 7.72) | 2.38 | 123 (57, 66) | 0.06 (-0.27, 0.38) |
| Total Difficulties Score | 71 (14) | 8.54 (6.79, 10.28) | 7.35 | 73 (7) | 8.33 (6.94, 9.72) | 5.94 | 123 (57, 66) | 0.14 (-0.19, 0.46) |
| Externalising Score | 71 (14) | 5.10 (3.92, 6.28) | 4.98 | 73 (7) | 5.08 (4.10, 6.06) | 4.18 | 123 (57, 66) | 0.11 (-0.22, 0.44) |
| Internalising Score | 71 (14) | 3.44 (2.59, 4.28) | 3.55 | 73 (7) | 3.25 (2.54, 3.95) | 3.00 | 123 (57, 66) | 0.12 (-0.21, 0.44) |
| Impact Score | 42 (8) | 1.62 (1.02, 2.22) | 1.94 | 34 (4) | 1.32 (0.77, 1.88) | 1.63 | 64 (34, 30) | 0.27 (-0.18, 0.73) |
| PARYC | | | | | | | | |
| Supporting Frequency | 38 (7) | 42.95 (41.07, 44.82) | 5.78 | 33 (5) | 41.03 (38.83, 43.23) | 6.33 | 59 (31, 28) | 0.24 (-0.23, 0.71) |
| Supporting Problematic | 38 (7) | 0.89 (0.46, 1.33) | 1.33 | 33 (5) | 1.33 (0.78, 1.89) | 1.59 | 59 (31, 28) | -0.26 (-0.73, 0.21) |
| Proactive Frequency | 38 (7) | 41.16 (38.87, 43.45) | 7.06 | 33 (5) | 39.52 (37.45, 41.58) | 5.92 | 59 (31, 28) | 0.29 (-0.18, 0.75) |
| Proactive Problematic | 38 (7) | 0.71 (0.16, 1.26) | 1.69 | 33 (5) | 1.09 (0.55, 1.63) | 1.55 | 59 (31, 28) | -0.26 (-0.73, 0.21) |
| Total Frequency | 38 (7) | 84.11 (80.39, 87.82) | 11.44 | 33 (5) | 80.55 (76.65, 84.44) | 11.19 | 59 (31, 28) | 0.29 (-0.18, 0.76) |
| Total Problematic | 38 (7) | 1.61 (0.69, 2.52) | 2.82 | 33 (5) | 2.42 (1.42, 3.42) | 2.87 | 59 (31, 28) | -0.28 (-0.75, 0.19) |

### Year 1 Phonics results

As seen in Table 15, there were no statistically significant differences at the conventional level seen on the results for the Year 1 phonics assessment for cohorts one to five.

**Table 15: Year 1 Phonics results (all eligible pupils)**

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| Year 1 Phonics | 285 (33) | 30.02 (28.83, 31.20) | 9.99 | 294 (22) | 29.93 (28.72, 31.14) | 10.37 | 524 (252, 272) | 0.02 (-0.15, 0.18) |

### Secondary analysis: Key Stage 1 assessment literacy results

Although not specified in the protocol, statistically significant effect sizes were seen for the KS1 reading scores (p = 0.01) and the reading and writing scores combined (p = 0.049) for those pupils for whom these results were available (cohorts one and two), representing three to four months progress in favour of the intervention condition (Table 16).

Because Key Stage results are an outcome of such interest to schools, we felt it was valid to include these results here as secondary analysis. It is important to note that the analysis is for two out of the six cohorts only (those for which KS1 data was available) and so does not give the full picture of the impact of the intervention on KS1 results.

**Table 16: Key Stage 1 assessment results (all eligible pupils)**

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| KS1 Assessment Reading/Writing Score | 114 (22) | 13.93 (13.24, 14.62) | 3.69 | 111 (14) | 13.08 (12.41, 13.75) | 3.54 | 189 (92, 97) | 0.26 (0.00, 0.52)* |
| KS1 Assessment Reading Score | 114 (22) | 14.67 (13.92, 15.41) | 3.9 | 111 (14) | 13.58 (12.87, 14.28) | 3.72 | 189 (92, 97) | 0.33 (0.07, 0.59)* |

** p ≤ 0.05*

**Sub-analysis**

**Pupils in receipt of Free School Meals**

As seen in Tables 17–24 there were no statistically significant differences found between control and intervention conditions for children in receipt of Free School Meals for the primary or the secondary outcomes at immediate post-test, six-month follow-up or at twelve-month follow-up.

*Outcomes at post-test*

**Table 17: Primary outcomes at post-test (all eligible FSM pupils)**

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| Letter ID | 88 (8) | 105.72 (103.72, 107.71) | 9.35 | 99 (8) | 106.38 (104.56, 108.21) | 9.10 | 171 (80, 91) | 0.02 (-0.27, 0.31) |
| Word ID | 86 (7) | 116.76 (113.64, 119.87) | 14.43 | 96 (7) | 117.46 (114.55, 120.37) | 14.25 | 168 (79, 89) | -0.00 (-0.29, 0.29) |
| Word Attack | 84 (7) | 117.14 (115.16, 119.12) | 9.07 | 95 (8) | 116.87 (114.89, 118.86) | 9.68 | 164 (77, 87) | 0.09 (-0.21, 0.38) |

**Table 18: Secondary Outcomes – social and behavioural measures at post-test (all eligible FSM pupils)**

| Outcome | Raw scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **SDQ** | | | | | | | | |
| Prosocial Score | 79 (9) | 6.63 (6.06, 7.21) | 2.54 | 83 (6) | 7.01 (6.48, 7.55) | 2.44 | 147 (70, 77) | -0.15 (-0.46, 0.16) |
| Total Difficulties Score | 79 (9) | 9.46 (8.09, 10.83) | 6.09 | 83 (6) | 9.52 (8.40, 10.63) | 5.07 | 147 (70, 77) | 0.00 (-0.30, 0.31) |
| Externalising Score | 79 (9) | 5.97 (4.96, 6.99) | 4.51 | 83 (6) | 5.64 (4.73, 6.54) | 4.12 | 147 (70, 77) | 0.08 (-0.23, 0.39) |
| Internalising Score | 79 (9) | 3.48 (2.80, 4.16) | 3.03 | 83 (6) | 3.88 (3.20, 4.56) | 3.11 | 147 (70, 77) | -0.11 (-0.41, 0.20) |
| Impact Score | 51 (6) | 1.12 (0.73, 1.51) | 1.39 | 48 (2) | 1.02 (0.70, 1.34) | 1.12 | 91 (45, 46) | 0.22 (-0.17, 0.62) |
| **PARYC** | | | | | | | | |
| Supporting Frequency | 33 (3) | 40.70 (38.80, 42.60) | 5.46 | 53 (5) | 40.60 (38.65, 42.56) | 7.12 | 78 (30, 48) | -0.08 (-0.51, 0.36) |
| Supporting Problematic | 33 (3) | 1.27 (0.81, 1.74) | 1.33 | 53 (5) | 1.23 (0.88, 1.57) | 1.27 | 78 (30, 48) | 0.01 (-0.43, 0.44) |
| Proactive Frequency | 34 (3) | 38.03 (35.67, 40.39) | 6.87 | 55 (5) | 38.53 (36.35, 40.71) | 8.09 | 81 (31, 50) | 0.15 (-0.28, 0.58) |
| Proactive Problematic | 34 (3) | 1.74 (1.19, 2.28) | 1.58 | 55 (5) | 1.18 (0.78, 1.58) | 1.48 | 81 (31, 50) | 0.16 (-0.27, 0.59) |
| Total Frequency | 33 (3) | 78.70 (74.97, 82.43) | 10.71 | 53 (5) | 79.17 (75.47, 82.87) | 13.47 | 78 (30, 48) | 0.05 (-0.39, 0.49) |
| Total Problematic | 33 (3) | 3.06 (2.34, 3.78) | 2.08 | 53 (5) | 2.43 (1.81, 3.06) | 2.28 | 78 (30, 48) | 0.07 (-0.37, 0.50) |

**Outcomes at six-month and twelve-month follow-up**

**Table 19: Primary outcomes at six months (all eligible FSM pupils)**

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| Letter ID | 66 (6) | 103.55 (101.83, 105.26) | 6.98 | 75 (6) | 101.35 (99.29, 103.41) | 8.93 | 129 (60, 69) | 0.30 (-0.04, 0.63) |
| Word ID | 63 (6) | 117.81 (113.96, 121.65) | 15.26 | 68 (6) | 116.03 (112.46, 119.60) | 14.74 | 119 (57, 62) | 0.11 (-0.24, 0.45) |
| Word Attack | 65 (6) | 118.31 (115.71, 120.91) | 10.49 | 75 (6) | 115.87 (113.37, 118.37) | 10.82 | 128 (59, 69) | 0.15 (-0.18, 0.48) |

**Table 20: Secondary outcomes—social and behavioural measures at six-month follow-up (all eligible FSM pupils)**

| Outcome | Raw scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **SDQ** | | | | | | | | |
| Prosocial Score | 57 (8) | 6.58 (5.87, 7.29) | 2.67 | 59 (6) | 7.39 (6.73, 8.05) | 2.53 | 102 (49, 53) | -0.15 (-0.51, 0.22) |
| Total Difficulties Score | 57 (8) | 10.67 (8.74, 12.59) | 7.27 | 59 (6) | 9.19 (7.75, 10.62) | 5.50 | 102 (49, 53) | 0.19 (-0.18, 0.55) |
| Externalising Score | 57 (8) | 6.89 (5.54, 8.25) | 5.11 | 59 (6) | 5.32 (4.30, 6.34) | 3.91 | 102 (49, 53) | 0.26 (-0.10, 0.63) |
| Internalising Score | 57 (8) | 3.77 (2.88, 4.66) | 3.35 | 59 (6) | 3.86 (3.00, 4.73) | 3.32 | 102 (49, 53) | 0.01 (-0.36, 0.37) |
| Impact Score | 38 (6) | 1.21 (0.69, 1.73) | 1.60 | 40 (3) | 1.18 (0.78, 1.57) | 1.24 | 69 (32, 37) | 0.20 (-0.25, 0.64) |
| **PARYC** | | | | | | | | |
| Supporting Frequency | 26 (3) | 44.04 (42.13, 45.95) | 4.86 | 32 (3) | 40.34 (38.33, 42.36) | 5.70 | 52 (23, 29) | 0.67 (0.14, 1.20)* |
| Supporting Problematic | 26 (3) | 0.92 (0.47, 1.38) | 1.16 | 32 (3) | 1.31 (0.84, 1.78) | 1.33 | 52 (23, 29) | -0.09 (-0.60, 0.43) |
| Proactive Frequency | 26 (3) | 39.38 (36.66, 42.11) | 6.94 | 33 (3) | 37.91 (35.83, 39.98) | 5.96 | 53 (23, 30) | -0.03 (-0.54, 0.49) |
| Proactive Problematic | 26 (3) | 1.58 (0.91, 2.24) | 1.70 | 33 (3) | 1.12 (0.53, 1.71) | 1.69 | 53 (23, 30) | 0.30 (-0.21, 0.82) |
| Total Frequency | 26 (3) | 83.42 (79.68, 87.16) | 9.54 | 32 (3) | 78.13 (74.59, 81.66) | 9.99 | 52 (23, 29) | 0.34 (-0.18, 0.86) |
| Total Problematic | 26 (3) | 2.50 (1.55, 3.45) | 2.42 | 32 (3) | 2.44 (1.51, 3.37) | 2.63 | 52 (23, 29) | 0.17 (-0.35, 0.69) |

** $p \leq 0.05$*

**Table 21: Primary outcomes at twelve-month follow-up (all eligible FSM pupils)**

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| Letter ID | 32 (4) | 97.09 (94.54, 99.65) | 7.22 | 31 (5) | 96.84 (93.74, 99.94) | 8.63 | 54 (28, 26) | 0.10 (-0.39, 0.60) |
| Word ID | 21 (3) | 107.48 (101.14, 113.81) | 14.51 | 24 (4) | 105.50 (99.99, 111.01) | 13.51 | 38 (18, 20) | 0.21 (-0.38, 0.80) |
| Word Attack | 32 (4) | 113.91 (109.70, 118.11) | 11.90 | 30 (5) | 110.20 (105.02, 115.38) | 14.18 | 53 (28, 25) | 0.23 (-0.27, 0.73) |

**Table 22: Secondary outcomes—social and behavioural measures at twelve-month follow-up (all eligible FSM pupils)**

| Outcome | Raw scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **SDQ** | | | | | | | | |
| Prosocial Score | 27 (3) | 7.26 (6.28, 8.24) | 2.55 | 25 (3) | 7.32 (6.51, 8.13) | 2.01 | 46 (24, 22) | -0.01 (-0.55, 0.53) |
| Total Difficulties Score | 27 (3) | 10.07 (7.50, 12.65) | 6.68 | 25 (3) | 9.48 (7.26, 11.70) | 5.55 | 46 (24, 22) | 0.12 (-0.42, 0.67) |
| Externalising Score | 27 (3) | 6.63 (4.75, 8.51) | 4.88 | 25 (3) | 5.80 (4.14, 7.46) | 4.14 | 46 (24, 22) | 0.24 (-0.31, 0.78) |
| Internalising Score | 27 (3) | 3.44 (2.26, 4.63) | 3.07 | 25 (3) | 3.68 (2.43, 4.93) | 3.12 | 46 (24, 22) | -0.10 (-0.65, 0.44) |
| Impact Score | 18 (2) | 1.39 (0.68, 2.10) | 1.50 | 14 (1) | 1.21 (0.42, 2.00) | 1.48 | 29 (16, 13) | 0.00 (-0.70, 0.70) |
| **PARYC** | | | | | | | | |
| Supporting Frequency | 12 (1) | 42.00 (37.50, 46.50) | 7.79 | 8 (1) | 41.75 (37.90, 45.60) | 5.44 | 18 (11, 7) | -0.17 (-1.07, 0.73) |
| Supporting Problematic | 12 (1) | 0.67 (0.10, 1.24) | 0.98 | 8 (1) | 1.25 (0.52, 1.98) | 1.04 | 18 (11, 7) | 0.16 (-0.73, 1.06) |
| Proactive Frequency | 12 (1) | 40.42 (35.69, 45.14) | 8.18 | 8 (1) | 37.75 (32.16, 43.34) | 7.91 | 18 (11, 7) | 0.14 (-0.76, 1.04) |
| Proactive Problematic | 12 (1) | 0.83 (-0.08, 1.75) | 1.59 | 8 (1) | 1.13 (0.09, 2.16) | 1.46 | 18 (11, 7) | 0.23 (-0.67, 1.12) |
| Total Frequency | 12 (1) | 82.42 (73.84, 90.99) | 14.85 | 8 (1) | 79.50 (70.64, 88.36) | 12.52 | 18 (11, 7) | -0.00 (-0.90, 0.89) |
| Total Problematic | 12 (1) | 1.50 (0.08, 2.92) | 2.47 | 8 (1) | 2.38 (0.87, 3.88) | 2.13 | 18 (11, 7) | 0.22 (-0.68, 1.12) |

**Year 1 Phonics results**

**Table 23: Year 1 Phonics results (all eligible FSM pupils)**

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| Year 1 Phonics | 82 (11) | 27.27 (24.55, 29.99) | 12.30 | 87 (7) | 27.53 (25.08, 29.98) | 11.42 | 151 (71, 80) | -0.11 (-0.41, 0.20) |

**Key Stage 1 assessment literacy results**

**Table 24: Key Stage 1 assessment results (all eligible FSM pupils)**

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| KS1 Assessment Reading/ Writing Score | 46 (7) | 13.28 (12.20, 14.36) | 3.67 | 43 (6) | 12.47 (11.49, 13.44) | 3.21 | 76 (39, 37) | 0.10 (-0.31, 0.52) |
| KS1 Assessment Reading Score | 46 (7) | 13.87 (12.72, 15.02) | 3.90 | 43 (6) | 12.86 (11.84, 13.88) | 3.34 | 76 (39, 37) | 0.17 (-0.25, 0.59) |

**Gender**

No differences were found for the primary and secondary outcomes at immediate post-test. Statistically significant differences at the conventional level were, however, found between the results of boys at six- and twelve-month follow-up, although follow-up results were not available for all boys in the trial. At six months, boys in the intervention condition demonstrated positive effects representing approximately three months' progress in Word Identification and Word Attack compared to those in the control condition (ES = 0.25; p = 0.04 and ES = 0.24; p = 0.05, respectively). At twelve-month follow-up, results representing approximately seven months' progress were seen for Word Attack (ES = 0.53, p = 0.02). In addition, although the KS1 SATs outcomes were not included in the protocol and results were only available for cohorts one and two, positive effects were also found for KS1 reading outcomes (ES = 0.41; p = 0.02), representing approximately five months' progress in favour of the intervention. There were no statistically significant differences found in any of the measures at any time point for girls. The results of this analysis are presented in Appendix D.

**Low achieving children at baseline (score of 95% or less at baseline)**

There were no statistically significant differences in primary and secondary (social and behavioural) outcomes at any time point except for the 'Supporting Frequency' (supporting positive behaviour) measure at six-month follow-up where an effect size of 0.59 was seen (p = 0.03). The results of this analysis are presented in Appendix E.

**On-treatment outcomes: parental level of participation in the SPOKES programme**

The SPOKES developers specified that attendance at five or more sessions of the programme would be regarded as 'completing' the programme. Using this measure, no statistically significant differences in primary literacy outcomes, or the Year 1 phonics assessment results, were found between those who completed the programme and the control condition. There were, however, positive, statistically significant differences in the non-pre-specified Key Stage 1 reading (p = 0.01) and reading/writing (p = 0.03) measures with corresponding effect sizes of 0.42 and 0.35 representing approximately five and four months' progress for those receiving the SPOKES programme respectively. Results are presented in Appendix F.

## Cost

The costs for SPOKES are somewhat different than for most of the EEF projects in that the intervention is delivered to the parents, not the children, thus there may not be a direct cost to the school. The programme could be funded by a school, or it could be funded by a social service organisation.

Details of the costs of implementing the SPOKES programme were obtained from Plymouth Parent Partnership. These costs are primarily staff costs to implement the weekly SPOKES sessions. The cost for the programme implementation in this study was approximately £804 per place, which could be considered per child, per parent, or per family. A breakdown of the total costs can be seen in Table 25.

**Table 25: Breakdown of costs for programme delivery (3 cohorts)**

| | Cost |
|---|---|
| Education psychologists (1 day per week) | £68,040 |
| Parenting Programme Facilitators (1.5 days per week) | £60,312 |
| Training | £5,200 |
| Administration | £10,145 |
| Supervision | £4,500 |
| Childcare (crèches for families) | £7,251 |
| Books | £7,085 |
| Resources | £2,250 |
| Venues* | £8,921 |
| **Total** | **£173,704** |

** Not all venues charged for room use. Cost per room was approximately £450 per group.*

These costs were calculated over three rounds of the programme (18 courses across the three cohorts—a total of 216 parental places). The total costs for these were £173,704.

This was an efficacy trial, designed to be delivered under ideal conditions with additional support by the developer. A small proportion of these costs, therefore, would not apply to a normal implementation of a scaled-up programme. For example, the training costs would be reduced over time once group leaders were already trained; in addition, if sessions were held in schools free-of-charge, venue costs could be reduced making the average cost per place £794. However, this is still a relatively expensive intervention, albeit comparable in cost to other parent programmes.

# Process evaluation

This section summarises the key findings from the process evaluation. It includes attendance data collected by group leaders, an overview of the responses of the subgroup of intervention parents who took part in the telephone interviews, summaries of the evaluations of SPOKES that parents completed at the end of the programme, and a sub-analysis of the weekly evaluation returns.

## Implementation

As this evaluation of SPOKES was an efficacy trial—a trial of the programme run in optimum conditions—high standards of implementation were required to ensure the programme was delivered as it was designed. Implementation fidelity was ensured by the developers and the Plymouth Parent Partnership through weekly supervision meetings with the developers and group leaders. These supervision sessions involved detailed observations of sessions recorded by group leaders, and interrogation of specific activities, with an emphasis on fidelity to the manual.

An important factor in terms of implementation fidelity is the dosage—in this case, attendance at parenting sessions. As the programme was not based around a particular school or location, strategies were put in place to boost attendance, for example through providing taxis for those who would otherwise struggle to attend and providing, where possible, city centre locations for sessions. In addition, a requirement for eligibility for inclusion in the trial was the ability to attend sessions if allocated to the intervention condition.

### Parental attendance

This section includes a summary of the number of sessions that parents attended, and the main reasons provided for any absences.

The intervention involved ten sessions with parents. As indicated in Table 26, attendance at SPOKES sessions was lower than anticipated. Over a quarter (27%) of parents assigned to the programme did not attend any sessions. Only 7% attended all sessions. As noted, attending five sessions was considered ‘completion’: just over half of participants (57%) who were allocated to the intervention achieved this. This remains relatively consistent across cohorts.

**Table 26: Attendance at SPOKES sessions per cohort as reported by the Plymouth Parent Partnership**

| Cohort | C1 | C2 | C3 | C4 | C5 | C6 | Total |
|---|---|---|---|---|---|---|---|
| Number of participants (intervention only)* | 50 | 65 | 54 | 69 | 57 | 75 | 370 |
| **Number of Sessions Attended** | | | | | | | |
| 0 | 17 | 15 | 19 | 18 | 13 | 18 | 100 |
| 1 | 5 | 3 | 2 | 5 | 3 | 7 | 25 |
| 2 | 0 | 5 | 0 | 2 | 2 | 4 | 13 |
| 3 | 1 | 1 | 2 | 2 | 2 | 2 | 10 |
| 4 | 1 | 2 | 3 | 2 | 1 | 2 | 11 |
| 5 | 2 | 1 | 5 | 4 | 2 | 3 | 17 |
| 6 | 2 | 5 | 3 | 2 | 4 | 5 | 21 |
| 7 | 4 | 7 | 4 | 3 | 8 | 5 | 31 |
| 8 | 7 | 9 | 8 | 13 | 9 | 10 | 56 |
| 9 | 11 | 10 | 5 | 10 | 11 | 14 | 61 |
| 10 | 0 | 7 | 3 | 8 | 2 | 5 | 25 |
| **Completers (attended 5+ sessions)** | | | | | | | |
| Number | 26 | 39 | 28 | 40 | 36 | 42 | 211 |
| Percentage (%) | 52 | 60 | 52 | 58 | 63 | 56 | 57 |

** Excludes ineligible participants and those who withdrew consent to use data.*

**End of programme and weekly evaluation forms**

One hundred and seventy-three end of programme evaluation forms were completed by participants. The original evaluation form was completed by 49 parents in the intervention condition in cohorts one and two. The form was then altered to make the questions more valuable for evaluation purposes. The revised evaluation form is given in Appendix G, along with the numbers and percentages for each Likert-scale response. This was completed by 124 parents in cohorts three to six.

Overall, it can be seen that parents were very happy with the programme, with the vast majority finding aspects of the SPOKES techniques and content 'useful' or 'very useful', the methods used in group sessions 'useful' or 'very useful', and group leaders 'helpful' or 'very helpful'.

Three to four per cent of respondents found 'using environmental print', 'visiting the library', and the 'one-to-one visit' 'not very useful' or 'not at all useful' (although only four parents per item expressed these views). All of the weekly evaluation forms were, however, filled in by intervention parents regarded to have 'completed' the programme.

The weekly evaluation forms from sessions one and two were analysed to see if there were any differences between the responses of 'completers' and 'non-completers'. These weeks were chosen because they were the only sessions for which there were regarded to be sufficient numbers of 'non-

completers' returning forms. The weeks were analysed separately because each session was on a different topic and this would ensure comparability between the two groups. A copy of the feedback form is provided in Appendix H. Two-hundred and forty-one respondents completed the feedback forms for session one (49 non-completers and 190 completers), and 185 for session two (20 non-completers, 165 completers).

Analysis of the main questions regarding enjoyment of the session and its usefulness indicated that in both sessions the mean response of completers was higher (and therefore more positive) than that of non-completers (four and three respectively for both questions in session one; and five and four respectively for both questions in session two). Given the small numbers and diversity of individual responses, however, it is difficult to explain the reasons for these differences, although there is some indication that a combination of dislike of role-play, personal circumstances (including personal shyness), and the conditions of the session (such as room size and the presence of other parents) may have been factors. For example, free responses to the open-ended question 'What did you like least about today's session?' for the week one evaluation form indicate that 15% of responses to this question by completers named the role-play element of the session (17 out of 107 responses) compared to 13% of non-completers (3 out of 23 responses), 21% of non-completers indicated personal circumstances (including shyness) compared to 5% of completers and 25% and 26% of completers and non-completers respectively mentioned the conditions of the sessions. In contrast 40% of completers compared to 25% of non-completers indicated 'nothing' in response to this question.

**Parent interviews**

Evaluation team staff called 130 parents from the intervention group to participate in brief interviews at each post-test time period. Interviewers attempted to reach parents up to four times but only reached 59 parents. Reasons for inability to contact remaining parents included having the wrong number, getting no answer, or the line being always busy.

**Attendance from parent interviews**

In contrast to recorded attendance by group leaders, in the telephone interviews the majority of parents reported missing very few sessions, and 80% of respondents (56 out of 71) reported that they attended five or more sessions—somewhat higher than the attendance recorded by the group leaders for the whole treatment group where only 57% were recorded as 'completing' the programme using this definition. This was to be expected as those who attended more often were more likely to participate in the telephone call. In addition, there is likely to be a relationship between attitudes towards the programme and attendance.

Although their attendance levels were high, the telephone interviewees did provide some reasons for non-attendance at sessions. The principle reasons given for missing sessions were their own or family illnesses, work, and holidays. Some parents felt that the initial invitation letter and consent forms were unclear about the requirements of their participation. When asked what would make it easier to attend a few cited the location and the timing of the sessions.

Of the 12 parents in the telephone interviews who reported low or non-attendance (for three interviewees attendance information provided was incomplete), five reported this as being due to their own or a child's illness, and three reported changes in employment (including study) and/or workload. Three parents also reported issues with caring for other children, despite provision being made for childcare during attendance at the programme to improve accessibility. It should be noted, however, that two of these parents also reported illness and starting studying as other life changes impacting on their ability to attend the programme. Of the three remaining parents, one missed the first two weeks and then felt she couldn't attend because she 'may have missed too much'. Two, however, felt that they did not think their child needed any extra support so did not attend: 'my child is not a struggling reader'. This might indicate that some schools nominated children who did not meet the specified criteria or used different criteria for recruitment to the programme.

Finally, due to the research design, there was a considerable gap between consent, randomisation, and the start date for the parenting sessions. When family illness or a house-move cropped up, families could not defer a term because of possible cohort effects. So, means of coping with changes in family circumstances that could be dealt with in a rolled out version of the programme were not possible in the study meaning these families were lost to the evaluation.

**Use of SPOKES strategies or activities**

During the interviews, parents reported using many of the strategies taught in the SPOKES programme. By far the most frequently mentioned was the use of Pause, Prompt and Praise. This must have been taught, and repeated, frequently as many parents used this phrase to report that they had been praising their children, and pausing and encouraging them to read on when they were stuck. Many of them cited literacy tips that they used when reading with their children, playing with their children, and ignoring their bad behaviour indicating that they used them frequently, and not only with their Year 1 'struggling reader'. One parent, for example, reported: 'I use [the techniques] with all our children during family time or homework time'.

When asked what strategies they did *not* use, there were a number of responses, with the two most frequently cited being working with letters and puppets.

**Overall parental satisfaction with the programme**

In the immediate post-programme interviews, parents were overwhelmingly positive about SPOKES. Fifty-six of the 59 parents interviewed responded enthusiastically about it. They felt that they, their children, and their families had benefitted from their participation: 'It made me feel like we were all in it together'; 'My son made big steps after this [the programme]'. Of particular importance was being able to use the ideas at home: '[I enjoyed] putting into place the ideas which [the group leaders] gave to us. It was nice to take them home and watch them work. I always looked forward to trying out what they suggested'. Most parents would recommend it to other families. Of the three who had reservations, two felt it would be better for younger children and one said that her daughter didn't need it: '[I] didn't dislike it but … [I] felt I didn't need to be there'.

In the parent interviews, parents were asked what they liked most about the SPOKES sessions. Responses were very positive and varied. The most often mentioned activities were group support, learning the Pause, Prompt and Praise strategies, behaviour management techniques, making puppets, and sharing ideas.

There were not many aspects of the programme that parents reported disliking. Of those that were reported, role-play within sessions was the most frequently given answer with 16 parents disliking this aspect, primarily because they were shy or lacked confidence.

**Follow-up interviews with parents**

The follow-up parent interviews which took place at six months (41 interviews, cohorts one to five) and 12 months after completion of the programme (nine interviews, cohorts one and two) revealed similar patterns of findings among the cohorts and across the two time periods regarding programme satisfaction and strategies used. However, because these took place some time after the end of the programme, the comments were somewhat more general than in the immediate post-test interviews.

Their overall reaction to the programme remained positive: parents really enjoyed it, reported learning a lot, and said they would recommend it to other parents. The activities that were reported most often as still being used by parents were Pause, Prompt, Praise, reading together, and increased wait time for the children to read words. At the six-month follow-up interview a few parents remembered not liking the role-plays, or sharing sessions with others, because they had felt self-conscious and shy.

**Suggestions from parents**

From the parent interviews, there were a number of suggestions about what to change about the programme but there was no consistency in those suggestions. Some thought SPOKES should be longer, some shorter; a few wanted more focus on literacy and less on parenting, and a few suggested that it would be more appropriate for younger children (such as those in reception or nursery). A few parents reported that the initial letter that they had received about the programme was not clear meaning that they did not really understand what they were committing to: '[The] letter that went out initially… wasn't clear what it was going to be about. A lot of parents didn't turn up because they thought it was just about buying books' and that messages from schools could differ: '[I] was told by the school that the course was only offered to kids with difficulties but then on the course no other parent was told this. They were just offered the course'.

# Conclusion

## Discussion

The Teaching and Learning Toolkit reports that overall parental involvement in children's schooling results in 'moderate impact for moderate cost, based on moderate evidence'. Two recent meta-analyses from the USA suggested that increasing parental involvement in primary and secondary schools had, on average, two to three months' positive impact (Jeynes, 2005; Nye, Schwartz, and Turner, 2006). These findings cover a wide variety of interventions at different age levels and for different subjects, some for regular students, and some for those struggling. The main results of this study found weaker effects than the previous studies quoted. In addition, previous research on a similar SPOKES programme (which had a greater focus on social behaviour) found it to be effective in improving both the literacy and behaviour of five- and six-year-old children at risk of social exclusion through antisocial behaviour (Sylva *et al*, 2008) with an effect size (*d*) of 0.3 for the British Ability Scales word reading sub-test (p. 447).

The overall intention-to-treat findings from this efficacy trial of SPOKES indicate that the programme did not have an impact on children's literacy. The only outcome measures that showed a statistically significant effect in favour of the treatment group were the KS1 reading and writing measures. These measures were not pre-specified in the trial protocol and the analysis was only able to be conducted on a sub-sample of participants (cohorts one and two). As the KS1 assessments were administered and marked by the children's teachers, they are less reliable than the standardised reading measures or than the end of Year 1 phonics assessment, both of which are more objective measures. This would be true across both control and intervention groups which somewhat increases the validity of these findings, although in some cases teachers may have been aware of parental attendance at SPOKES sessions. The primary outcomes measures (WRMT-R sub-tests) can be considered more reliable than the KS1 results and other teacher-completed or parent-completed measures (such as the SDQ and the PARYC) as the individual assessors were blind to condition.

There were, however, some positive findings in the subgroup analysis relating to the impact of SPOKES on boys, particularly in the longer term. There were statistically significant findings at the conventional level for some, although not all, primary outcomes at both the six-month and twelve-month follow-up rounds of data collection, and in the Key Stage 1 reading assessments. Although not all boys were included in these analyses, the findings suggest an impact of the SPOKES programme on some aspects of boys' literacy of between three and seven months' additional progress.

The secondary, social, and behavioural outcomes in this evaluation were inconsistent. However these, and the baseline measure, can be considered to be within the normal range which may suggest that enacting substantial improvement may be more difficult than, for example, in the previous SPOKES trials that targeted children with behaviour and literacy difficulties.

The process evaluation findings demonstrated that the programme was positively perceived by the parents who attended sessions, that they enjoyed their participation and, in fact, believed that it had improved their children's literacy. Although there are indications that parents introduced techniques learnt through the SPOKES programme in the home (such as Pause, Prompt, Praise) this is only through self-report and does not indicate the extent of any changes which occurred. The telephone interviews indicated that for the majority of those who were low- or non-attenders this was due to primarily unforeseen life factors, including illness and employment changes. Yet this was not a representative sample: in total only 57% of the parents participated in five or more of the ten parenting sessions—not a very high participation rate. This might account for the limited impact on children's reading. Attendance at other parenting programmes has varied across studies but the retention rates in this evaluation do appear to be lower than other parent programmes, and in other studies higher attendance was correlated with positive effects on parenting (Reid *et al.*, 2004) although this is not

evident in this evaluation. This may, however, reflect the very low criteria for being considered a 'completer'—attending only five of the ten sessions.

## Limitations

There were, however, some limitations to the evaluation which may have impacted on the outcomes. These were primarily related to recruitment to the programme, a delay between recruitment and implementation, and low attendance.

The identification by schools of struggling readers does appear to have been problematic, with the majority of children not being defined as such by the baseline measure (or, in some cases, by their parent), but rather by the children's teacher (as discussed in the 'eligibility' section above). Because only a low proportion of children scored poorly on the standardised baseline measure (for this evaluation a 'poor' score was a score of 95% or less of the mean at baseline, which was the case for 28% of the total sample for analysis), the programme can be said to have not been primarily delivered to its intended target population. This could have reduced any impacts, both overall and on those children who were struggling readers (because the group-based nature of the programme would have meant that parents would not have the same shared experience to draw upon, for example in group discussions and role-play activities). In addition, between the baseline pre-test measure and the post-test, those who were lost from the control condition had, on average, lower baseline attainment than those retained in the study, and those pupils lost from the intervention condition with higher baseline attainment were lost to follow up. Within this context it should also be noted that children in the control condition did receive free books and newsletters in order to encourage recruitment and retention to the study.

The early recruitment of parents to the programme in order to facilitate the evaluation (in particular to allow for baseline measures to be collected) meant that there may have been a long time lag between recruitment and the start of the programme for some parents. Consequently, when there was a change in circumstances, such as family illness, house-move, or change in employment circumstances—more likely to happen over a longer time period—families were lost to the evaluation as the research design meant that families could not defer a term because of possible cohort effects. Therefore, ways of coping with changes in family circumstances that could be incorporated into a rolled out version of the programme were not possible in the study, meaning these families were lost to the evaluation. This impacted on recruitment and may also have impacted on the profile of those recruited (such as in favour of less struggling readers). Also, parents did not necessarily attend the programme close to their own locality or with other parents from their child's school. As indicated above, attendance on the programme was low and it could be queried if the definition of 'completer' as having attended five or more sessions could be considered sufficient to have had programme impacts. Consequently the programme could be considered not to have been delivered to the optimal design desired in an efficacy trial.

As differences between schools can be large, not taking school into account in the analyses could mean that the effect of treatment appeared to be a lot smaller, and so potentially not statistically significant, than would have been the case in a clustered analysis. However, we chose not to conduct a clustered analysis due to the small number of pupils recruited from many of the schools, and because group attendance for intervention parents was not linked to the school their child attended. Conversely, although smaller than anticipated numbers were generally recruited from each school, the recruitment process was the same for all parents, which may have led to some contamination effects (for example, parents allocated to different conditions may have discussed the programme and received the same information regarding the programme at the 'coffee morning' events). In the analyses all statistically significant differences were in favour of the intervention, yet because so many statistical analyses were conducted, the likelihood of finding some statistically significant differences by chance increases. Importantly, while the baseline and outcome measures used were all designed to assess literacy and behavioural outcomes in the target population relevant to this study, the correlation between the pre-test and the primary outcome was lower than expected in this study.

Further limitations to the design which could have impacted on outcomes include a possible cohort effect and the mean delay in some post-testing resulting from the delivery model of the programme. Given that the programme was delivered across the school year, the term in which it was delivered could have had an impact on outcomes. In other words, there could be a time in the school year during which the programme could have higher impact on student outcomes than at other stages. The research design also meant that some delay in post-testing did occur (for example, for those children whose parents participated in the programme during the Summer term post-testing was delayed by the long summer break). In the analysis, the impact of cohort on treatment effect was, however, inconsistent and therefore excluded from the model. However, one cohort (cohort six) was not included in the six-month follow-up and only two cohorts had completed the programme by the time of the twelve-month follow-up round of data collection. Consequently these findings may not be representative of the whole sample.

Finally, the findings of this evaluation also relate to a particular geographical location, primarily Plymouth, which limits the generalisation of the findings to a wider population. In particular, one aspect not addressed by this evaluation was the prevalence of other parenting programmes in the Plymouth area—such as the Incredible Years, parts of which were used in the development of SPOKES)—that may also have had an impact on outcomes.

## Future research and recommendations

Positive effects were found on the Key Stage 1 literacy assessments for children who had reached the end of KS1, and on some medium term outcomes for boys. While these results did not include all cohorts of children involved in the evaluation, they do suggest that wider outcomes should be measured as they imply that the greatest added value of the programme might be in developing the wider literacy skills used at KS 1 rather than those measured by the primary outcomes. The positive follow-up outcomes for a subsample of boys in this study (at six- and twelve-month follow up), taken in conjunction with the KS1 results, also suggest that the SPOKES programme had a greater impact on boys than on the sample as a whole. Consequently, future research on the SPOKES programme would be warranted with particular focus on both the wider outcomes of the programme and its impact on boys' literacy.

Given the low level of attendance of intervention parents in this evaluation, future research should also consider factors relating to recruitment, attendance, and implementation fidelity. Prior to any future evaluation, the SPOKES programme might benefit from a 'service design' exercise. This would be particularly important in considering any possible roll-out of the programme as SPOKES sessions are quite expensive to run and time-intensive for parents. Definitions of 'completers', those who have completed the programme, may need to be restricted to those attending a higher number of sessions. If attendance was higher, and the definition of 'completers' required a higher attendance rate, there might be more impact of the intervention.

In particular, service deliverers, the developer, and referrers, need to be clear on the recruitment process and who is targeted. The fact that a lot of the participating children were not 'struggling readers' as originally intended for the programme may also suggest revisions are needed to the way in which low attaining pupils are identified in schools (such as through the Early Years Foundation Stage profile) and how 'struggling reader' is defined. A future evaluation may benefit from following up on children's comprehension skills, a more advanced and critical feature of reading, in addition to the early literacy skills assessed as part of this evaluation.

Perhaps the developers and providers should try to find ways to increase the participation rate as it might then have greater impact on parents' support for their children's literacy and thus improve children's reading achievement, even though in this case a dose-response relationship was not found. In particular, parents found the programme to be successful in improving their children's reading, and they reported benefitting from the programme themselves.

Consequently, we feel it would be worth refining the programme and evaluating it again, to include:

- more attention to fidelity through a more thorough process evaluation to assess both what is happening during SPOKES sessions and implications for scalability of the final model; and
- wider outcome measures, with perhaps more focus on the literacy achievement of boys.

# References

Dunn, L. M., Dunn, L. M., Whetton, C. and Burley, J. (1997) *The British Picture Vocabulary Scale: Second Edition*, London, UK: GL Assessment.

Clay, M. (1991) *Becoming literate: The construction of inner control*, Auckland, NZ: Heinemann.

Gardner, F. (March, 2013) 'Parenting interventions for child problem behaviours: For whom, where, and why do they work?', Presented to the South Wales Conference, Cardiff, Wales.

Jeynes, W. H. (2005) 'A Meta-Analysis of the Relation of Parental Involvement to Urban Elementary School Student Academic Achievement', Urban Education, 40 (3), 237–269.

McNaughton, S., Glynn,T. and Robinson, V. (1987). *Pause prompt and praise: Effective remedial reading tutoring*, Birmingham, UK: Positive Products.

Nye, C., Schwartz, J. and Turner, H. (2006) 'Approaches to Parent Involvement for Improving the Academic Performance of Elementary School Age Children: A Systematic Review'. *Campbell Systematic Reviews*, 2 (4).

Reid, M., Webster-Stratton, C. and Bydar, N. (2004) 'Halting the development of externalizing behaviors in Head Start children', *Journal of Clinical Child and Adolescent Psychology,* 33*,* 279–291.

Scott, S., Sylva, K., Beckett, V., Kallitsogloua, A., Doolana, M. and Ford, T. (2012) 'Should parenting programmes to improve children's life chances address child behaviour, reading skills, or both? Rationale for the Helping Children Achieve trial', *European Journal of Development Psychology*, 9 (1), 47–60.

Scott, S. and O'Connor, T. (2012) 'An experimental test of differential susceptibility to parenting among emotionally-dysregulated children in a randomized controlled trial for oppositional behaviour', *The Journal of Child Psychology and Psychiatry,* 53 (11), 1,184–1,193.

Sylva, K., Scott, S., Totsika, V., Ereky-Stevens, K. and Crook, C. (2008) 'Training parents to help their children to read: A randomized control trial', *British Journal of Educational Psychology*, 78, 435–455.

Sylva, K., Crook, C., Roberts, F. and Price, J. (2013) *SPOKES Literacy Programme – Working Manual for Group Leaders*, Oxford, UK: Authors.

Woodcock, R. W. (1998) *Woodcock Reading Mastery Tests*. Circle Pines, MN: American Guidance Service.

# Appendix A: Parent information sheet and consent form

## PARENT INFORMATION SHEET

**Invitation**
Dear Parent,

We are delighted that your school has agreed to participate in a study of children's literacy development starting in January 2013. The study is being run by Plymouth Parent Partnership, the University of Oxford and the University of York. We very much hope you and your child would like to take part, but before you decide, it is important that you understand why the study is being done and what it will involve for you and your child.

**What will happen if my child and I take part?**
If you would like to take part in the study, you will be offered participation in one of two programmes: one involves taking part in a Supporting Parents on Kids' Education (SPOKES) Literacy parent group, and the other involves literacy packs for your family. To make a fair allocation to each programme, the selection will be decided at random using a computer. Whichever programme you are offered, we would visit your child in school on four occasions, once before the programme and three times afterwards. The visits would focus on literacy activities and games. We have used these activities in our research before and know that children find them fun and enjoy doing them. For most children each visit will consist of a short one-to-one session with a researcher. For some children, each visit will consist of three short sessions. Before and after both programmes, we would like to ask some parents a few questions about things you do with your child at home, and we hope to do this on the telephone at your convenience.

**What are we trying to find out?**
The study will help us to work out the best ways to help parents support their children's reading at home, and what are the key aspects in children's literacy development. The SPOKES Literacy parent programme is a 10-week group intervention that helps parents to develop skills to support their child's reading. If you are offered this programme, you will be invited to attend 10 weekly group sessions about strategies to use when reading with your child. If you are offered the Literacy pack programme, we will send book packs to your family in your child's book bag over a 10-week period.

**What happens to the results of the study?**
A summary of findings will be given to participants' schools at the end of the study and will be available to interested families. We also aim to publish our findings in scientific journals, but this may be one or two years from the end of the study. No individual child will be identifiable in any published report of this work. Any personal information will be destroyed at the end of the study.

**Do I have to take part?**
It is entirely up to you whether or not you decide to be part of this research project and attend these sessions. If you do decide to participate, please keep this information sheet to refer to, and please sign the enclosed consent forms. You are free to withdraw from the research at any time and you do not need to give a reason. Your withdrawal from the study will not affect the support that your child will receive from the school.

**What should I do next?**
If you are happy for your child to participate in this research, then please fill in the enclosed form and return it to a Plymouth Parent Partnership staff member or your child's teacher. If you have any questions about this research project then please do not hesitate to contact us, our contact details are below.

We look forward to working with you and your child.

Sincerely,
Plymouth Parent Partnership, the Oxford Team and the York Team

**Plymouth Parent Partnership:**
Mrs Angela Archer
Jan Cutting Healthy Living Centre
Scott Business Park
Beacon Park Rd, Plymouth PL2 2PQ
Tel: 01752 258933
Email: parentpartnership@plymouth.gov.uk

**University of Oxford:**
Professor Kathy Sylva
Department of Education
15 Norham Gardens, Oxford, OX2 6PY
Tel: 01865 274008
Email: Kathy.sylva@education.ox.ac.uk

## Consent Form - Parents

I have read and understood the information about this study, and have had the opportunity to ask questions. I understand who will have access to information provided and what will happen to the data at the end of this study I also understand that I can withdraw at any time by informing the research team of my decision.

I understand that the study has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee and the University of York Education Ethics Committee.

**I agree to myself and my child participating in the project:**

Name: _____

Child's name: _____

Child's date of birth: _____

Child's gender: M / F

Name of school: _____

Date: _____

Signature: _____

Name (member of research team): _____

Date: _____

Signature:

# Appendix B: Parent telephone interview schedule

**Participant ID:…….**

**Date and time of call:** ……/……../…….. (DD/MM/YY) …..……..am/pm

**SEMI-STRUCTURED POST-TEST INTERVIEW**

**Area 1: Factors influencing attendance (do not read out this line)**

Qu 1: Did you miss any of the 10 sessions?

Qu 2: What were some of the things that made it difficult to attend?

Qu 3: Can you think of ways which would make it easier, or more attractive, for parents to attend?

**Area 2: Parent satisfaction of the SPOKES programme (do not read out this line)**

Qu 4: What part of the programme did you like most and why?

Qu 5: What part of the programme did you like least and why?

Qu 6: if there was anything you could change about the programme what would it be?

**Area 3: SPOKES strategies/recommended activities implemented by the parent (do not read out this line)**

Qu 7: Which techniques or activities that you learned in the programme have you used most often with CHILD's NAME (no prompting)?

Qu 8: Can you give some examples of how/when you use the techniques or activities you learned (if they are used)?

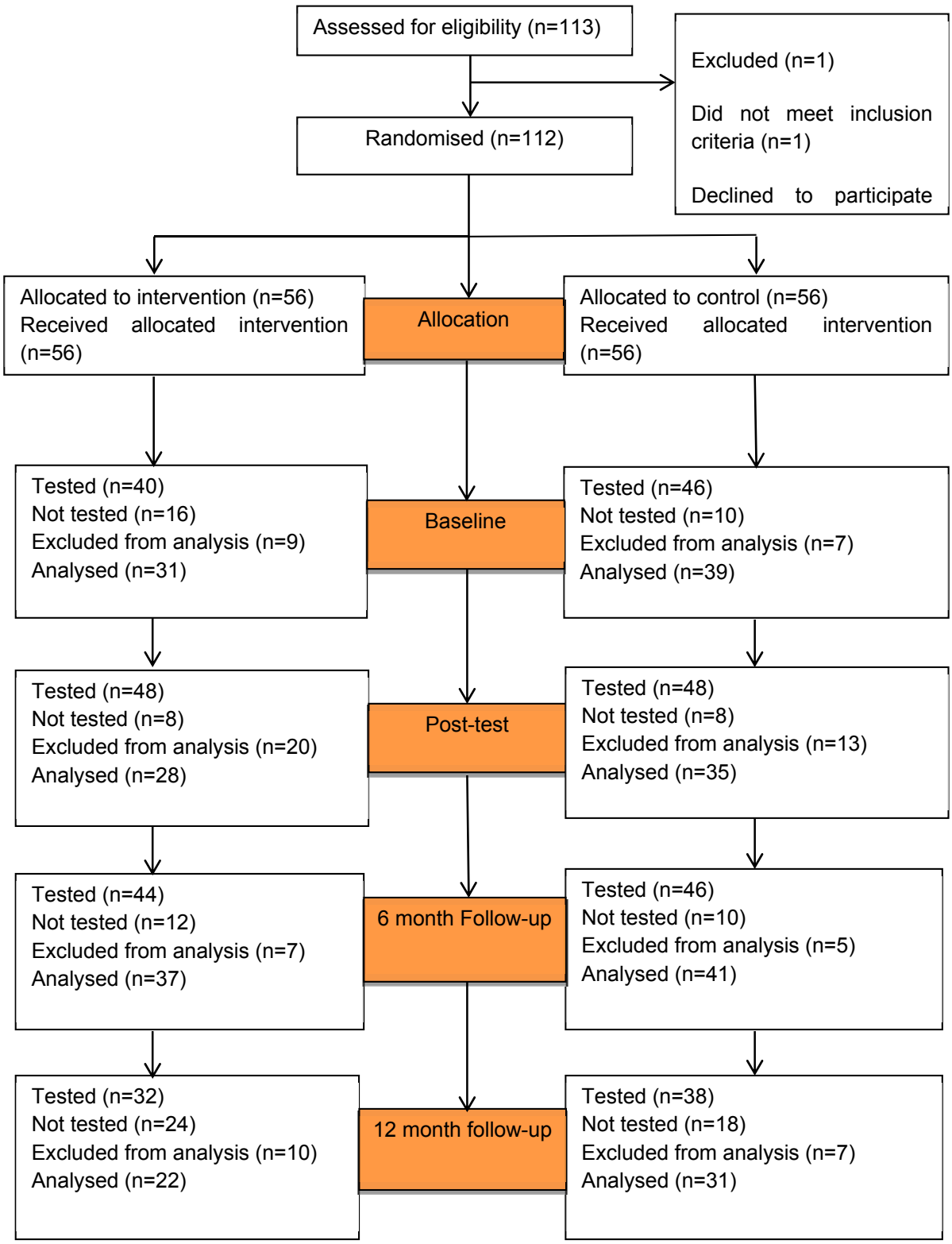
Qu 9a: What techniques or activities learned on the course did you try that you no longer use?

Qu 9b: (if there is a response to 9a) Why do you no longer use them?

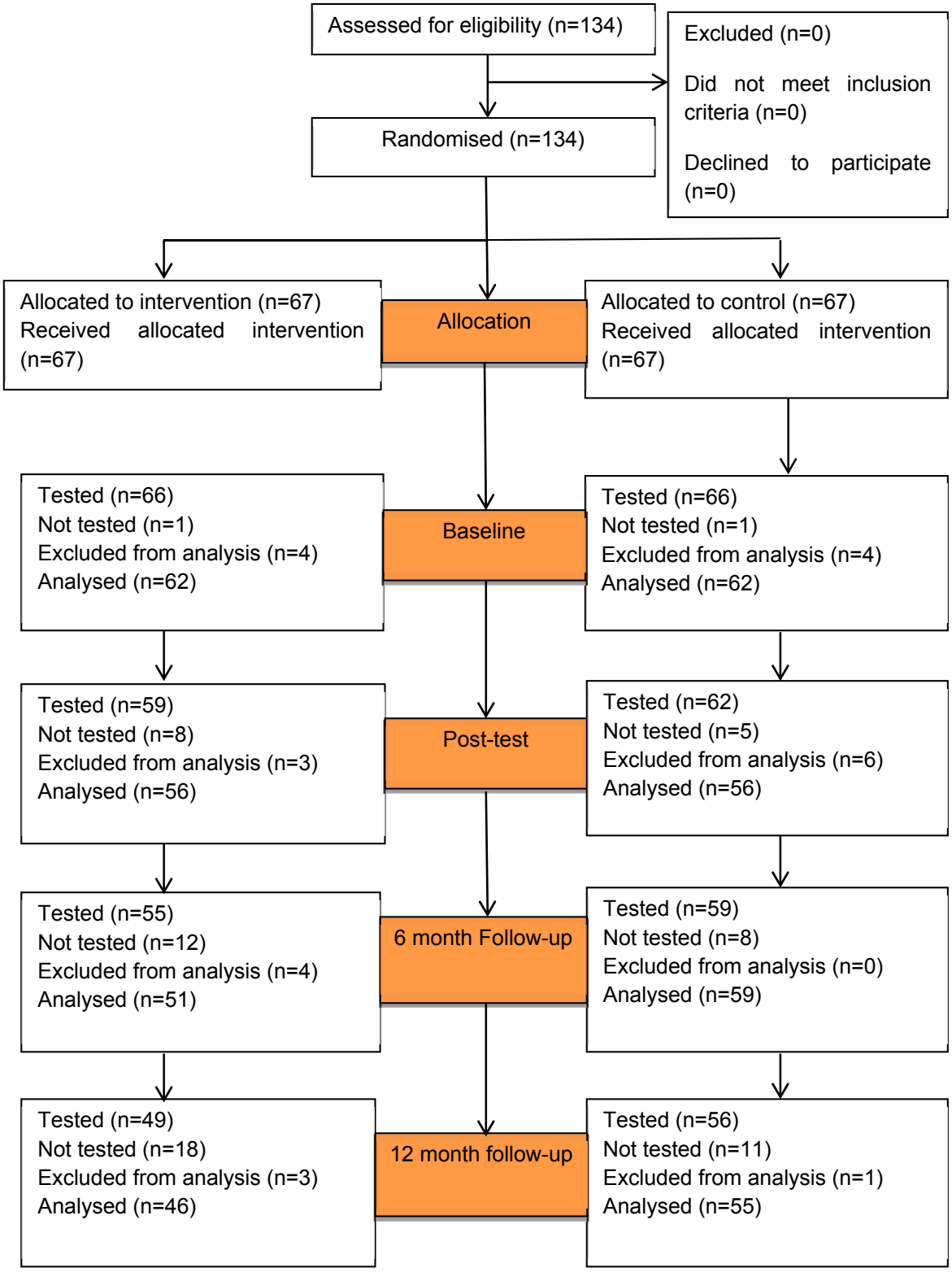
Qu 10: Can you think of a short sentence to sum up your overall view of the SPOKES programme?

# Appendix C: Consort diagrams by cohort

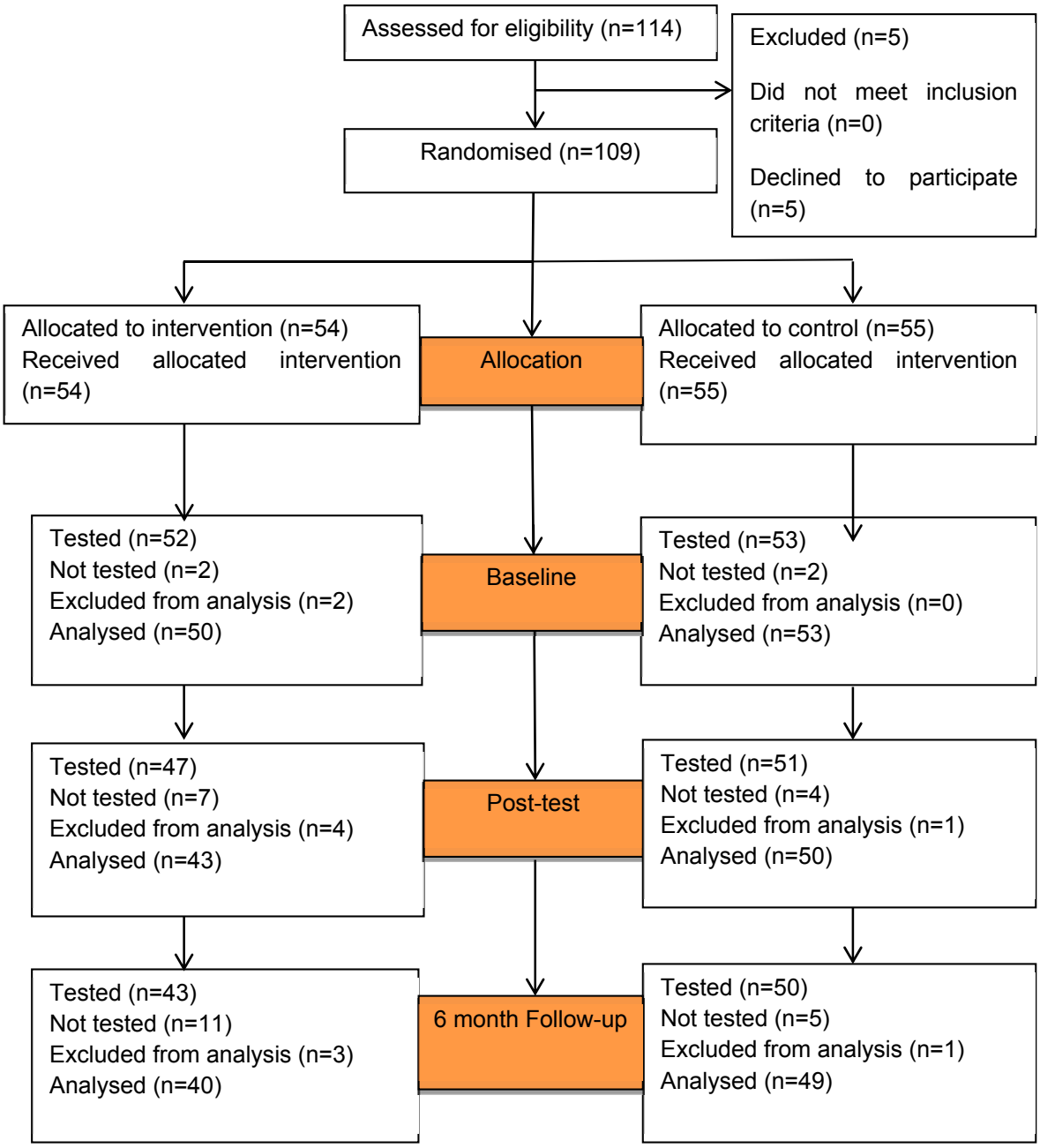
Cohort 1 (Spring Term 2013)



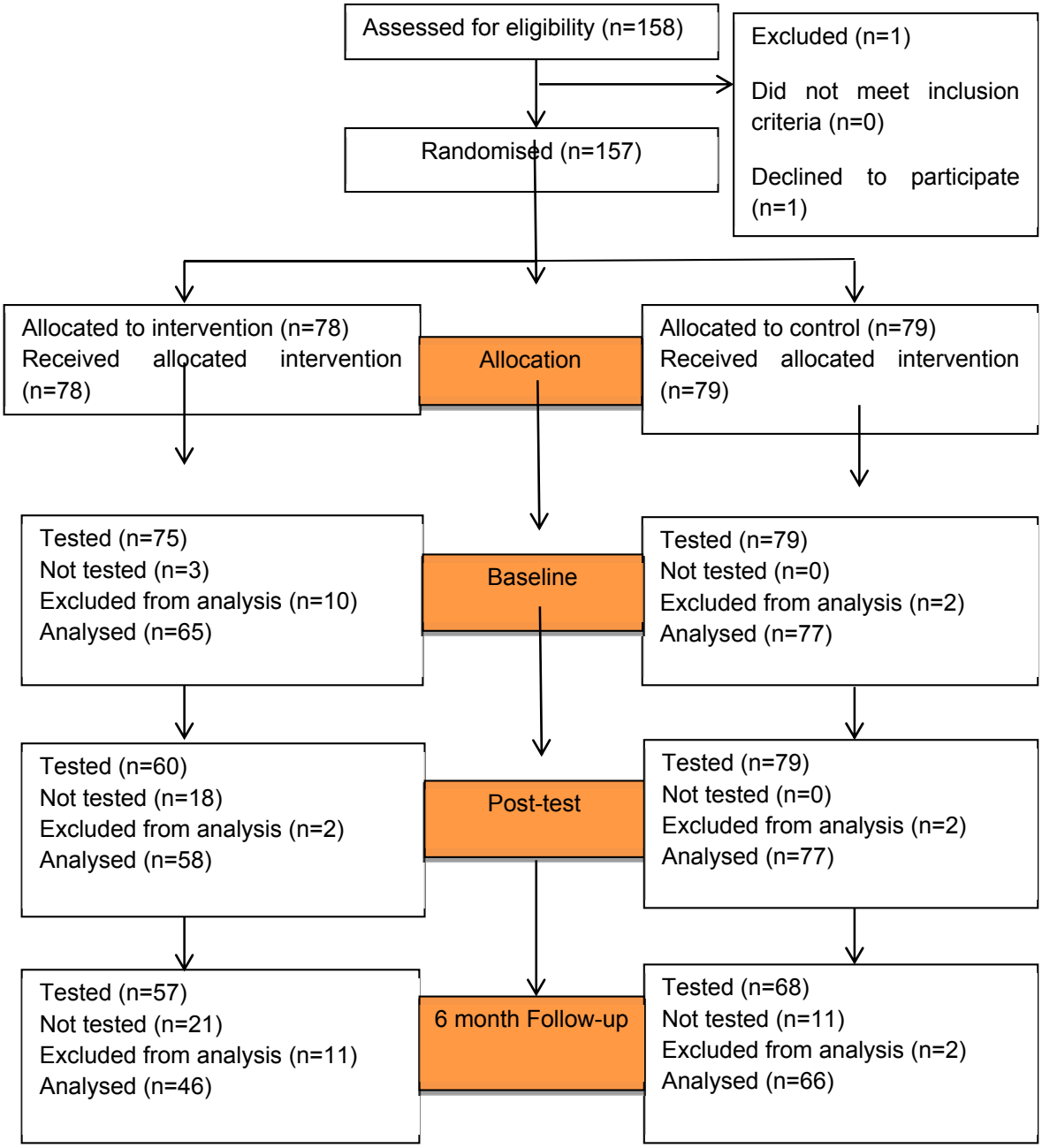
Cohort 2 (Summer Term 2013)



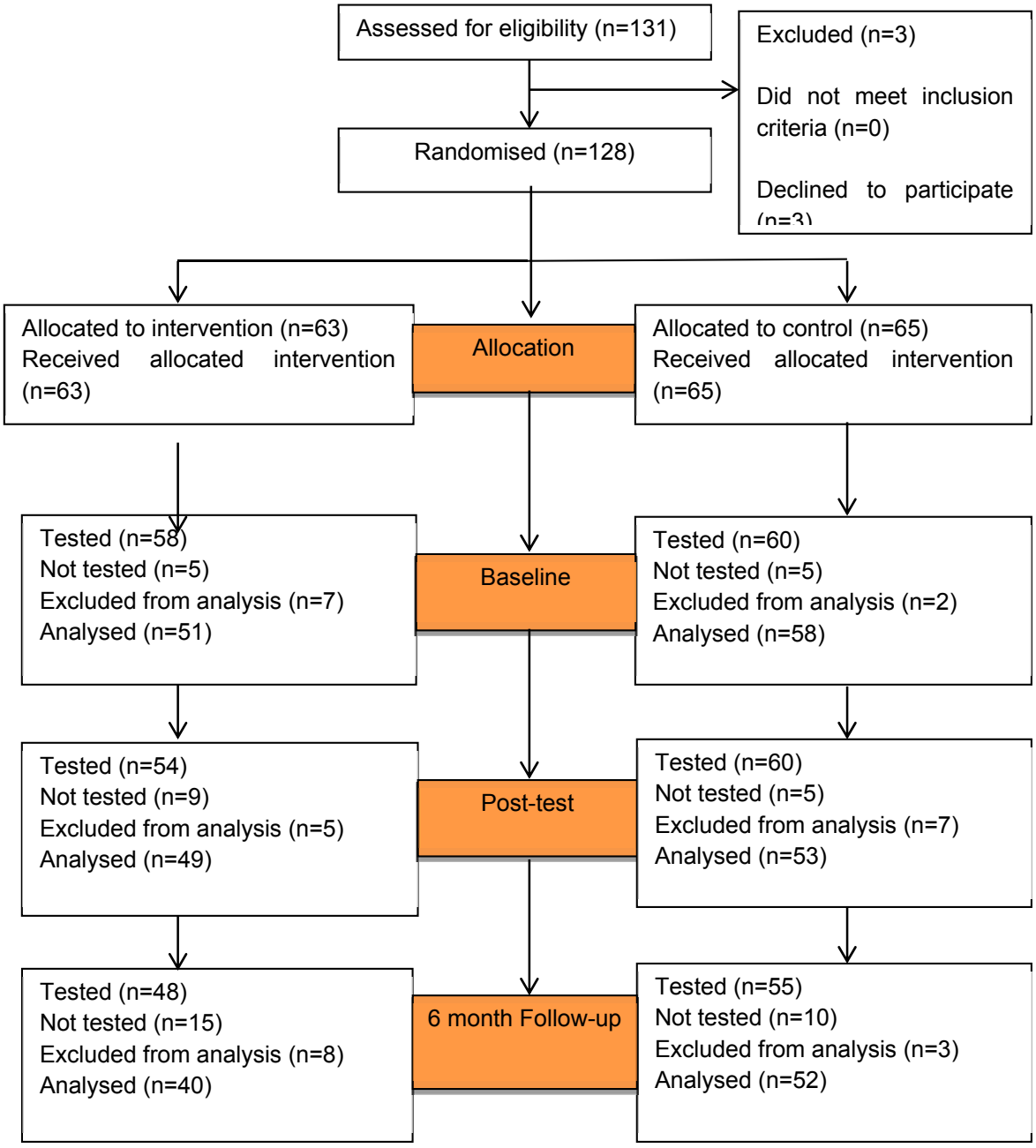
Cohort 3 (Autumn Term 2013)



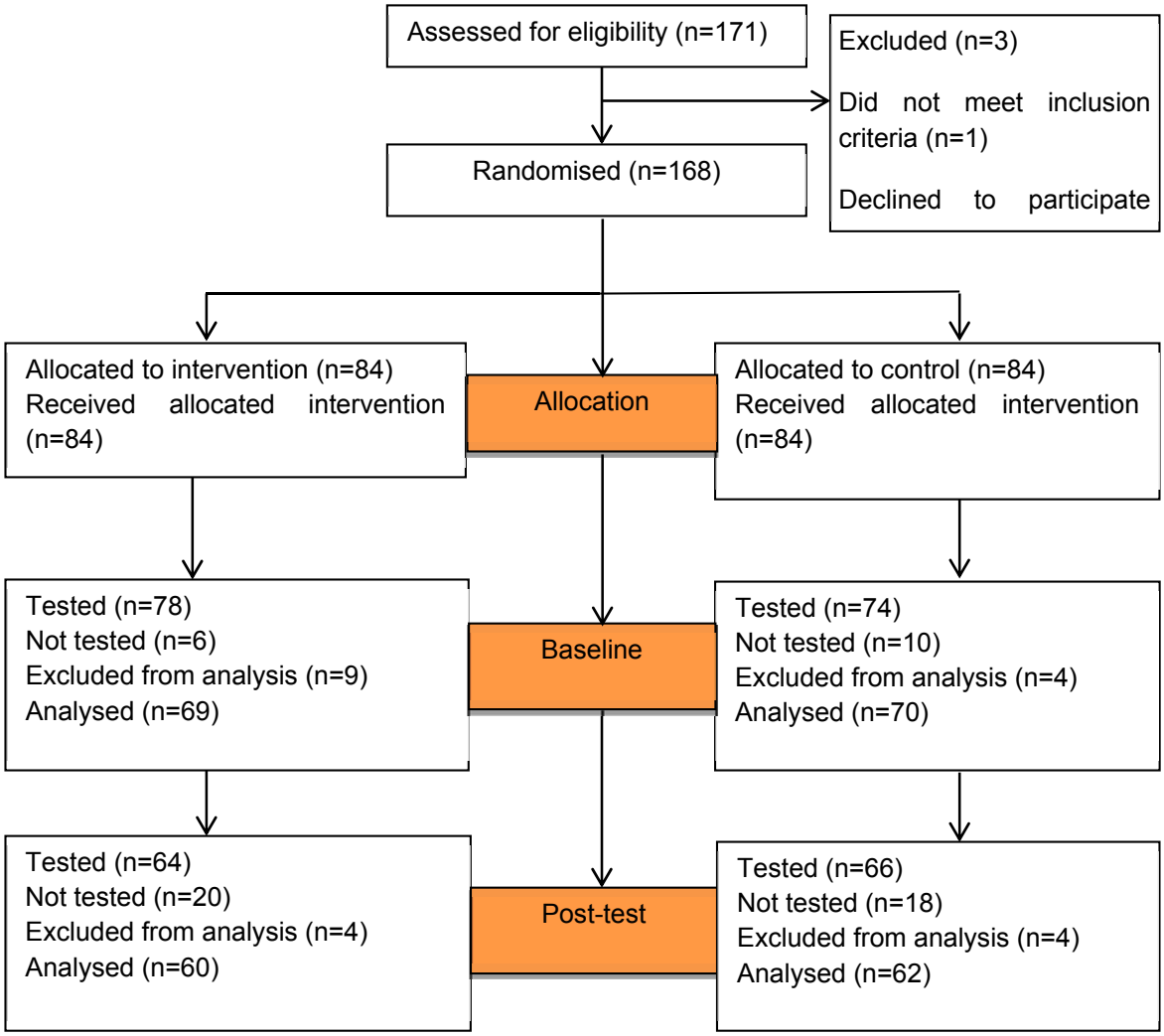
Cohort 4 (Spring Term 2014)



Cohort 5 (Summer Term 2014)



Cohort 6 (Autumn Term 2014)



# Appendix D: Outcomes according to gender

Table D.1: Primary Outcomes at post-test (all eligible boys)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **Letter ID** | 160 (14) | 107.06 (105.73, 108.39) | 8.42 | 220 (13) | 106.73 (105.69, 107.77) | 7.73 | 353 (146, 207) | 0.04 (-0.16, 0.25) |
| **Word ID** | 158 (13) | 120.91 (118.86, 122.95) | 12.87 | 215 (12) | 119.25 (117.52, 120.97) | 12.65 | 348 (145, 203) | 0.11 (-0.09, 0.32) |
| **Word Attack** | 154 (13) | 118.86 (117.47, 120.25) | 8.61 | 216 (12) | 118.04 (116.88, 119.20) | 8.51 | 345 (141, 204) | 0.07 (-0.14, 0.27) |

Table D.2: Secondary Outcomes – social and behavioural measures at post-test (all eligible boys)

| Outcome | Raw scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| SDQ | | | | | | | | |
| **Prosocial Score** | 139 (12) | 6.86 (6.43, 7.30) | 2.55 | 180 (8) | 6.37 (5.98, 6.75) | 2.58 | 299 (127, 172) | 0.17 (-0.06, 0.39) |
| **Total Difficulties Score** | 139 (12) | 8.41 (7.45, 9.37) | 5.67 | 181 (8) | 8.50 (7.61, 9.39) | 5.98 | 300 (127, 173) | 0.02 (-0.20, 0.24) |
| **Externalising Score** | 139 (12) | 5.28 (4.58, 5.98) | 4.11 | 181 (8) | 5.53 (4.90, 6.16) | 4.22 | 300 (127, 173) | -0.05 (-0.27, 0.17) |
| **Internalising Score** | 139 (12) | 3.13 (2.64, 3.62) | 2.90 | 181 (8) | 2.97 (2.49, 3.44) | 3.18 | 300 (127, 173) | 0.10 (-0.12, 0.32) |
| **Impact Score** | 73 (8) | 1.25 (0.90, 1.60) | 1.49 | 93 (6) | 1.26 (0.96, 1.55) | 1.43 | 152 (65, 87) | 0.17 (-0.14, 0.48) |
| PARYC | | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Supporting Frequency** | 89 (8) | 41.10 (39.93, 42.27) | 5.51 | 141 (9) | 40.84 (39.85, 41.84) | 5.93 | 213 (81, 132) | -0.00 (-0.27, 0.26) |
| **Supporting Problematic** | 89 (8) | 1.16 (0.89, 1.42) | 1.24 | 141 (9) | 1.21 (1.00, 1.43) | 1.29 | 213 (81, 132) | 0.00 (-0.26, 0.27) |
| **Proactive Frequency** | 87 (8) | 38.14 (36.56, 39.71) | 7.35 | 140 (9) | 39.49 (38.46, 40.53) | 6.11 | 210 (79, 131) | -0.10 (-0.37, 0.17) |
| **Proactive Problematic** | 87 (8) | 1.30 (0.98, 1.62) | 1.48 | 140 (9) | 1.22 (0.98, 1.47) | 1.45 | 210 (79, 131) | 0.00 (-0.27, 0.27) |
| **Total Frequency** | 87 (8) | 79.15 (76.75, 81.55) | 11.19 | 138 (9) | 80.41 (78.66, 82.16) | 10.27 | 208 (79, 129) | -0.08 (-0.35, 0.19) |
| **Total Problematic** | 87 (8) | 2.48 (1.99, 2.98) | 2.31 | 138 (9) | 2.46 (2.05, 2.86) | 2.38 | 208 (79, 129) | 0.01 (-0.26, 0.27) |

Table D.3: Primary Outcomes at 6-month follow-up (all eligible boys)

| **Outcome** | **Standardised scores** | | | | | | **Effect Size** | |
|---|---|---|---|---|---|---|---|---|
| | **Intervention** | | | **Control** | | | | |
| | **N (missing)** | **Mean (95% CI)** | **SD** | **N (missing)** | **Mean (95% CI)** | **SD** | **N in model (Int, Con)** | **Effect Size (95% CI)** |
| **Letter ID** | 131 (11) | 103.72 (102.49, 104.95) | 7.03 | 161 (9) | 102.44 (101.20, 103.68) | 7.85 | 272 (120, 152) | 0.19 (-0.04, 0.42) |
| **Word ID** | 121 (11) | 121.23 (118.63, 123.83) | 14.31 | 146 (9) | 117.81 (115.45, 120.16) | 14.24 | 247 (110, 137) | 0.25 (0.00, 0.49)* |
| **Word Attack** | 131 (11) | 118.86 (117.15, 120.57) | 9.78 | 160 (9) | 116.76 (115.22, 118.30) | 9.75 | 271 (120, 151) | 0.24 (0.00, 0.47)* |

*** p≤0.05**

Table D.4: Secondary Outcomes – social and behavioural measures at 6-month follow-up (all eligible boys)

| **Outcome** | **Raw scores** | | | | | | **Effect Size** | |
|---|---|---|---|---|---|---|---|---|
| | **Intervention** | | | **Control** | | | | |
| | **N (missing)** | **Mean (95% CI)** | **SD** | **N (missing)** | **Mean (95% CI)** | **SD** | **N in model (Int, Con)** | **Effect Size (95% CI)** |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| SDQ | | | | | | | | |
| **Prosocial Score** | 103 (11) | 6.69 (6.14, 7.24) | 2.77 | 122 (8) | 7.04 (6.59, 7.50) | 2.52 | 206 (92, 114) | -0.08 (-0.35, 0.18) |
| **Total Difficulties Score** | 103 (11) | 9.77 (8.41, 11.13) | 6.91 | 122 (8) | 8.18 (7.08, 9.28) | 6.09 | 206 (92, 114) | 0.29 (0.02, 0.55)* |
| **Externalising Score** | 103 (11) | 5.91 (4.97, 6.85) | 4.77 | 122 (8) | 5.04 (4.29, 5.79) | 4.15 | 206 (92, 114) | 0.20 (-0.06, 0.47) |
| **Internalising Score** | 103 (11) | 3.85 (3.17, 4.54) | 3.49 | 122 (8) | 3.14 (2.58, 3.70) | 3.09 | 206 (92, 114) | 0.29 (0.02, 0.55)* |
| **Impact Score** | 56 (6) | 1.61 (1.13, 2.08) | 1.79 | 64 (4) | 1.28 (0.93, 1.64) | 1.42 | 110 (50, 60) | 0.26 (-0.10, 0.62) |
| PARYC | | | | | | | | |
| **Supporting Frequency** | 66 (8) | 42.09 (40.69, 43.50) | 5.71 | 80 (7) | 41.84 (40.79, 42.88) | 4.68 | 131 (58, 73) | 0.04 (-0.29, 0.37) |
| **Supporting Problematic** | 66 (8) | 0.86 (0.53, 1.20) | 1.35 | 80 (7) | 1.29 (0.97, 1.61) | 1.42 | 131 (58, 73) | -0.29 (-0.62, 0.04) |
| **Proactive Frequency** | 64 (8) | 40.02 (38.47, 41.56) | 6.19 | 80 (7) | 40.65 (39.45, 41.85) | 5.38 | 129 (56, 73) | -0.12 (-0.45, 0.21) |
| **Proactive Problematic** | 64 (8) | 0.95 (0.58, 1.32) | 1.47 | 80 (7) | 1.19 (0.80, 1.57) | 1.72 | 129 (56, 73) | -0.18 (-0.51, 0.15) |
| **Total Frequency** | 63 (8) | 81.89 (79.22, 84.56) | 10.61 | 78 (7) | 82.56 (80.67, 84.46) | 8.36 | 126 (55, 71) | -0.08 (-0.41, 0.25) |
| **Total Problematic** | 63 (8) | 1.84 (1.21, 2.48) | 2.52 | 78 (7) | 2.50 (1.86, 3.14) | 2.80 | 126 (55, 71) | -0.27 (-0.60, 0.07) |

*** p≤0.05**

Table D.5: Primary Outcomes at 12-month follow-up (all eligible boys)

| **Outcome** | **Standardised scores** | | | | | | **Effect Size** | |
|---|---|---|---|---|---|---|---|---|
| | **Intervention** | | | **Control** | | | | |
| | **N (missing)** | **Mean (95% CI)** | **SD** | **N (missing)** | **Mean (95% CI)** | **SD** | **N in model (Int, Con)** | **Effect Size (95% CI)** |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Letter ID** | 44 (8) | 97.20 (94.98, 99.43) | 7.38 | 55 (5) | 94.75 (92.50, 96.99) | 8.31 | 86 (36, 50) | 0.14 (-0.26, 0.54) |
| **Word ID** | 34 (7) | 111.26 (106.41, 116.12) | 14.16 | 41 (5) | 105.83 (101.06, 110.60) | 15.26 | 63 (27, 36) | 0.30 (-0.16, 0.76) |
| **Word Attack** | 44 (8) | 115.55 (112.08, 119.01) | 11.49 | 53 (5) | 108.26 (104.21, 112.32) | 14.77 | 84 (36, 48) | 0.53 (0.13, 0.94)* |

*** p≤0.05**

Table D.6: Secondary Outcomes – social and behavioural measures at 12-month follow-up (all eligible boys)

| **Outcome** | **Raw scores** | | | | | | **Effect Size** | |
|---|---|---|---|---|---|---|---|---|
| | **Intervention** | | | **Control** | | | | |
| | **N (missing)** | **Mean (95% CI)** | **SD** | **N (missing)** | **Mean (95% CI)** | **SD** | **N in model (Int, Con)** | **Effect Size (95% CI)** |
| SDQ | | | | | | | | |
| **Prosocial Score** | 40 (7) | 6.78 (5.78, 7.77) | 3.13 | 45 (4) | 6.78 (6.03, 7.52) | 2.50 | 74 (33, 41) | -0.00 (-0.43, 0.42) |
| **Total Difficulties Score** | 40 (7) | 9.67 (7.29, 12.06) | 7.54 | 45 (4) | 9.64 (7.78, 11.51) | 6.26 | 74 (33, 41) | 0.15 (-0.28, 0.58) |
| **Externalising Score** | 40 (7) | 5.73 (4.13, 7.32) | 5.05 | 45 (4) | 5.96 (4.69, 7.22) | 4.24 | 74 (33, 41) | 0.10 (-0.33, 0.53) |
| **Internalising Score** | 40 (7) | 3.95 (2.65, 5.25) | 4.10 | 45 (4) | 3.69 (2.76, 4.62) | 3.13 | 74 (33, 41) | 0.15 (-0.27, 0.58) |
| **Impact Score** | 28 (4) | 1.75 (0.94, 2.56) | 2.14 | 25 (3) | 1.64 (0.99, 2.29) | 1.63 | 46 (24, 22) | 0.20 (-0.35, 0.74) |
| PARYC | | | | | | | | |
| **Supporting Frequency** | 26 (3) | 42.62 (40.28, 44.95) | 5.96 | 22 (4) | 41.09 (38.08, 44.10) | 7.05 | 41 (23, 18) | 0.09 (-0.48, 0.66) |
| **Supporting Problematic** | 26 (3) | 0.77 (0.28, 1.26) | 1.24 | 22 (4) | 1.36 (0.57, 2.16) | 1.87 | 41 (23, 18) | -0.16 (-0.72, 0.42) |
| **Proactive Frequency** | 26 (3) | 41.31 (38.26, 44.36) | 7.78 | 22 (4) | 40.14 (37.98, 42.29) | 5.05 | 41 (23, 18) | 0.16 (-0.41, 0.73) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Proactive Problematic** | 26 (3) | 0.54 (0.03, 1.05) | 1.30 | 22 (4) | 1.14 (0.41, 1.86) | 1.70 | 41 (23, 18) | -0.35 (-0.93, 0.22) |
| **Total Frequency** | 26 (3) | 83.92 (78.95, 88.89) | 12.67 | 22 (4) | 81.23 (76.46, 85.99) | 11.17 | 41 (23, 18) | 0.14 (-0.43, 0.71) |
| **Total Problematic** | 26 (3) | 1.31 (0.39, 2.23) | 2.35 | 22 (4) | 2.50 (1.10, 3.90) | 3.28 | 41 (23, 18) | -0.27 (-0.84, 0.30) |

Table D.7: Year 1 Phonics results (all eligible boys)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **Year 1 Phonics** | 149 (13) | 29.87 (28.24, 31.49) | 9.93 | 170 (8) | 29.21 (27.64, 30.77) | 10.19 | 298 (136, 162) | 0.12 (-0.11, 0.34) |

Table D.8: KS1 Assessment results (all eligible boys)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **KS1 Assessment Reading/Writing Score** | 62 (10) | 13.60 (12.69, 14.50) | 3.55 | 64 (5) | 12.48 (11.67, 13.30) | 3.25 | 111 (52, 59) | 0.28 (-0.07, 0.63) |
| **KS1 Assessment Reading Score** | 62 (10) | 14.65 (13.68, 15.61) | 3.81 | 64 (5) | 13.00 (12.11, 13.89) | 3.55 | 111 (52, 59) | 0.41 (0.05, 0.76)* |

* p≤0.05

Table D.9: Primary Outcomes at post-test (all eligible girls)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **Letter ID** | 162 (17) | 108.81 (107.52, 110.09) | 8.18 | 136 (14) | 108.06 (106.59, 109.53) | 8.56 | 267 (145, 122) | 0.10 (-0.13, 0.33) |
| **Word ID** | 156 (15) | 121.78 (119.54, 124.01) | 13.93 | 129 (14) | 121.08 (118.74, 123.41) | 13.25 | 256 (141, 115) | 0.01 (-0.22, 0.25) |
| **Word Attack** | 158 (16) | 119.03 (117.60, 120.45) | 8.93 | 131 (14) | 118.89 (117.24, 120.55) | 9.47 | 259 (142, 117) | -0.01 (-0.24, 0.22) |

Table D.10: Secondary Outcomes – social and behavioural measures at post-test follow-up (all eligible girls)

| Outcome | Raw scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| SDQ | | | | | | | | |
| **Prosocial Score** | 132 (15) | 7.55 (7.13, 7.96) | 2.37 | 121 (11) | 7.76 (7.36, 8.16) | 2.20 | 227 (117, 110) | -0.05 (-0.30, 0.19) |
| **Total Difficulties Score** | 132 (15) | 6.25 (5.32, 7.18) | 5.34 | 121 (11) | 6.51 (5.52, 7.51) | 5.47 | 227 (117, 110) | -0.07 (-0.32, 0.17) |
| **Externalising Score** | 132 (15) | 3.67 (3.04, 4.29) | 3.60 | 121 (11) | 3.32 (2.67, 3.97) | 3.56 | 227 (117, 110) | 0.06 (-0.19, 0.31) |
| **Internalising Score** | 132 (15) | 2.58 (2.07, 3.10) | 2.96 | 121 (11) | 3.19 (2.62, 3.76) | 3.13 | 227 (117, 110) | -0.20 (-0.45, 0.05) |
| **Impact Score** | 47 (9) | 0.94 (0.60, 1.27) | 1.15 | 41 (3) | 0.73 (0.39, 1.07) | 1.10 | 76 (38, 38) | 0.24 (-0.19, 0.66) |
| PARYC | | | | | | | | |
| **Supporting Frequency** | 81 (7) | 41.06 (39.98, 42.15) | 4.89 | 79 (9) | 41.57 (40.13, 43.00) | 6.38 | 144 (74, 70) | -0.03 (-0.34, 0.28) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Supporting Problematic** | 81 (7) | 1.05 (0.78, 1.32) | 1.20 | 79 (9) | 1.01 (0.75, 1.28) | 1.17 | 144 (74, 70) | 0.03 (-0.28, 0.34) |
| **Proactive Frequency** | 82 (7) | 38.63 (37.21, 40.05) | 6.43 | 80 (9) | 39.45 (37.61, 41.29) | 8.22 | 146 (75, 71) | -0.06 (-0.37, 0.25) |
| **Proactive Problematic** | 82 (7) | 1.01 (0.66, 1.36) | 1.58 | 80 (9) | 0.89 (0.57, 1.20) | 1.41 | 146 (75, 71) | 0.07 (-0.24, 0.38) |
| **Total Frequency** | 80 (7) | 79.79 (77.73, 81.84) | 9.20 | 79 (9) | 81.15 (78.15, 84.15) | 13.32 | 143 (73, 70) | -0.05 (-0.36, 0.26) |
| **Total Problematic** | 80 (7) | 2.09 (1.58, 2.60) | 2.29 | 79 (9) | 1.91 (1.42, 2.40) | 2.17 | 143 (73, 70) | 0.05 (-0.26, 0.36) |

Table D.11: Primary Outcomes at 6-month follow-up (all eligible girls)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **Letter ID** | 110 (11) | 104.72 (103.54, 105.90) | 6.19 | 110 (12) | 103.98 (102.47, 105.49) | 7.91 | 197 (99, 98) | 0.10 (-0.17, 0.36) |
| **Word ID** | 98 (11) | 119.45 (116.99, 121.91) | 12.16 | 102 (13) | 119.72 (116.99, 122.44) | 13.77 | 176 (87, 89) | -0.07 (-0.34, 0.21) |
| **Word Attack** | 108 (10) | 119.03 (117.35, 120.71) | 8.73 | 110 (13) | 118.55 (116.64, 120.47) | 10.05 | 195 (98, 97) | -0.05 (-0.32, 0.22) |

Table D.12: Secondary Outcomes – social and behavioural measures at 6-month follow-up (all eligible girls)

| Outcome | Raw scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| SDQ | | | | | | | | |
| **Prosocial Score** | 84 (11) | 8.15 (7.64, 8.67) | 2.36 | 92 (12) | 7.70 (7.19, 8.20) | 2.43 | 153 (73, 80) | 0.30 (-0.00, 0.59) |
| **Total Difficulties Score** | 84 (11) | 7.10 (5.61, 8.58) | 6.80 | 92 (12) | 6.55 (5.34, 7.77) | 5.84 | 153 (73, 80) | 0.09 (-0.21, 0.38) |
| **Externalising Score** | 84 (11) | 3.85 (2.94, 4.75) | 4.14 | 92 (12) | 3.50 (2.73, 4.27) | 3.71 | 153 (73, 80) | 0.06 (-0.24, 0.35) |
| **Internalising Score** | 84 (11) | 3.25 (2.49, 4.01) | 3.46 | 92 (12) | 3.05 (2.34, 3.77) | 3.42 | 153 (73, 80) | 0.09 (-0.20, 0.39) |
| **Impact Score** | 37 (6) | 0.86 (0.52, 1.21) | 1.06 | 38 (6) | 1.34 (0.91, 1.77) | 1.32 | 63 (31, 32) | -0.29 (-0.74, 0.17) |
| PARYC | | | | | | | | |
| **Supporting Frequency** | 49 (6) | 42.43 (40.91, 43.95) | 5.32 | 58 (6) | 41.31 (39.72, 42.90) | 6.06 | 95 (43, 52) | 0.17 (-0.21, 0.55) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Supporting Problematic** | 49 (6) | 0.94 (0.60, 1.28) | 1.18 | 58 (6) | 0.72 (0.47, 0.98) | 0.97 | 95 (43, 52) | 0.13 (-0.26, 0.51) |
| **Proactive Frequency** | 48 (5) | 39.42 (37.69, 41.14) | 5.97 | 58 (6) | 39.88 (38.10, 41.66) | 6.77 | 95 (43, 52) | -0.10 (-0.48, 0.28) |
| **Proactive Problematic** | 48 (5) | 0.88 (0.52, 1.23) | 1.21 | 58 (6) | 0.71 (0.37, 1.05) | 1.30 | 95 (43, 52) | 0.13 (-0.26, 0.51) |
| **Total Frequency** | 47 (5) | 81.83 (79.10, 84.56) | 9.35 | 58 (6) | 81.19 (78.13, 84.25) | 11.65 | 94 (42, 52) | 0.03 (-0.35, 0.42) |
| **Total Problematic** | 47 (5) | 1.74 (1.25, 2.24) | 1.70 | 58 (6) | 1.43 (0.94, 1.92) | 1.86 | 94 (42, 52) | 0.18 (-0.21, 0.56) |

Table D.13: Primary Outcomes at 12-month follow-up (all eligible girls)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **Letter ID** | 35 (7) | 97.83 (95.27, 100.38) | 7.56 | 36 (7) | 97.81 (94.60, 101.01) | 9.61 | 57 (28, 29) | 0.08 (-0.39, 0.55) |
| **Word ID** | 28 (7) | 111.68 (106.04, 117.32) | 14.92 | 28 (6) | 112.29 (106.46, 118.11) | 15.42 | 43 (21, 22) | 0.24 (-0.29, 0.77) |
| **Word Attack** | 34 (7) | 113.47 (109.05, 117.89) | 12.90 | 36 (7) | 116.06 (111.35, 120.76) | 14.11 | 56 (27, 29) | -0.15 (-0.62, 0.32) |

Table D.14: Secondary Outcomes – social and behavioural measures at 12-month follow-up (all eligible girls)

| Outcome | Raw scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| SDQ | | | | | | | | |
| **Prosocial Score** | 31 (7) | 8.29 (7.52, 9.06) | 2.15 | 28 (3) | 7.79 (7.01, 8.56) | 2.06 | 49 (24, 25) | 0.41 (-0.11, 0.93) |
| **Total Difficulties Score** | 31 (7) | 7.06 (4.57, 9.56) | 6.95 | 28 (3) | 6.21 (4.41, 8.02) | 4.77 | 49 (24, 25) | -0.04(-0.55, 0.47) |
| **Externalising Score** | 31 (7) | 4.29 (2.55, 6.03) | 4.84 | 28 (3) | 3.68 (2.26, 5.10) | 3.75 | 49 (24, 25) | 0.03 (-0.48, 0.54) |
| **Internalising Score** | 31 (7) | 2.77 (1.84, 3.71) | 2.60 | 28 (3) | 2.54 (1.52, 3.55) | 2.69 | 49 (24, 25) | -0.14 (-0.65, 0.37) |
| **Impact Score** | 14 (4) | 1.36 (0.56, 2.16) | 1.50 | 9 (1) | 0.44 (-0.44, 1.33) | 1.33 | 18 (10, 8) | 0.46 (-0.39, 1.30) |
| PARYC | | | | | | | | |
| **Supporting Frequency** | 12 (4) | 43.67 (40.45, 46.88) | 5.57 | 11 (1) | 40.91 (37.96, 43.86) | 4.89 | 18 (8, 10) | 0.34 (-0.49, 1.16) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Supporting Problematic** | 12 (4) | 1.17 (0.28, 2.05) | 1.53 | 11 (1) | 1.27 (0.73, 1.82) | 0.90 | 18 (8, 10) | -0.24 (-1.06, 0.58) |
| **Proactive Frequency** | 12 (4) | 40.83 (37.67, 43.99) | 5.47 | 11 (1) | 38.27 (33.76, 42.79) | 7.48 | 18 (8, 10) | 0.20 (-0.62, 1.02) |
| **Proactive Problematic** | 12 (4) | 1.08 (-0.28, 2.44) | 2.35 | 11 (1) | 1.00 (0.24, 1.76) | 1.26 | 18 (8, 10) | 0.00 (-0.82, 0.82) |
| **Total Frequency** | 12 (4) | 84.50 (79.49, 89.51) | 8.68 | 11 (1) | 79.18 (72.17, 86.20) | 11.63 | 18 (8, 10) | 0.30 (-0.52, 1.13) |
| **Total Problematic** | 12 (4) | 2.25 (0.12, 4.38) | 3.70 | 11 (1) | 2.27 (1.09, 3.45) | 1.95 | 18 (8, 10) | -0.10 (-0.92, 0.72) |

Table D.15: Year 1 Phonics results (all eligible girls)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **Year 1 Phonics** | 136 (20) | 30.18 (28.45, 31.92) | 10.10 | 124 (14) | 30.91 (29.01, 32.81) | 10.58 | 226 (116, 110) | -0.11 (-0.35, 0.13) |

Table D.16: KS1 Assessment results (all eligible girls)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **KS1 Assessment Reading/Writing Score** | 52 (12) | 14.33 (13.26, 15.39) | 3.85 | 47 (9) | 13.89 (12.79, 15.00) | 3.78 | 78 (40, 38) | 0.21 (-0.19, 0.60) |
| **KS1 Assessment Reading Score** | 52 (12) | 14.69 (13.52, 15.87) | 4.24 | 47 (9) | 14.36 (13.24, 15.48) | 3.85 | 78 (40, 38) | 0.19 (-0.21, 0.58) |

# Appendix E: Outcomes for children with low baseline scores (95% of mean or less)

Table E.1: Primary Outcomes at pre-test (all eligible pupils with 95% of mean or less at baseline)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **Letter ID** | 83 (1) | 104.69 (102.86, 106.51) | 8.31 | 107 (2) | 105.15 (103.54, 106.76) | 8.31 | 187 (82, 105) | -0.07 (-0.35, 0.22) |
| **Word ID** | 82 (1) | 114.83 (111.80, 117.85) | 13.70 | 107 (2) | 115.96 (113.72, 118.20) | 11.57 | 186 (81, 105) | -0.10 (-0.39, 0.19) |
| **Word Attack** | 81 (1) | 115.43 (113.39, 117.48) | 9.21 | 106 (2) | 115.36 (113.53, 117.18) | 9.39 | 184 (80, 104) | 0.02 (-0.27, 0.31) |

Table E.2: Secondary Outcomes – social and behavioural measures at post-test (all eligible pupils with 95% of mean or less at baseline)

| Outcome | Raw scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| SDQ | | | | | | | | |
| **Prosocial Score** | 69 (0) | 6.42 (5.76, 7.08) | 2.75 | 96 (1) | 6.77 (6.21, 7.33) | 2.74 | 164 (69, 95) | -0.17 (-0.45, 0.14) |
| **Total Difficulties Score** | 69 (0) | 9.32 (7.86, 10.78) | 6.07 | 96 (1) | 8.97 (7.74, 10.19) | 6.01 | 164 (69, 95) | 0.10 (-0.21, 0.41) |
| **Externalising Score** | 69 (0) | 5.68 (4.66, 6.70) | 4.22 | 96 (1) | 5.49 (4.61, 6.37) | 4.33 | 164 (69, 95) | 0.12 (-0.19, 0.43) |
| **Internalising Score** | 69 (0) | 3.64 (2.89, 4.39) | 3.11 | 96 (1) | 3.48 (2.82, 4.14) | 3.22 | 164 (69, 95) | 0.03 (-0.28, 0.34) |
| **Impact Score** | 39 (0) | 1.44 (0.93, 1.94) | 1.59 | 50 (0) | 1.50 (1.08, 1.92) | 1.50 | 89 (39, 50) | 0.14 (-0.28, 0.56) |
| PARYC | | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Supporting Frequency** | 41 (0) | 40.85 (39.15, 42.56) | 5.46 | 62 (2) | 41.71 (40.18, 43.24) | 6.01 | 101 (41, 60) | -0.17 (-1.07, 0.73) |
| **Supporting Problematic** | 41 (0) | 1.24 (0.83, 1.66) | 1.32 | 62 (2) | 1.32 (0.96, 1.68) | 1.41 | 101 (41, 60) | 0.16 (-0.73, 1.06) |
| **Proactive Frequency** | 40 (0) | 38.23 (35.95, 40.50) | 7.18 | 62 (2) | 38.69 (36.72, 40.67) | 7.79 | 100 (40, 60) | 0.14 (-0.76, 1.04) |
| **Proactive Problematic** | 40 (0) | 1.38 (0.85, 1.90) | 1.66 | 62 (2) | 1.26 (0.82, 1.69) | 1.72 | 100 (40, 60) | 0.23 (-0.67, 1.12) |
| **Total Frequency** | 40 (0) | 78.88 (75.49, 82.26) | 10.70 | 62 (2) | 80.40 (77.24, 83.56) | 12.44 | 100 (40, 60) | -0.00 (-0.90, 0.89) |
| **Total Problematic** | 40 (0) | 2.65 (1.93, 3.37) | 2.27 | 62 (2) | 2.58 (1.88, 3.28) | 2.76 | 100 (40, 60) | 0.22 (-0.68, 1.12) |

Table E.3: Primary Outcomes at 6-month follow-up (all eligible pupils with 95% of mean or less at baseline)

| **Outcome** | **Standardised scores** | | | | | | **Effect Size** | |
|---|---|---|---|---|---|---|---|---|
| | **Intervention** | | | **Control** | | | | |
| | **N (missing)** | **Mean (95% CI)** | **SD** | **N (missing)** | **Mean (95% CI)** | **SD** | **N in model (Int, Con)** | **Effect Size (95% CI)** |
| **Letter ID** | 62 (0) | 101.52 (99.46, 103.58) | 8.11 | 89 (2) | 101.45 (99.69, 103.21) | 8.31 | 149 (62, 87) | 0.03 (-0.30, 0.35) |
| **Word ID** | 55 (0) | 114.36 (110.43, 118.29) | 14.58 | 79 (2) | 113.67 (110.61, 116.73) | 13.61 | 132 (55, 77) | 0.03 (-0.32, 0.37) |
| **Word Attack** | 63 (0) | 115.76 (112.71, 118.82) | 12.13 | 89 (2) | 115.10 (112.85, 117.35) | 10.63 | 150 (63, 87) | 0.00 (-0.32, 0.32) |

Table E.4: Secondary Outcomes – social and behavioural measures at 6-month follow-up (all eligible pupils with 95% of mean or less at baseline)

| **Outcome** | **Raw scores** | | | | | | **Effect Size** | |
|---|---|---|---|---|---|---|---|---|
| | **Intervention** | | | **Control** | | | | |
| | **N (missing)** | **Mean (95% CI)** | **SD** | **N (missing)** | **Mean (95% CI)** | **SD** | **N in model (Int, Con)** | **Effect Size (95% CI)** |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| SDQ | | | | | | | | |
| **Prosocial Score** | 52 (0) | 7.06 (6.27, 7.84) | 2.84 | 76 (2) | 7.11 (6.47, 7.74) | 2.75 | 126 (52, 74) | 0.04 (-0.32, 0.39) |
| **Total Difficulties Score** | 52 (0) | 10.06 (7.93, 12.19) | 7.67 | 76 (2) | 8.46 (7.03, 9.89) | 6.24 | 126 (52, 74) | 0.22 (-0.13, 0.58) |
| **Externalising Score** | 52 (0) | 5.85 (4.45, 7.24) | 5.03 | 76 (2) | 5.01 (4.02, 6.01) | 4.34 | 126 (52, 74) | 0.16 (-0.19, 0.52) |
| **Internalising Score** | 52 (0) | 4.21 (3.15, 5.28) | 3.84 | 76 (2) | 3.45 (2.72, 4.18) | 3.18 | 126 (52, 74) | 0.22 (-0.13, 0.58) |
| **Impact Score** | 33 (0) | 1.73 (1.14, 2.32) | 1.70 | 48 (2) | 1.35 (0.96, 1.75) | 1.38 | 79 (33, 46) | 0.29 (-0.15, 0.74) |
| PARYC | | | | | | | | |
| **Supporting Frequency** | 27 (0) | 43.74 (41.74, 45.74) | 5.20 | 45 (2) | 41.40 (39.81, 42.99) | 5.33 | 70 (27, 43) | 0.59 (0.10, 1.08)* |
| **Supporting Problematic** | 27 (0) | 0.89 (0.40, 1.38) | 1.28 | 45 (2) | 1.33 (0.84, 1.83) | 1.67 | 70 (27, 43) | -0.20 (-0.68, 0.27) |
| **Proactive Frequency** | 24 (0) | 40.67 (38.10, 43.23) | 6.28 | 47 (2) | 39.87 (38.04, 41.71) | 6.29 | 69 (24, 45) | 0.20 (-0.29, 0.70) |
| **Proactive Problematic** | 24 (0) | 0.71 (0.23, 1.18) | 1.16 | 47 (2) | 1.23 (0.69, 1.78) | 1.87 | 69 (24, 45) | -0.11 (-0.60, 0.38) |
| **Total Frequency** | 24 (0) | 84.17 (79.93, 88.41) | 10.38 | 45 (2) | 81.36 (78.13, 84.58) | 10.83 | 67 (24, 43) | 0.40 (-0.10, 0.90) |
| **Total Problematic** | 24 (0) | 1.67 (0.76, 2.57) | 2.22 | 45 (2) | 2.60 (1.63, 3.57) | 3.26 | 67 (24, 43) | -0.17 (-0.66, 0.33) |

*** p≤0.05**

Table E.5: Primary Outcomes at 12-month follow-up (all eligible pupils with 95% of mean or less at baseline)

| **Outcome** | **Standardised scores** | | | | | | **Effect Size** | |
|---|---|---|---|---|---|---|---|---|
| | **Intervention** | | | **Control** | | | | |
| | **N (missing)** | **Mean (95% CI)** | **SD** | **N (missing)** | **Mean (95% CI)** | **SD** | **N in model (Int, Con)** | **Effect Size (95% CI)** |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Letter ID** | 22 (0) | 95.59 (92.30, 98.89) | 7.73 | 33 (0) | 93.61 (90.23, 96.98) | 9.69 | 55 (22, 33) | 0.16 (-0.38, 0.70) |
| **Word ID** | 14 (0) | 106.93 (100.25, 113.60) | 12.49 | 22 (0) | 101.82 (94.81, 108.83) | 16.45 | 36 (14, 22) | 0.35 (-0.33, 1.02) |
| **Word Attack** | 22 (0) | 110.32 (104.88, 115.76) | 12.75 | 31 (0) | 105.68 (100.46, 110.89) | 14.52 | 53 (22, 31) | 0.31 (-0.24, 0.86) |

Table E.6: Secondary Outcomes – social and behavioural measures at 12-month follow-up (all eligible pupils with 95% of mean or less at baseline)

| **Outcome** | **Raw scores** | | | | | | **Effect Size** | |
|---|---|---|---|---|---|---|---|---|
| | **Intervention** | | | **Control** | | | | |
| | **N (missing)** | **Mean (95% CI)** | **SD** | **N (missing)** | **Mean (95% CI)** | **SD** | **N in model (Int, Con)** | **Effect Size (95% CI)** |
| SDQ | | | | | | | | |
| **Prosocial Score** | 21 (0) | 7.29 (5.96, 8.61) | 3.04 | 24 (0) | 6.88 (5.74, 8.01) | 2.77 | 45 (21, 24) | 0.09 (-0.49, 0.68) |
| **Total Difficulties Score** | 21 (0) | 7.10 (4.39, 9.80) | 6.20 | 24 (0) | 8.50 (5.97, 11.03) | 6.20 | 45 (21, 24) | 0.02 (-0.57, 0.60) |
| **Externalising Score** | 21 (0) | 4.33 (2.71, 5.96) | 3.72 | 24 (0) | 5.21 (3.54, 6.88) | 4.09 | 45 (21, 24) | 0.14 (-0.44, 0.73) |
| **Internalising Score** | 21 (0) | 2.76 (1.07, 4.45) | 3.87 | 24 (0) | 3.29 (1.97, 4.61) | 3.24 | 45 (21, 24) | -0.13 (-0.72, 0.46) |
| **Impact Score** | 13 (0) | 1.08 (0.28, 1.88) | 1.44 | 10 (0) | 1.10 (-0.11, 2.31) | 1.91 | 23 (13, 10) | -0.45 (-1.28, 0.39) |
| PARYC | | | | | | | | |
| **Supporting Frequency** | 6 (0) | 41.50 (34.48, 48.52) | 8.60 | 13 (0) | 39.92 (35.17, 44.68) | 8.57 | 19 (6, 13) | 0.11 (-0.86, 1.08) |
| **Supporting Problematic** | 6 (0) | 1.00 (-0.37, 2.37) | 1.67 | 13 (0) | 1.54 (0.49, 2.59) | 1.90 | 19 (6, 13) | -0.38 (-1.36, 0.60) |
| **Proactive Frequency** | 6 (0) | 40.50 (32.42, 48.58) | 9.89 | 13 (0) | 37.69 (33.30, 42.08) | 7.91 | 19 (6, 13) | 0.10 (-0.87, 1.07) |
| **Proactive Problematic** | 6 (0) | 1.17 (-0.79, 3.13) | 2.40 | 13 (0) | 1.62 (0.45, 2.78) | 2.10 | 19 (6, 13) | -0.59 (-1.57, 0.40) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Total Frequency** | 6 (0) | 82.00 (68.06, 95.94) | 17.08 | 13 (0) | 77.62 (68.99, 86.24) | 15.55 | 19 (6, 13) | 0.11 (-0.86, 1.08) |
| **Total Problematic** | 6 (0) | 2.17 (-1.12, 5.45) | 4.02 | 13 (0) | 3.15 (1.04, 5.26) | 3.80 | 19 (6, 13) | -0.51 (-1.49, 0.47) |

# Appendix F: Outcomes according to programme 'completion'

Table F.1: Primary Outcomes at post-test (all eligible 'completers' and control pupils)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **Letter ID** | 193 (16) | 108.37 (107.24, 109.50) | 7.85 | 356 (27) | 107.24 (106.38, 108.09) | 8.07 | 506 (177, 329) | 0.10 (-0.08, 0.28) |
| **Word ID** | 188 (14) | 122.16 (120.34, 123.99) | 12.52 | 344 (26) | 119.93 (118.54, 121.32) | 12.89 | 492 (174, 318) | 0.06 (-0.12, 0.24) |
| **Word Attack** | 187 (15) | 119.52 (118.24, 120.81) | 8.81 | 347 (26) | 118.36 (117.41, 119.32) | 8.88 | 493 (172, 321) | 0.04 (-0.13, 0.22) |

Table F.2: Primary Outcomes at 6-month follow-up (all eligible 'completers' and control pupils)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **Letter ID** | 152 (14) | 104.23 (103.18, 105.28) | 6.46 | 271 (21) | 103.07 (102.11, 104.03) | 7.90 | 388 (138, 250) | 0.12 (-0.08, 0.32) |
| **Word ID** | 137 (14) | 121.72 (119.47, 123.96) | 13.12 | 248 (22) | 118.59 (116.81, 120.38) | 14.05 | 349 (123, 226) | 0.14 (-0.07, 0.35) |
| **Word Attack** | 151 (14) | 119.50 (117.92, 121.07) | 9.66 | 270 (22) | 117.49 (116.29, 118.70) | 9.90 | 385 (137, 248) | 0.15 (-0.05, 0.35) |

Table F.3: Primary Outcomes at 12-month follow-up (all eligible 'completers' and control pupils)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Letter ID** | 50 (9) | 97.72 (95.64, 99.80) | 7.37 | 91 (12) | 95.96 (94.09, 97.83) | 8.92 | 120 (41, 79) | 0.13 (-0.22, 0.47) |
| **Word ID** | 39 (8) | 111.72 (106.91, 116.52) | 15.00 | 69 (11) | 108.45 (104.71, 112.19) | 15.54 | 89 (31, 58) | 0.13 (-0.27, 0.52) |
| **Word Attack** | 49 (9) | 115.37 (111.72, 119.01) | 12.76 | 89 (12) | 111.42 (108.25, 114.58) | 14.93 | 117 (40, 77) | 0.24 (-0.11, 0.59) |

Table F.4: Year 1 Phonics results (all eligible 'completers' and control pupils)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **Year 1 Phonics** | 164 (15) | 31.03 (29.54, 32.52) | 9.55 | 294 (22) | 29.93 (28.72, 31.14) | 10.37 | 421 (149, 272) | 0.04 (-0.15, 0.23) |

Table F.5: KS1 assessment results (all eligible 'completers' and control pupils)

| Outcome | Standardised scores | | | | | | Effect Size | |
|---|---|---|---|---|---|---|---|---|
| | Intervention | | | Control | | | | |
| | N (missing) | Mean (95% CI) | SD | N (missing) | Mean (95% CI) | SD | N in model (Int, Con) | Effect Size (95% CI) |
| **KS1 Assessment Reading/Writing Score** | 64 (11) | 14.55 (13.69, 15.40) | 3.41 | 111 (14) | 13.08 (12.41, 13.75) | 3.54 | 150 (53, 97) | 0.35* (0.04, 0.66) |
| **KS1 Assessment Reading Score** | 64 (11) | 15.38 (14.45, 16.30) | 3.70 | 111 (14) | 13.58 (12.87, 14.28) | 3.72 | 150 (53, 97) | 0.42* (0.11, 0.73) |

*p≤0.05

SPOKES

# Appendix G: End of Programme evaluations (Cohorts 3-6)

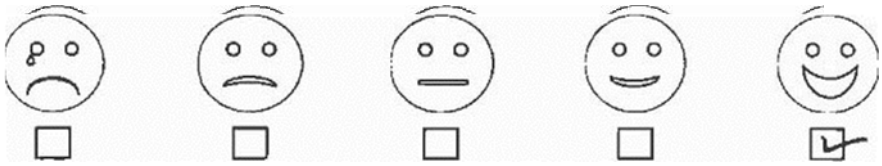
*We would like to know what you thought of the SPOKES Literacy Programme. The following questionnaire is part of our evaluation of SPOKES and your views will help us to improve the programme we offer. Please answer as honestly as possible. Your co-operation is greatly appreciated and all responses are strictly confidential.*

**Name:** **Date:** _____

## A: Qverall SPOKES programme

*Please indicate* **how enjoyable** *you found the overall SPOKES programme. Please tick the box below the smiley face that best represents how you feel.*

**1. Overall, how *enjoyable* did you find the SPOKES programme?**



| | | | | |
|---|---|---|---|---|
| 0 (0%) | 0 (0%) | 1 (1%) | 20 (%16) | 101 (83%) |

## B: SPOKES techniques/content

*Please indicate* **how *usefulyou* *found*** *the specific techniques. Please circle the response that best describes what you think* **(**1= not at all useful, 5= very useful)

**1. Talking about the book before reading with your child**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 2 (2%) | 12 (10%) | 32 (26%) | 78 (63%) | |

**2. Pausing and prompting if your child makes a mistake when reading**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 0 (0%) | 3 (2%) | 14 (11%) | 107 (86%) | |

**3. Praising your child for getting a word right or for 'having a go' at a new word**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 1 (1%) | 0 (0%) | 1 (1%) | 9 (7%) | | |

**4. Using Environmental Print**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 1 (1%) | 3 (3%) | 24 (21%) | 40 (34%) | 34 (91%) | |

**5. Encouraging writing at home**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 0 (0%) | 3 (2%) | 30 (25%) | 88 (73%) | |

**6. Using letters and sounds to read**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 0 (0%) | 10 (8%) | 30 (24%) | 83 (67%) | |

**7. Setting aside a specific time for reading**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 1 (1%) | 8 (7%) | 23 (19%) | 91 (74%) | |

**8. Playing with your child**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 1 (1%) | 5 (4%) | 22 (18%) | 95 (77%) | |

**9. Visiting the library**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 2 (2%) | 2 (2%) | 16 (15%) | 32 (30%) | 56 (52%) | |

**10. One-to-one visit**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 2 (3%) | 1 (1%) | 6 (8%) | 19 (25%) | 47 (63%) | |

**Any other comments about the SPOKES techniques:**

## C: Group sessions

*Please indicate* **how useful** *you found each of the following methods. Please circle the response that most clearly describes your opinion.* (1= not at all useful, 5= very useful)

**1. Information provided by the group leader**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 0 (0%) | 3 (2%) | 12 (10%) | 108 (88%) | |

**2. Demonstration of skills through videos**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 1 (1%) | 18 (15%) | 39 (32%) | 64 (52%) | |

### 3. Group discussion of skills

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 0 (0%) | 0 (0%) | 22 (18%) | 100 (82%) | |

**4. Role plays and practising during the group sessions**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 0 (%) | 2 (2%) | 11 (9%) | 30 (24%) | 80 (50%) | |

**5. Practising the reading and other literacy skills with your child at home**

| *1=not at all useful* | **1** | **2** | **3** | **4** | **5** | *5=very useful* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 0 (0%) | 1 (%) | 12 (%) | 110 (%) | |

**Any other comments about the group sessions:**

D: Group leaders

*Please tell us your views about* your group leaders. *Please circle the response that best describes how you feel.* (1 = low, 5 = high)

**1. The overall leading of the group**

| *1=unhelpful* | **1** | **2** | **3** | **4** | **5** | *5=very helpful* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 0 (0%) | 1 (1%) | 8 (7%) | 113 (93%) | |

**2. The leaders' preparation**

| *1=unhelpful* | **1** | **2** | **3** | **4** | **5** | *5=very helpful* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 0 (0%) | 1 (1%) | 10 (8%) | 112 (92%) | |

**3. Regarding the leaders' interest in me and my child, I was….**

| *1=not satisfied* | 1 | 2 | 3 | 4 | 5 | *5=very satisfied* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 0 (0%) | 1 (1%) | 10 (8%) | 112 (92%) | |

**Any other comments about the group leaders:**

## E: My child

*Please tell us about* **your child's reading.** *Please circle the response that best describes how you feel.* (1= got much worse, 5= got much better)

**1. Since beginning the literacy programme my child's reading has……**

| *1=unhelpful* | 1 | 2 | 3 | 4 | 5 | *5=very helpful* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 0 (0%) | 6 (5%) | 48 (40%) | 68 (57%) | |

**2. Since the beginning of the programme my ability to help my child read has…**

| *1=got much worse* | 1 | 2 | 3 | 4 | 5 | *5=got much better* |
|---|---|---|---|---|---|---|
| | 0 (0%) | 0 (0%) | 2 (2%) | 28 (23%) | 93 (76%) | |

**Any other comments about your child's reading:**

## F: General

*Finally, we would be grateful if you would answer a few* **general questions** *about SPOKES:*

1. **Which part of the programme was mast helpful to you?**

2. **Which part of the programme was least helpful to you?**

3. **What was most enjoyable?**

4. **What did you like least about the programme?**

5. **How could the programme have been improved to help you more?**

6. **What would you say to a new parent thinking of participating in SPOKES?**

7. **Do you have any other comments?**

***Thank you for completing this form!***
***The SPOKES team***

## Appendix H: Weekly Evaluation forms

### *SPOKES Literacy Programme*

*Your comments on today's session*

**Please rate below your views on today's session** *(1 = low, 5 = high)*

*Did you find today's session enjoyable?*

*1* *2* *3* 4 5

*Overall how useful was today's topic to you?*

*1* *2* *3* 4 5

*What did you like most about today's session?*

*What did you like least about today's session?*

*1 would like more information in future sessions on……..…*

**Comments**

**Name:**

**Date:**

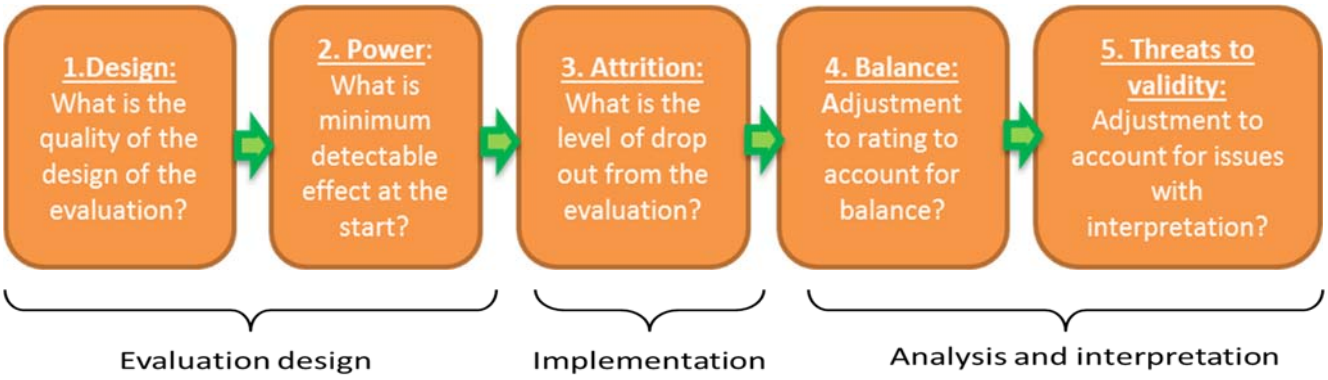
# Appendix F: Cost rating

Cost ratings are based on the approximate cost per pupil per year of implementing the intervention over three years. Cost ratings are awarded using the following criteria.

| Cost | Description |
|---|---|
| £ | *Very low:* less than £80 per pupil per year. |
| £ £ | *Low:* up to about £200 per pupil per year. |
| £ £ £ | *Moderate:* up to about £700 per pupil per year. |
| £ £ £ £ | *High:* up to £1,200 per pupil per year. |
| £ £ £ £ £ | *Very high:* over £1,200 per pupil per year. |

# Appendix G: Padlock rating

**7th July 2016 Complete by Elena Rosa Brown**




| Rating | 1. Design | 2. Power (MDES) | 3. Attrition | 4. Balance | 5. Threats to validity |
|---|---|---|---|---|---|
| 5 🔒 | Fair and clear experimental design (RCT) | < 0.2 | < 10% | Well-balanced on observables | No threats to validity |
| 4 🔒 | Fair and clear experimental design (RCT, RDD) | < 0.3 | < 20% | | |
| 3 🔒 | Well-matched comparison (quasi-experiment) | < 0.4 | < 30% | | |
| 2 🔒 | Matched comparison (quasi-experiment) | < 0.5 | < 40% | | |
| 1 🔒 | Comparison group with poor or no matching | < 0.6 | < 50% | | |
| 0 🔒 | No comparator | > 0.6 | > 50% | Imbalanced on observables | Significant threats |

The final security rating for this trial is 3 🔒.

This document is available for download at 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