Success for All

Queen's University Belfast
Andy Biggart and Liam O'Hare

Education Endowment Foundation

Evaluation Summary	
Age range	Reception/Year 1
Number of	1250 pupils
Number of	50 schools
Design	Whole school randomised controlled trial
Primary	Reading

Amendments

Following initial planning and development work between CESI and IEE as well as the extension of the study with the addition of a second cohort and changes to EEF guidance the following changes have been made to the initial protocol:

- Section 1 – The centre name of the indepent evaluators has been revised from the Centre for Effective Education (CEE) to the Centre for Evidence and Social Innovation (CESI), following the establishment of this new interdisciplinary research centre at Queen's Univerity. Seaneen Sloan has also been added to the evaluation team as the main fieldworker and person responsible for liason with IEE over data collection and preparation.

- Section 2.2.1 - Random allocation has been updated to reflect the recruitment of a second cohort following the early drop-out of seven schools from the intervention group in the first cohort. Withdrawn schools from the first cohort are being followed up on an intent to treat to basis.

- Section 2,2,1 – The method of random allocation has been confirmed with schools being allocated in pairs based on a ranking of Key Stage 2 results.

- Section 2.2.4 - Greater clarification has been provided in relation to the agreed analysis plan and the specification of the multi-level models.

- Section 2.2.4 - The inclusion of a secondary outcome (Attention Deficit related behaviours) had originally been suggested by the independent evaluators based on a previous literacy evaluation. This was primarily of academic interest and has not been funded or collected by the IEE.

- Section 2.2.4 - The proposed sub-group and exploratory analyses have been updated and clarified. This now includes the EEF requirement for analysis in relation to pupils eligibile for free school meals. In-line with EEF guidance sub-group analysis has now been restricted to prior attainment at baseline.

- Section 2.2.4 – The handling of missing data has been clarified to include a sensitivity analysis using multiple imputation if missing data >5%.

- Section 3 - The timing of the survey of teachers in Cohort 1 has been revised due to recruitment and retention issues experienced, and a survey and telephone interviews of the second cohort have been added to the process evaluation.

- Section 6 - The timeline has been updated to reflect the above changes..

1. Project team

The project team will be drawn from senior and experienced staff within the Centre for Evidence and Social Innovation (CESI) at Queen's University Belfast. The CESI have considerable experience in the conduct and analysis of randomised control trials and cluster randomised control trials in educational and community settings. They will draw on their expertise in relation to trials of literacy interventions having conducted 6 randomised control trials in the area of literacy over the past 5 years. The Centre has developed particular expertise in working closely with programme developers in the design, analysis and reporting of trials, while providing a thorough and robust independent evaluation of programme impacts. The team has also experience in accounting for the nested nature of the data when schools are randomised at the school level through the use of multi-level modelling.

Project Team:

Dr Andy Biggart: Principal Investigator: is a Fellow of the CESI. He is an experienced research project manager on large scale mixed methods studies. He has recently been Principal Investigator for a number of large scale randomized trial evaluations in Ireland and the U.K., including a three year study of the Doodle Den balanced literacy after-school program. He is also a UK partner on a major comparative European Union funded study into educational disadvantage in Europe. He has over 20 year's research experience primarily in the conduct of evaluation and policy relevant educational research on children and young people.

Dr Sarah Miller: Co-investigator is a Fellow of the CESI and a psychologist with a strong quantitative and statistical background. Her experience of supervising and conducting research projects spans randomised controlled trials, systematic reviews and large scale surveys. She is Principal Investigator of an RCT of the Lifestart Parenting Programme as well as an RCT evaluation of the Business in the Community Pupil Mentoring Scheme. Dr Miller is experienced in the measurement of developmental outcomes in young children including cognitive development, and literacy outcomes.

Dr Liam O'Hare is Senior Research Fellow within the CESI, and he has been Principal Investigator on a range of RCT evaluations and has particular expertise in psychometric measurement and testing.

Dr Seaneen Sloan is a Research Fellow in the CESI, and has experience in the conduct of fieldwork and the analysis of randomised control trials in relation to early child development and in relation to both academic and socio-emotional well-being.

Advisors:

Professor Paul Connolly – is Professor of Education and Dean of Research in the Faculty of Arts, Humanities and Social Sciences at Queen's University Belfast. He was the founding Director of the former Centre for Effective Education and is currently the Director of the Centre for Evidence and Social Innovation. Through the Centre, he has led a number of large-scale randomised trials in Northern Ireland, the Republic of Ireland and England. Professor Connolly also has a strong interest in

systematic reviews and meta-analysis and was Co-Chair of the Education Coordinating Group of the international Campbell Collaboration.

Professor Carol McGuinness: is Professor of Psychology at Queen's University Belfast, and is an expert in early years curriculum and literacy. She is author of the influential report, *From Thinking Skills to Thinking Classrooms,* which was commissioned by the UK Department of Education (1999) and recent research funded through ESRC TLRP programme evaluated the impact of a methodology to enhance children thinking skills in primary school classrooms.

Examples of relevant projects conducted by the team:

Literacy Randomised Trials:

- A three year randomised control trial evaluation of the Doodle Den Balanced Literacy After- school Programme with 5 and 6 year old children in Dublin (Biggart, Connolly, O'Hare & Kerr).

- A Cluster Randomised Controlled Trial Evaluation of Booktime England: A Book Gifting Intervention for Reception-Aged Children (Biggart, O'Hare & Connolly).

- A Cluster Randomised Controlled Trial Evaluation of Booktime Northern Ireland: A Book Gifting Intervention for Reception-Aged Children. (Connolly & O'Hare).

- Business in The Community Time to Read Progamme (Miller and Connolly).

- A Randomised Controlled Trial Evaluation of Bookstart+: A Book Gifting Intervention for Two- Year-Old Children. (Connolly & O'Hare).

Other cluster randomised trials:

- A cluster randomised controlled trial evaluation and cost-effectiveness analysis of the Roots of Empathy schools-based programme for improving social and emotional wellbeing outcomes among 8-9 year olds in Northern Ireland (Miller & Connolly).

- Cluster Randomised Controlled Trial Evaluation of the Effects of Sesame Tree on 5-6 Year Old Children's Attitudes and Awareness (Connolly).

- Cluster Randomised Controlled Trial Evaluation of the Effects of the Sesame Tree Outreach Pack on 5-6 Year Old Children's Attitudes and Awareness (Connolly).

- Cluster Randomised Controlled Trial Evaluation of the Effects of the 'Eager and Able to Learn' Programme on 2-3 Year Old Children's Early Dispositions Towards Education. (Connolly & Miller).

- Cluster Randomised Controlled Trial Evaluation of the Effects of the Media Initiative for Children on 3-4 Year Old Children's Attitudes Towards Diversity and the Republic of Ireland. (Connolly & Miller).

Roles and responsibilities:
Dr Andy Biggart is the Principal Investigator and will have responsibility for overseeing all aspects of the design, randomisation and the analysis and write up of the data. He will be supported by Dr Sarah Miller who has extensive experience in the conduct of CRCT's and multi-level modelling in relation to CRCT's. Dr Liam O'Hare will provide support in relation to psychometric measurement and the conduct of the process evaluation surveys through Questback. Dr Seanean Sloan will be responsible for the conduct of the fieldwork and liaison with the IEE over data collection.

2. Impact evaluation summary main study

2.1 Context and design rationale

The great emphasis that has been placed on the development of children's early literacy skills is very much related to its role as a gateway subject. Literacy skills are widely recognised as an important precursor to general academic achievement as well as in relation to broader participation in society. Longitudinal studies have also shown that children who fail to gain adequate basic literacy skills at an early stage are unlikely to catch-up later (Brooks, 2007; Francis, Shaywitz, Stuebing, Shaywitx & Fletcher 1996; Juel, 1988).

There has, however, been considerable debate and controversy over the best approaches for the teaching of literacy, especially among struggling beginning readers. This has been exemplified by what has been termed the 'literacy wars' between whole-language approaches and those that advocate the teaching of phonics. A systematic review limited to evidence from randomized trials concluded there was evidence that systematic phonics instruction was more effective than whole language or word approaches (Torgerson, Brooks & Hall, 2003). Although the importance of the teaching of phonics has a strong evidence base, many reviews have concluded that it is insufficient on its own. Cowen (2003) for example, synthesized six major research studies which considered the early stages of learning to read and concluded that direct phonics teaching should not be taught on its own, in isolation from meaning and understanding. A number of national reviews, including the Rose Report, support this increasing consensus that a variety of approaches are required, including the use of systematic phonics, to support the literacy needs of all children (Rose, 2006; NICHD, 2000; Rowe, 2005).

This more balanced approach is reflected in the Success for All programme which is a multi- component programme that includes the teaching of synthetic phonics with other evidenced based components, but also includes a major focus on professional staff development and school-wide structures. These have been identified as important components of successful literacy programmes (Slavin et al, 2007). The main impact evaluation will focus upon providing an independent assessment of the effectiveness of the main Success for All programme in improving struggling readers literacy skills in English schools. The Success for All programme has good evidence of effectiveness within the United States where a large number of evaluations have recorded positive improvements in children's literacy with pooled average effect sizes of around +0.5. However, many of these studies have involved quasi-experimental matched non-randomised designs, which have a tendency to inflate effect sizes (Wilson & Lipsey, 2007). There has been one major cluster randomised trial of Success for All in 41 schools in the US (Borman et al, 2007), the results of which were also positive but with more modest effect sizes (0.2-0.33). A number of smaller scale studies have been conducted in the UK of the programme which have shown positive effects in the English context (Slavin, Wordworth & Jones-Hill, 2005: Harris, Hopkins & Wordworth, 2001) although another evaluation found mixed effects (Tymms & Merrell, 2001).

While there is a substantial body of existing evidence of the effectiveness of Success for All both in the US and the UK, it has only been evaluated through one large scale RCT in the US. There is therefore a need for the conduct of a high quality rigorous independent randomised control trial to strengthen the evidence behind the programme and its implementation in the UK context.

2.2. RCT: Impact of Main Success for All Programme

The Institute of Effective Education have been provided with funding from the EEF to undertake the fieldwork for the evaluations and this proposal outlines the independent oversight of the work and the respective roles of CESI and IEE.

2.2.1 Pre trial support and design

- **Sampling of Schools**: The schools sample will be geographically dispersed throughout England although will be concentrated in the North and the Midlands.

- **Random allocation of schools to intervention groups.** 50 schools were proposed for the random allocation. Schools would also need to agree to take part in the evaluation for the full two years and to gain parental consent prior to randomisation. The randomisation process involved ranking schools in terms of the proportion of pupils above Level 4 in English and Maths at Key Stage 2 (averaged over 3 years where possible) and schools were then randomly allocated in pairs.

- Recruitment was challenging due to the level of commitment required from schools. An initial batch of 39 schools were randomised in a first round of schools. Following some early drop-out of intervention schools (n=7) and the initial difficulties in recruiting the number of target schools, EEF agreed for an additional round of recruitment and the introduction of a second cohort of schools the following year. An additional group of 14 schools have been recruited as a second cohort to maintain study power. All schools in both cohorts will be followed up on an intent-to-treat basis.

- Schools that form part of the control group will be compensated £2,000/year for their participation in the evaluation.

- **Measures used in the impact assessment.** These were originally pre-specified by the IEE for evaluation (outlined below), these seem appropriate and there are advantages of using similar measures for the purposes of synthesis. IEE should seek joint agreement with the CESI over the final format of measures to be assessed providing full details of the final testing materials, format of collection and the details of the way in which any additional demographic data on the children will be obtained.

2.2.2 Measures

The pre-specified measures are age appropriate instruments that have been used widely in evaluations of literacy abilities in younger children, and have been normed on large samples. All display satisfactory reliability. The British Picture Vocabulary Scale (BPVS) provides a standardised pre-test score.This will provide a baseline measure in the analysis of post-test outcomes.

Post-test measures

Pupils will be post-tested using a variety of measures at several time points. The assessments will use the Woodcock Reading Mastery Test lll sub-scales: Letter Identification, Word Identification and Word Attack. These sub-tests assess basic literacy skills at the end of Reception. Further post-testing will be conducted at the end of Year 1. This testing will include the Woodcock sub-scales Word Identification and Word Attack as well as Passage Comprehension. The Woodcock scales are age appropriate measures although they have not been normed on a UK based sample. However, the 3rd edition has been approved for use in the UK by the SpLD Assessment Standards Committee (with the caution that it may require revision for Americanisms and other cultural features).

Woodcock Reading Mastery lll Sub-scales

Letter Identification (End of Reception)
Word Identification (End of Reception)
Word Attack (End of Reception)
Word Identification (End of Year 1)

Word Attack (End of Year 1)
Passage Comprehension (End of Year 1)

Key Stage 1 assessments (Phonics check Year 1)

2.2.3 Test administration

Test administration has already been funded to be collected by the IEE. In particular, it will be important that the test administrators are blinded to treatment group. Following test-administration the raw test questionnaires should be sent to the CESI for data entry. For quality control purposes we feel it is important for the main study that the data entry is conducted by the independent evaluation team, but protocols would be put in place for the early sharing of data at the conclusion of the study.

2.2.4 Data Analysis

Analysis will be conducted on an intention-to-treat basis on a combined dataset of the two cohorts. The initial characteristics of the intervention and control groups will be compared at baseline in relation to their core characteristics: e.g. gender, FSM eligibility, and mean scores on the pre-test.

The main effects of the intervention will be estimated using multilevel modelling to take account of the clustered nature of the data and a series of models will be estimated for each outcome (where pupil is level 1 and school is level 2). Firstly, a simple analysis will be conducted: the relevant outcome measure at post-test forming the dependent variable and the independent variables including a dummy variable representing whether the child was a member of the intervention or control group (coded '1' and '0' respectively) and pupils' baseline scores on the pre-test (BPVS). Then, a series of pupil level and school level characteristics will be added as covariates to control for any baseline differences in the variables and to accommodate variables used in the allocation process (school-level attainment at KS2). If necessary, standard errors will be bootstrapped as a test of robustness.

The main focus for the analysis will be the estimated coefficient associated with the dummy variable that represents the difference in mean scores on the respective outcome variables between the intervention and control groups, once baseline scores and other covariates are controlled for. These coefficients will then be used to estimate the effect size of the programme in relation to the respective outcome variables as the standardised mean differences between the two groups at post-test (Hedges' g).

To estimate the effect of the intervention for children eligible for FSM the main analysis will be repeated on a subsample of the students that were identified as eligible for FSM.

An additional sub-group analysis will examine the differential response to the intervention according to different abilities at baseline (BPVS).

Further exploratory analysis on the intervention group will examine variation in outcomes according to the rated level of implementation in different schools.

If the proportion of missing data is low (less than 5%) we will use list-wise deletion of data, otherwise data will be imputed using multiple imputation which will be presented as a sensitivity analysis.

Primary Outcome Measures at Post-test:

A range of primary literacy outcomes will be assessed at two different time points (End of Reception Class and at the end of Year 1).

Woodcock Reading Mastery subscales:

End of Reception
Letter Identification
Word Identification
Word Attack

End of Year 1

Word Identification
Word Attack
Passage Comprehension

Secondary Outcome:

A secondary outcome is the Key Stage 1 (Phonics Check) which will be gathered from NPD data following completion during Year 1.

As stated above the BPVS measure will act a pre-test score for all statistical models.

2.2.5 Sample Size

Assuming an average class size of 25 and 50 schools the main evaluation of SfA should have an overall sample size of around 1250. Power calculations suggest that with 50 schools we should be able detect effect sizes of around 0.2. Using Optimal Design software with the parameters outlined below the current study design has a power of 80% to detect an effect size of 0.22.

The following parameters were used in the power calculation:

- Significance level (α) = 0.05
- Power (P) = 80%
- Cell size (n) = 25
- ICC (ρ) = 0.10
- Proportion of variation at level 2 (R^2_{l2}) = 0.60

3. Process evaluation summary

CESI will undertake a detailed process evaluation of the main programme to include a survey of key stake-holders, observation of the programme in a sample of schools and in-depth interviews to examine implementation and fidelity to inform scale-up and wider roll-out.

CESI will conduct an online questionnaire survey of teaching staff and Success for All Coordinators in each participating school responsible for leading the intervention. The survey will focus on implementation with key issues (to include fidelity, exposure, quality of delivery and participant and deliverer engagement) after the programme had bedded down (towards the end of the first year of implementation).

The survey will be repeated with teaching staff among the second cohort of 7 intervention schools towards the end of the second year and follow-up telephone interviews targeted at SfA Coordinators and Year 1 teachers in each of the schools. These will focus on any issues that arose over the course of the intervention year: how well the approach worked overall, views of what worked well and what did not work so well.

All responses to on-line surveys will be hosted by the on-line questionnaire provider Questback. Questback has been used by the CESI to conduct on-line evaluations of this kind in other research projects. Response rates would be maximised by follow-up telephone calls to any non-respondents, who would be encouraged to complete the questionnaire by phone. Using this method in previous studies we have achieved up to 100% response rates. Closed questions will be analysed using descriptive statistics and open questions will be thematically coded.

3.1 Observations and in-depth interviews

A small number of schools will be selected for observation and in-depth interviews by CEE (maximum 10) with the lead member of staff in selected schools, or other appropriate stakeholders. This will gather in-depth information on the extent to which the scheme was implemented as envisaged, what issues were encountered, and for whom, and how these were addressed. In addition observations will be conducted in these schools during the delivery of the programme in order to examine fidelity and any problems encountered such as problems with the timing of activities or issues with differentiation

according to ability.

SFA trainers will also collect data on implementation which will feed into the main impact assessment.

4. Main issues or risks to the evaluation and how they would be addressed.

Risk	Assessment	Countermeasures and contingency plan
Schools decide they no longer want to be part of evaluation following randomisation	***Likelihood:*** Low ***Impact:*** Low	Control schools are being provided with £2,000 compensation
IEE and CEE have differences of opinion on trial design, measures or approach to analysis	***Likelihood:*** Medium ***Impact:*** Medium	Early project initiation meeting with IEE and EEF to finalise project design and agree measures. CEE staff have experience of working closely with programme developers in a flexible way while maintaining the robustness of the study design and independence of evaluation.
Differential Pupil Attrition from control and intervention groups	***Likelihood: Low*** ***Impact: Low***	With a well-designed trial of this size we would expect some attrition but with this sample size this should be evenly matched between control and intervention schools. Imputation methods used if required
Lack of study power	***Likelihood: Low*** ***Impact: Low***	Some smaller observed effect sizes may not be significant. This will be dealt with in the interpretation of the impact results.
Data protection and ethics	***Likelihood:*** Low. ***Impact:*** High.	Robust data protection and ethical procedures are in place at CEE. York have already obtained ethical approval for the study testing and data sharing protocols will be established.
Staffing issues: staff leaving/unavailable over extended duration of project	***Likelihood*:** Medium (turnover low) ***Impact:*** High.	Succession planning has been built into team roles. Large team can absorb problems in the short-term. Sufficient numbers of experienced staff in senior roles to cover others in the team.

6. Timeline (See below)

References

Brooks, G. (2007). *What works with children with literacy difficulties?: The effectiveness of interventions.* Slough: National Foundation for Educational Research.

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A. and Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology*, 88, 3-17.

Juel, C. (1988). Learning to Read and Write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 80 (4), 437-447.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Rose, J. (2006). *Independent Review of the Teaching of Early Reading*. London: Department for Education and Skills.

Rowe, K. (2005). *National Inquiry into the Teaching of Literacy (Australia), "Teaching Reading"* (2005). Retrieved 12th August, 2012 from http://research.acer.edu.au/tll_misc/5.

Slavin, R. E.; Lake, C.; Cheung, A. and Davis, S. (2009). Effective reading programs for the elementary grades. *Review of Educational Research,* 79 (4), 1391-1465.

Slavin, R., Woedworth, J. and Jones-Hill, M. (2005). Success for All in England: Implementation and outcomes of a comprehensive literacy reform for primary schools, MD: Success for All Foundation.

Torgreson, C. J., Brooks, G., and Hall, J. (2003). *A Systematic Review of the research Literature in the Teaching of Reading and Spelling.* Department of Education and Skills Research Report No. RR711. Sheffield: University of Sheffield.

Tymms, P. and Mercell, C. (2001). *Success for All Quantitative Report: Second Year.* Durham: CEM, Durham University.

Wilson, S. and Lipsey, M. (2007). School based intervention for aggressive and disruptive behaviour: Update on a Meta-Analysis. *Journal of Preventative Medicine*, 33 (3), 130-143

Timeline

	2013						2014						2015						2016					
	Jan-Feb	March- April	May- June-	July- Aug	Sept- Oct-	Nov- Dec	Jan-Feb	March- April	May- June-	July- Aug	Sept- Oct-	Nov- Dec	Jan-Feb	March- April	May- June-	July- Aug	Sept- Oct-	Nov- Dec	Jan- Feb	March- April	May- June-	July- Aug	Sept- Oct	Nov- Dec
1.EvaluationPlanning																								
Evaluation Design (CESI/IEE)																								
Recruitment of Schools (IEE)																								
Ethical Approval (IEE & QUB)																								
Randomisation of Schools (CESI)																								
2.Impact Evaluation																								
Administration of pre- test (IEE)																								
Administration of post- test (IEE)																								
Collection of KS results (IEE)																								
Data Entry (CESI)																								
Final Analysis and Report (CESI)																								
3.Process Evaluation																								
Process Design (CESI)																								
Survey administration (CESI)																								
Telephone Follow-up (CESI)																								
Survey Analysis (CESI)																								
In-depth Interviews and observation (CESI)																								