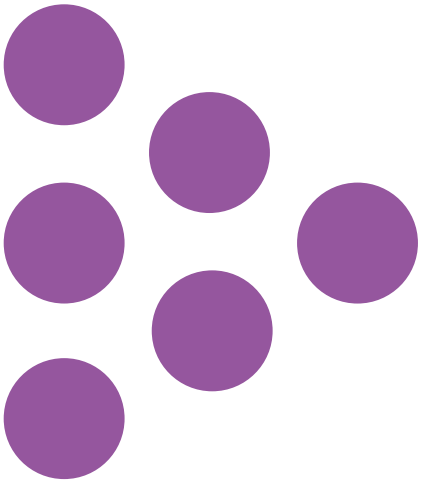
## Randomised Controlled Trial (RCT) of Families Connect

# Statistical Analysis Plan

**Public: 07/01/2020**



# Contents

# 1 Evaluation Summary

| RCT evaluation of The Families Connect Programme | |
|---|---|
| Developer | Save the Children UK (SCUK) |
| Evaluator | National Foundation for Educational Research (NFER), with Queen's University Belfast supporting the process evaluation |
| Trial Registration Number | ISRCTN88158874 |
| Principal investigator | Ben Styles |
| Trial manager | Pippa Lord |
| Trial statistician | Connie Rennie |
| SAP version / date | V3 25.10.19 |
| **Trial Design** | |
| Age range | 4 – 6 year olds (Reception and Year 1 in England and Wales; P1 and P2 in Scotland; Y1 and Y2 in Northern Ireland) |
| Number of schools | 25 schools |
| Number of pupils | 400 pupils (16 per school) |
| Design | Family-randomised efficacy trial |
| Primary outcome | Pupils' receptive vocabulary using the BPVS3 six months after programme delivery |
| Secondary outcomes | Pupils' overall receptive vocabulary using BPVS3 immediately and six months after programme delivery<br>Pupils' receptive vocabulary using the BPVS3 immediately after programme delivery<br>Pupils' maths attainment using the PUMA immediately and six months after programme delivery<br>Pupils' social and emotional outcomes using the teacher-completed SDQ and CSS questionnaire immediately and six months after delivery |
| Other outcomes | Increased confidence and engagement in child's learning as measured by parental perceptions of parent efficacy, home learning environment and parent role construction scales using a parent questionnaire administered at baseline and immediately after programme delivery |

# 2 About the intervention and delivery

The Families Connect trial is a family level randomised efficacy trial, with two main arms, (intervention and control). Measurement involves trial participants from across the UK in Reception and Year 1 in England and Wales, P1 and P2 in Scotland and Y1 and Y2 in Northern Ireland. Families Connect (FC) is a parental engagement programme that has been designed by Save the Children UK (SCUK) to develop the skills and confidence of families in disadvantaged areas, and provide them with the resources to actively engage their children in learning in the home.

The FC programme consists of two phases. Phase 1 involves preparation for delivery. This includes the recruitment of schools, and within the schools, the recruitment of families through newsletters, texts and 'coffee mornings' which are hosted by the schools for the parents. Furthermore Community Practitioners (CPs) are trained across two days to deliver the programme. Phase 2 involves an 8-week programme where parents and children attend 8 two-hour sessions delivered by the CPs. Each session is split into two parts, the first hour involves CPs working only with the parents. In the second hour, the children are invited to join the session, and parents get the opportunity to practise what they have covered directly with their child. It is intended that the skills learnt during the sessions will be used further at home. And the intended long-term effects of FC are that children will have a greater chance of achieving their potential and doing better at school, as their parents will do more in the home to support their learning.

# 3 Study design

### Aims

The evaluation aims to explore the impact of FC on pupils' language, numeracy, social and emotional development, and on parents' level of engagement with their child's learning at home. The evaluation aims to establish to what extent FC is achieving its intended outcomes, and to what extent this would warrant further investment. As such, the specific research questions of the evaluation are listed below.

**The primary research question is**:

Does the programme make a difference to children's language development?

**The secondary research questions are:**

Does the programme make a difference to children's numeracy development?

Does the programme make a difference to children's social and emotional development?

Does the programme make a difference to parental engagement with children's learning?

**To address the primary and secondary research questions,** the trial will evaluate:

i. Receptive vocabulary for all pupils in the trial as measured by the British Picture Vocabulary Scale 3 (BPVS3), six months after programme delivery.

ii. Receptive vocabulary for all pupils in the trial as measured by the British Picture Vocabulary Scale 3 (BPVS3), immediately after programme delivery.
iii. Receptive vocabulary for all pupils in the trial as measured by the BPVS3, over both time points after programme delivery.
iv. Numeracy development as measured by the Hodder PUMA, immediately and six months after programme delivery.
v. Social and emotional development as measured by the teacher completed Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997), and the Child Softer Skills (CSS) (developed by SCUK) scale immediately and six months after programme delivery.
vi. Parental engagement in pupils' learning as measured by the Parental Perceptions of Parent Efficacy (PES), Parent Role Construction (PRC) and the Home Learning Environment KS1 (HLE) scales immediately after programme delivery.

## Trial design

The FC trial is a family-level randomised efficacy trial, with two main arms, (intervention and control). Schools that sign up to the trial aim to recruit 16 parents of pupils who are in Reception or Year 1 in England and Wales, P1 or P2 in Scotland and Y1 or Y2 in Northern Ireland. The parents may have more than one child (e.g. twins or siblings). Due to the nature of the intervention, it would not be possible to randomise children within a family, hence the family level randomisation, i.e. siblings will be randomised to the same group. Families will be randomised to either the intervention group, or the control group. Intervention group families will receive the intervention starting in January 2019. During this academic year the control group families will continue Business as Usual. Control group families will receive the intervention in the autumn term 2019 (i.e. near the start of the following academic year), after all trial measurement has been completed.

Due to difficulties recruiting schools initially, a booster recruitment phase of around 40 families was commissioned. This group was randomised in the same way as the main group mentioned above. Families randomised to the intervention group in the booster phase will start the intervention in the summer term of 2019. (The control group families in the booster phase will receive the intervention in the spring term 2020.) All pupil results will be analysed together with the two blocks being taken into account in the models.

## Eligible population

The following eligibility criteria are in place for the trial:

### Schools

Primary schools with Reception and Year 1 classes, in schools with over 20% free school meals (FSM) eligibility in England, and over 25% FSM eligibility in Wales; with Y1 and Y2 in schools in Northern Ireland with over 40% FSM eligibility; and with P1 and P2 in Scotland, in areas of disadvantage determined in consultation with the local SCUK manager. Schools who have taken part in one previous cycle of FC prior to the trial can take part, i.e. if a school has taken part in two previous cycles of FC, it will not be eligible.

### Families

Families with a child/children in Reception or Year 1 (in England and Wales) (the equivalent of P1/P2 in Scotland and Y1/Y2 Northern Ireland) in the academic year 2018/19. Families must not have taken part in FC before (for example with an older sibling). This information was collected on the baseline proforma.

### Age range

Although the trial is aimed at 4-6 year olds in R/Y1 in England/Wales (or equivalent year groups in the other countries), in practice some four year olds and six years old have been recruited to the trial from current nursery and Year 2 classes. As the primary outcome measure is adaptive, these pupils will still be eligible to take part in the trial.

### Disadvantage

FC is a universal programme. It is open to all and often runs with a mix of families from different backgrounds. This efficacy trial will not test the best combination of families (that may be better suited to a future school-randomised trial), but will aim to collect data on levels of disadvantage, for example household income (collected on a parent questionnaire using income bands) or FSM/Pupil Equity Fund eligibility of the families joining to monitor whether at least 20 per cent of them are from disadvantaged backgrounds.

### SEN

Children with special education needs (SEN) are eligible. Consideration will be needed as to the suitability of the outcome measures, in particular for those with visual impairments for the primary outcome (which uses visual cue cards).

### EAL

Where English is spoken by the family members as an additional language, schools and CPs will need to make local arrangements to include them in delivery (as per usual practice) and the trial. Parents whose first language is not English may need local support to complete the parent questionnaires (for example translators in situ). As the primary outcome for the trial (BPVS3) is administered in English and requires English vocabulary responses, Welsh-only speaking schools were not be eligible to take part in the trial.

## Recruitment

FC was promoted to families via newsletters, school texts and/or emails to parents, noticeboards, and coffee mornings/afternoons. The coffee morning/afternoons/open evenings are an open event where parents can find out about the programme, talk to CPs or SCUK and sign up to the trial. CPs could use their experience and knowledge of the families at their schools to identify/select families that they feel would ‘particularly benefit’ from the programme. CPs were asked to ensure families had a ‘cooling off’ period between being informed about the trial and signing up. The trial involves opt-in consent from families as the

legal basis under which NFER and partners are processing data. This is particularly in light of the personal sensitive data being collected via the SDQ.

# 4 Calculation of sample size

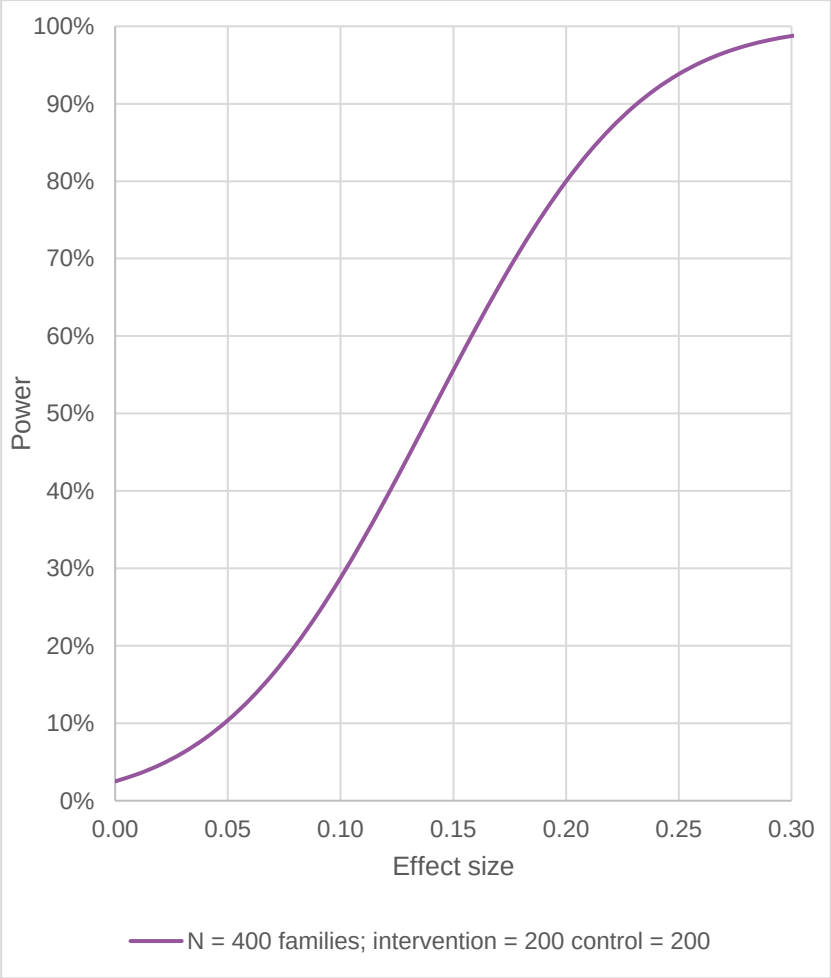
The trial is designed to measure 400 families from 25 schools (16 families per school) at analysis. This is based on both sample size calculations and pragmatic delivery reasons – the intervention is usually delivered in groups of between eight and ten families. 16 families per school will allow two groups of eight to take part – one as the intervention group, one as the waitlist control. A minimum of 12 families per school is recommended for pragmatic delivery reasons. This would result in more schools needing to be recruited, however this would not have an impact on the power calculations as this is a family level randomised trial[1]. Sample size calculations, informed by the analysis of SCUK data in advance of the trial (Rennie and Styles, 2018), are presented below. The following assumptions were used:

- a correlation between pre and both post-BPVS3 scores of 0.7 (secondary data analyses revealed correlations between baseline and follow-up of .75 and .76 depending on sample used (see Rennie and Styles, 2018, Technical Appendix on Secondary Data Analysis)
- an anticipated effect size of 0.2 (secondary analyses revealed a Hedge's G quasi-effect size of 0.29)
- no design effect through randomising within schools and only being concerned with internal validity i.e. no need to generalise the results of the trial to a wider population
- probability 0.05 of a Type I error
- 80 per cent power.

With these assumptions the model requires a minimum of 400 families to be randomised into two equally sized groups and analysed for the trial. These families will be distributed across approximately 25 schools based on an average attendance of eight families per delivery cycle and therefore 16 per school including controls. The power curve is displayed below in Figure 1.

[1] As noted in the protocol, this would, however, have a practical and budgetary impact on data collection and test administration with schools – and there would be an upper limit to the number of schools that can be included within the budget.

**Figure 1. Power curve**




It is important to stress this is the required analysed intention-to-treat sample size for the trial, with data points at baseline, follow-up one and follow-up two. Loss to follow-up may be caused by full school withdrawal, family drop-out, or pupil absence.

In order to account for this, SCUK had aimed to recruit an additional school per region (i.e. '25+5', so a total of 30 schools, six schools per region) and up to 20 families per school to allow for some pre- and post-randomisation drop out; the latter of which can be as high as six per cent in schools with a disadvantaged intake (DfE National Statistics). To allow for this we intended at least 440 families go forward to be randomised. Due to difficulties in securing this many families, a booster recruitment phase was commissioned to achieve the minimum intended sample. 378 families (391 pupils) were randomised from the original recruitment phase, and 105 families (108 pupils) were randomised in the booster phase. Final figures are shown below in Table 2.

# 5 Baseline respondents

All teachers and families (that had consented to join the trial) were asked to provide pupil administrative data (e.g. names, DoB, gender, FSM), to fill out the baseline surveys and complete the baseline tests prior to randomisation. This included a teacher questionnaire (TQ), a parent questionnaire (PQ) and the BPVS3. The PUMA test was not taken at baseline as the test was not appropriate to the age group. If administrative data and the primary outcome measure (BPVS3) were returned, a pupil was put forward for randomisation. If the BPVS3 was not returned due to a pupil being absent on testing day, this pupil was still included in the randomisation list. If the pupil had left school, they were removed from the trial. As such the following numbers refer to respondents who were included in the master trial list, and therefore put forward to randomisation for the trial.

**Table 1: Baseline survey response figures**

| Randomisation Phase | No. of pupils with family consent | Randomised | No. of baseline BPVS3 tests returned (from randomised group) |
|---|---|---|---|
| **1** | 396 | 391 | 382 (missing = 9) |
| **2** | 108 | 108 | 100 (missing = 8) |
| **Total** | 504 | 499 | 482 (missing = 17) |

# 6 Randomisation

As mentioned previously, consented families who returned their administrative data and BPVS3 records were put forward for randomisation. If a pupil was absent on testing day, they were still put forward for randomisation. Efforts were made to get the baseline measurements, however 17 records were not collectable. It was felt that ethically this should not be justification to remove the families from the trial – as parents had consented for their child to be part of the trial, other data was being collected about their child, and they had not withdrawn. If a pupil withdrew from the trial prior to baseline testing/randomisation, or had left school, they were removed from the master pupil list prior to randomisation. Randomisation was stratified by school. Due to the nature of the intervention involving parents working with their child(ren) and intending to impact on parental behaviour, it was not possible to randomise at a pupil level in case siblings should be allocated to different groups. Therefore randomisation was conducted at the family level. So as to equally distribute participants, and because over 2% of the families had more than one child taking part (see protocol), randomisation was also stratified by whether or not the family had more than one child.

A further randomisation was conducted to allocate families who were recruited during the second phase of recruitment. As such, randomisation took place in two phases, the first in January 2019, the second in March 2019. They were carried out by an NFER statistician using a full SPSS syntax audit trail. The results are presented below in Table 2.

**Table 2: Randomisation figures**

| Randomisation Phase | Families (I:C) | Pupils (I:C) |
|---|---|---|
| 1 | 378 (189:189) | 391 (193:198) |
| 2 | 105 (53:52) | 108 (54:54) |
| Total | 483 (242:241) | 499 (247:248) |

# 7 Outcome measurement

1. The British Picture Vocabulary Scale (3rd Ed. BPVS3) (Dunn et al., 2009) was chosen as it fits with the FC theory of change. Furthermore it has strong psychometric and implementation properties. It received ratings of 3/3 and 2/3 from the EEF database of early years' measures for the above properties respectively. Furthermore it has been used in previous Save the Children evaluations and aligns well with the language focus of the FC programme. It is a one-to-one teacher conducted assessment that measures the child's receptive vocabulary. The assessment will be administered by external test administrators. For this trial the raw score will be used as this still reflects the adaptive aspect of the test. The BPVS3 is a test appropriate for ages three to 18+. As such, as per the BPVS3 manual (Dunn et al., 2009), start sets are determined by the pupil's age at the time of testing. Raw scores are calculated by taking the highest mark of the highest set reached (ceiling set), and subtracting all the mistakes made between the basal set (the lowest set in which no more than one mistake is made) and the ceiling set. Start sets were calculated by an NFER statistician in SPSS using a full audit trail, using pupils' dates of birth, and a proposed date of test within two weeks of actual testing. As per the BPVS3 manual, although the start set is determined by age, the basal set may be lower than the start set if the pupil makes more than one mistake in the start set.
2. The Hodder Progress in Understanding Maths Assessment (Hodder PUMA test, McCarty and Cooke, 2015) was chosen as it is a nationwide standardised test which

has been aligned to the national curriculum. It aligns well with the FC programme to improve general attainment in numeracy. As the test is designed to evaluate Reception children in the summer, the youngest reception participants being measured at first follow-up (April/May 2019) may struggle with the test. Teachers and students will be made aware that this is to be expected to mitigate any negative impact this may have on the pupils and schools. Year 1 students, and all students at second follow-up should find the test to be of a suitable difficulty level. The assessment will be administered by external test administrators. The total raw scores will be used as the outcome measure.

3. The Strengths and Difficulties Questionnaire (SDQ, Goodman, 1997) was chosen as it is a reliable measure of pupils' emotional and social wellbeing and it has been used in previous FC cycles. The questionnaire consists of 25 items, split into 5 subscales with 5 items each (emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, prosocial behaviour). In addition, the SDQ consists of an impact supplement with a further five items. For this trial we will use three measures: the total difficulties score (the sum of emotional symptoms, conduct problems, hyperactivity/inattention and peer relationship problems), the prosocial score, and the impact score. These measures will be computed using the syntax published by Youth in Mind[2]. The SDQ will be teacher assessed.

4. The Child Softer Skills scale (CSS) is a bespoke 12 item scale designed by SCUK to be an age appropriate measure of children's attitudes and behaviours towards learning (Bradley et al., 2016). It was chosen as it aligns well with the intended impact of the FC programme, and had been used in previous FC cycles. The raw score of the 12 items summed will be used as the outcome measure. It will be teacher assessed.

5. The Key Stage 1 Home Learning Environment Scale (HLE) was chosen as it is a measure that captures the frequency of a range of general and work specific interactions between parents and pupils at home (Sylva *et al*., 2008). The activities in the scale align well with the activities practiced in the FC programme. It consists of 4 factors with three items in each; 'Home computing', 'One-to-one interaction', Expressive play' and 'Enrichment outings'. The outcome measure for this trial will be the raw sum of the 12 items. It is a parent-assessed measure.

6. The Parent Role Construction (PRC) scale is a subscale (Role Activity Beliefs) of the original Parental Role Construction for Involvement Scale (Hoover-Dempsey and Sandler, 2005). It is a reliable measure that captures a parent's belief about what they should be doing with regards their child's education, and can be used as an independent scale. It is a parent assessed measure that consists of 10 items. For the

[2] https://sdqinfo.org/c1.html

trial, the sum of the 10 items will be used as the outcome measure.

7. The Parent Efficacy Scale (PES) is a reliable measure of parents' beliefs about their ability to influence their child's educational outcomes (Hoover-Dempsey and Sandler, 2005). It is a parent assessed measure that consists of seven items. Some items are reverse coded. For the trial the sum of the seven items (reversed where appropriate) will be used as the outcome measure.

**Table 3: Measurement time points**

| Outcome | Measured by | Baseline | Follow-up 1 Immediately after programme delivery | Follow-up 2 Six months after programme delivery |
|---|---|---|---|---|
| 1. BPVS3 | Administrator | Y[1] | Y[2] | Y[2] |
| 2. PUMA | Administrator | N/A | Y[2] | Y[2] |
| 3. SDQ | Teacher | Y | Y | Y |
| 4. CSS | Teacher | Y | Y | Y |
| 5. HLE | Parent/guardian | Y | Y | N/A |
| 6. PRC | Parent/guardian | Y | Y | N/A |
| 7. PES | Parent/guardian | Y | Y | N/A |

Table notes: [1]administered by SCUK staff trained in BPVS3 administration and not directly connected with the school; [2]administered by NFER test administrators trained in BPVS3 and PUMA administration.

Measures five, six and seven are collected via the parent questionnaire (PQ) which will be measured at baseline and immediately after programme delivery. The PQ follow-up measures will be completed at the end of the 8 week programme (in Week 8, or immediately after for the intervention parents, and during a convened coffee morning during this period for control parents, noting that control parents must not attend a programme session to do this).

# 8 Analysis

As family was the unit for randomisation, all analysis of pupil outcomes will be at the family level. For families who have more than one sibling taking part in FC, the mean response will be used. However, note that the outcomes are described as pupil outcomes throughout.

### Primary outcome analysis

The primary outcome analysis of the BPVS3 raw scores will be 'intention-to-treat' (ITT). The primary model will be a multilevel model with three levels (school, pupil and time point). Pupils who have measurements at baseline and follow-up one and/or follow-up two will be included in the model, regardless of whether their school implemented the intervention, or the family took part.

The dependent variable for the model will be the BPVS3 raw scores at follow-up one and follow-up two with the following covariates:

- A dummy variable indicating family group allocation
- A dummy variable indicating randomisation phase
- A dummy variable indicating whether the family has more than one child
- Prior attainment as measured by the baseline measurement of the BPVS3
- A dummy time variable indicating 2nd follow-up
- An interaction variable time*intervention

The main effect of schools will be estimated in this model, i.e. school slopes will be fixed.

This model will determine whether the FC programme has an overall impact on pupils' receptive vocabulary, and if any impact has enhanced or attenuated over time through the use of the interaction term. Note, although this model will provide information on effects at both time points, and an overall effect, the primary outcome of interest will be the second time point. The overall and immediate effects will be treated as secondary outcomes.

### Further analysis of the BPVS3

The secondary ITT analysis of BPVS3 raw scores will be another multilevel model with three levels (school, pupil and time point). The dependent variable for the model will be the BPVS3 raw scores at follow-up one and follow-up two with the following covariates:

- A dummy variable indicating family group allocation
- A dummy variable indicating randomisation phase
- A dummy variable indicating whether the family has more than one child
- Prior attainment as measured by the baseline measurement of the BPVS3
- A dummy time variable indicating 2nd follow-up
- An interaction variable time*intervention

Whereas in the primary model, only the main effect of schools will be estimated, in this model the main effect of school and school by treatment effects will be estimated, i.e. school slopes will be random. This will be achieved by making the intervention variable random at the school level.

This model will determine whether the FC programme has an overall impact on pupils' receptive vocabulary, and if any impact has enhanced or attenuated over time through the use of the interaction term. Furthermore the model will explore potential differential effects of the intervention across schools.

## Effect Size

Effect sizes and confidence intervals will be calculated for all outcome models. The effect of the intervention will be determined by converting the coefficients (in the respective models) of the intervention group variable into Hedges g effect sizes. This will be done using the following formula:

$$g = \frac{\bar{x}_1 - \bar{x}_2}{S^*}$$

Where $\bar{x}_1 - \bar{x}_2$ denotes the model coefficient representing the mean difference between the intervention and control groups, while adjusting for the model covariates.

$S^*$ is the standard deviation. For all models this will be the square-root of the total school plus pupil variance from a model without covariates i.e. for repeated measures models it will not include the time-level variance. This is to enable comparisons with simpler models with only one follow-up time point.

Confidence intervals for the effect sizes will be derived by multiplying the standard error of the intervention group model coefficient or relevant contrast by 1.96. These will be converted to effect size confidence intervals using the same formula as the effect size itself.

## Sub-group analysis

To investigate whether the FC programme has differential effects for families from disadvantaged backgrounds, a subgroup analysis will be conducted on the primary outcome. A multilevel model with three levels (school, pupil, time point) will be run with the following covariates:

- A dummy variable indicating family group allocation
- A dummy variable indicating randomisation phase
- A dummy variable indicating whether the family has more than one child
- Prior attainment as measured by the baseline measurement of the BPVS3
- A dummy time variable indicating 2nd follow-up
- A dummy disadvantage variable indicating total household income below £20,000[3] PA
- An interaction variable income*intervention

This model will determine whether the FC intervention has a differential effect on receptive vocabulary for disadvantaged families compared to non-disadvantaged families.

To investigate whether the FC programme has differential effects for pupils with SEN, a subgroup analysis will be conducted on the primary outcome at time point two. A multilevel model with three levels (school, pupil, time point) will be run with the following covariates:

- A dummy variable indicating family group allocation
- A dummy variable indicating randomisation phase

[3] Note that this threshold does not take into consideration the size of household, therefore should be considered an approximate measure of disadvantage.

- A dummy variable indicating whether the family has more than one child
- Prior attainment as measured by the baseline measurement of the BPVS3
- A dummy time variable indicating 2nd follow-up
- A dummy SEN variable indicating if the pupil has special educational needs as noted by the parent (if only one sibling has SEN only the data from that child will be included in this model, not the average outcome)
- An interaction variable SEN*intervention

This model will determine whether the FC intervention has a differential effect on receptive vocabulary for pupils with SEN compared to pupils without SEN.

### Imbalance at baseline

We will not explore imbalance at baseline on BPVS3 score since this is a covariate in the final model. Using a baseline comparison table we will explore differences in proportions of FSM eligibility, EAL, SEN and Age on analysed groups.

### Missing Data

We will assess the level and pattern of missing data from the primary model. In order to assess the missingness mechanism, we will run a logistic regression model on whether a case had follow-up data for the primary outcome, regressed on the covariates of the primary outcome model plus other variables. For this project we are not accessing NPD data, and are therefore limited to using the data collected. We propose a methodology that includes all the variables included in the primary analysis model, plus other variables collected such as FSM, household income bands, EAL, SEN, guardian's level of education, if the guardian changed during the programme, and pupil's age. We will run a logistic regression to check if any of the above variables predict missingness at follow-up.

Under the 'missing at random' assumption we would expect a completer's analysis to be unbiased. If the extent of dropout was unequal between the groups, the missing not at random (MNAR) assumption is likely to hold and we will conduct sensitivity analysis. This will be done after running multiple imputation. If the aforementioned variables are found to be predictive of non-response, they will be used in the multiple imputation process using the mice package in R. The number of datasets is dependent on the amount of missing data but a minimum would be five datasets, with a minimum of ten iterations. These iterations are necessary as with only one dataset, the parameter estimates have more sampling variability. Multiple iterations also help in generating the estimates of the standard errors to accurately reflect the uncertainty about the missing values (Allison, 2012). The model would then be extended using a weighting approach according to Carpenter et al. (2007). Missing data analysis will only be possible in cases where we have the data for the aforementioned variables.

If there is only a small amount of missing baseline data, we will use the simpler methods described in White and Thompson (2005).

### Effects in the presence of non-compliance

Due to family schedules, it is likely that not all families will attend all 8 FC programme sessions. With the use of a Family Register, SCUK will record how many sessions each pupil attends. We will produce descriptive statistics on this attendance data. The main analysis will be followed by a CACE analysis (Complier Average Causal Effect) in order to assess the effect of non-compliance on the outcome measure where data from the Family Register will be used to determine the extent of each pupil's involvement. Families are asked to attend a minimum of one session from each of the three programme topics and a minimum of five sessions in total. This attendance guidance came from the Families and Schools Together programme. Two measures of compliance will be used. The first is a binary variable indicating whether five sessions were attended, (with at least one session from each of the three topic areas was attended) or not. The second measure is a continuous variable, indicating how many sessions were attended from zero to eight, regardless of which sessions they were.

Families may potentially have unobserved characteristics that have an influence on both the compliance with the intervention and academic attainment. Therefore, a two-stage least squares model will be used to calculate the CACE estimate (Angrist and Imbens, 1995). The first stage of the model will be compliance regressed on all covariates that are used in the main primary outcome model and the group allocation variable. The second stage of the model will regress the primary outcome on the covariates used in the main model and will also include a covariate representing the pupil's estimated level of compliance from the first stage of the model and an interaction term between the estimated compliance and the pupil's group allocation. The coefficient of the interaction term is the CACE estimate of the compliance effect. In the event that there are no confounding factors affecting compliance and attainment the CACE estimate will be equal to the intention-to-treat estimate. We will use the R package ivpack to perform the CACE analysis on the primary outcome only.

### Secondary outcome analysis: pupils

The secondary outcome analysis of numeracy will be ITT analysis of the raw PUMA score at both follow-up time points. The PUMA test was administered twice after the delivery of the intervention. It was not measured at baseline as it is not age appropriate for the beginning of reception. As such, the baseline BPVS3 score will be included as a covariate as a proxy measure of prior attainment. All pupils with measurements at baseline BPVS3 and either or both of the two PUMA follow-ups will be included in the model. The model will be a multi-level model with three levels (school, pupil and time point) and will include the following covariates:

- A dummy variable indicating family group allocation
- A dummy variable indicating randomisation phase
- A dummy variable indicating whether the family has more than one child
- Prior attainment as measured by the baseline measurement of the BPVS3
- A dummy time variable indicating 2nd follow-up
- An interaction variable time*intervention

This model will determine whether the FC intervention has an impact on pupils' numeracy, and if any impact has enhanced or attenuated over time.

The secondary outcome analysis of social and emotional development will be ITT analysis of the total difficulties score, the prosocial score, and the impact score of the SDQ (as computed using the syntax published by Youth in Mind). Three multilevel models with three levels (school, pupil and time point) will be run. All pupils with measurements at baseline and either or both follow-up time points, for each scale, will be included in each model. Each model will have the following covariates:

- A dummy variable indicating family group allocation
- A dummy variable indicating randomisation phase
- A dummy variable indicating whether the family has more than one child
- Baseline measures of the respective scales
- A dummy time variables indicating 2nd follow-up
- An interaction variable time*intervention

These models will determine whether the FC intervention has an impact on pupils' social and emotional wellbeing, and if any impact has enhanced or attenuated over time.

Furthermore, the secondary outcome analysis of social and emotional development will be ITT analysis of the total raw score of the child softer skills (CSS) scale. A multilevel model with three levels (school, pupil and time point) will be run. All pupils with measurements at baseline and either or both follow-up points will be included in the analysis. The model will have the following covariates:

- A dummy variable indicating family group allocation
- A dummy variable indicating randomisation phase
- A dummy variable indicating whether the family has more than one child
- Baseline measure of the CSS scale
- A dummy time variable indicating 2nd follow-up
- An interaction variable time*intervention

These models will also determine whether the FC intervention has an impact on pupils' social and emotional wellbeing, and if any impact has enhanced or attenuated over time.

### Secondary outcome analysis: parents

The secondary outcome analysis of parental engagement in pupils' learning will be ITT analysis of the PES total raw score, the PRC total raw score and HLE total raw score. Three, two level (school and pupil) multilevel models will be run on the respective outcomes, measured immediately after programme delivery. All pupils with data on the respective outcome at baseline and follow-up will be included in the model. The following covariates will be included:

- A dummy variable indicating family group allocation
- A dummy variable indicating randomisation phase

- A dummy variable indicating whether the family has more than one child
- Baseline measures of the respective scales

These models will determine whether the FC intervention has a short term effect on parent's engagement in their child's learning at home.

# 9 References


Allison, P. (2012). Why maximum likelihood is better than multiple imputation. *Statistical Horizons*.

Angrist, J. D. and Imbens, G. W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American statistical Association*, *90*(430), 431-442. Available: https://scholar.harvard.edu/imbens/files/wo-stage_least_squares_estimation_of_average_causal_effects_in_models_with_variable_treatment_intensity.pdf [08 March 2019].

Bradley, C., Wilson, L., Anfield, G. and Magness, J. (2016) Families Connect: Evaluation of Summer Delivery 2016. Save the Children [online]. Available: https://www.savethechildren.org.uk/what-we-do/uk-work/in-schools/families-connect [08 March 2019].

Carpenter, J. R., Kenward, M. G. and White, I. R. (2007). Sensitivity analysis after multiple imputation under missing at random: a weighting approach. *Statistical methods in medical research*, *16*(3), 259-275. Available: https://journals.sagepub.com/doi/abs/10.1177/0962280206075303

Department for Education (2016). Pupils' absence in schools in England: 2015 to 2016. UK Government. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/602320/SFR14_2017_Text.pdf [08 March 2019]

Dunn, L.M., Dunn, L.M., Burge, B. and Styles, B. (2009). British Picture Vocabulary Scale (3rd Ed. BPVS3). GL/Assessment. www.gl-assessment.co.uk.

Goodman, R. (1997). 'The strengths and difficulties questionnaire: a research note', The Journal of Child Psychology and Psychiatry, 38, 5, 581-586 [online]. Available: http://onlinelibrary.wiley.com/doi/10.1111/j.1469-7610.1997.tb01545.x/abstract [08 March 2019].

Hoover-Dempsey, K. and Sandler, H. (2005). The Social Context of Parental Involvement: A Path to Enhanced Achievement. Final Performance Report for OERI Grant #R305T010673. Retrieved from The Family-School Partnership Lab, Vanderbilt University, U.S. [online]. Available: http://www.vanderbilt.edu/Peabody/family-school/Reports [08 March 2019].

McCarty, C. and Cooke, C. (2015). Progress in Understanding Maths Assessment (PUMA). RS Assessment, from Hodder Education

Rennie, C. and Styles, B. (2018). Evaluation of data from the Families Connect Programme: A Technical Appendix [online]. Available [forthcoming].

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. and Taggart, B. (2008). Final Report from the Primary Phase: Preschool, School and Family Influences on Children's Development During Key Stage 2 (Age 7-11). Nottingham: DfES [online]. Available: http://ro.uow.edu.au/cgi/viewcontent.cgi?article=2806&context=sspapers [08 March 2019].

White, I.R. and Thompson, S.G. (2005) Adjusting for partially missing baseline measurements

in randomized trials Statist. Med. 2005; 24:993–1007

NFER
National Foundation for Educational Research

# Evidence for excellence in education



The Mere, Upton Park, Slough, Berks SL1 2DQ

**T:** +44 (0)1753 574123 • **F:** +44 (0)1753 691632 • enquiries@nfer.ac.uk

www.nfer.ac.uk