



E
E
F
ducation
ndowment
oundation

Zippy's Friends

Evaluation report and executive summary

June 2018

Independent evaluators:

Seaneen Sloan, Aideen Gildea, Sarah Miller, Allen Thurston



**QUEEN'S
UNIVERSITY
BELFAST**



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
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 evaluation is based at the Centre for Evidence and Social Innovation (CESI) at QUB and includes Dr Seaneen Sloan, Dr Aideen Gildea, Dr Sarah Miller, and Professor Allen Thurston.

Contents

Executive summary.....	4
Introduction	6
Methods	10
Impact evaluation	19
Process evaluation.....	29
Conclusion.....	43
References.....	46
Appendix A: EEF cost rating.....	49
Appendix B: Security classification of trial findings.....	50
Appendix C: Theory of change	52
Appendix D: Summary of previous evaluations of Zippy's Friends	53
Appendix E: Memorandum of Understanding for schools	56
Appendix E: Parent information and consent form.....	60
Appendix F: Randomisation procedure	63
Appendix G: Analysis code.....	65
Appendix H: Multilevel models	66
Appendix I: Implementation survey responses	68
Appendix J: Distribution of primary outcomes	70

Executive summary

The project

Zippy's Friends is a whole-class social and emotional learning (SEL) programme that aims to develop children's coping skills, including their ability to manage stressful situations. It is an internationally-used, well-developed programme aimed at class teachers of children in Years 1 and 2, run by the independent charity Partnership for Children.

There is indicative evidence that the social skills and self-regulation skills developed through Zippy's Friends could also improve children's academic attainment. This project aimed to test the impact of Zippy's Friends on academic attainment (in reading) and emotional self-regulation among Year 2 pupils.

Zippy's Friends consists of 24 weekly 45-minute sessions delivered over one school year. Six modules cover stories about a pet stick insect called Zippy and a group of young children who are his friends. Children are actively engaged in the stories through questioning, and learning is consolidated using games, activities, role-playing, and discussion. Zippy's Friends was delivered by class teachers during the 2016/2017 school year.

The evaluation involved 83 schools from five English local authorities, and 3,904 children. The evaluation was designed as an efficacy trial, with schools randomly allocated to an intervention group to teach Zippy's Friends or a control group to continue teaching as normal. Interviews, observations, and focus groups were conducted with teachers, local co-ordinators, pupils, and parents. A standard practice questionnaire was administered to all schools at the beginning and end of the trial. Intervention group teachers also completed module report forms (delivery logs) and implementation surveys.

Key conclusions

1. The project found no evidence that Zippy's Friends improved reading outcomes or emotional self-regulation for Year 2 pupils. These findings have low to moderate and low security, respectively.
2. There was no effect on reading attainment for pupils who had ever been eligible for free school meals. There was a small negative impact on emotional self-regulation for these pupils, but this result has lower security than the overall findings because of the smaller number of pupils.
3. Children receiving Zippy's Friends made small improvements in teacher-reported self-regulated learning, compared to other pupils.
4. The programme was very well received by teachers: all survey respondents felt Zippy's Friends had benefits for the children involved.
5. Evidence suggests that the impact of social and emotional learning programmes can take time to feed into academic outcomes. It is therefore recommended that academic outcomes are followed up at a later time point.

EEF security rating

These findings have a low to moderate security rating (low for the emotional self-regulation outcome). The trial was an efficacy trial, which tested whether the intervention worked under developer-led conditions in a number of schools. It was a well-designed, two-armed randomised controlled trial. There is evidence that schools randomised to the control condition increased their SEL provision over the course of the trial such that they delivered a similar amount of new provision to the intervention group. Intervention schools increased SEL from 41 to 65 minutes, whereas control group schools increased SEL from 37 to 55 minutes per week. This is a threat to the validity of the trial because it makes it harder to accurately estimate the full impact of Zippy's Friends compared to business as usual.

There were slightly fewer schools involved in the trial than necessary to reliably detect a small effect size (of under 0.2 standard deviations). 15% and 28% of the pupils who started the trial were not included in the final analysis of reading outcomes emotional self-regulation outcomes, respectively. This was because data was not available, generally due to pupils being absent, moving schools, or not completing all of the questions. The large amount of data missing for emotional self-regulation limits our confidence in the findings, particularly as the data was more likely to be missing for pupils with poorer performance on pre-test measures, and pupils eligible for free school meals ('FSM pupils') or with EAL status (English as an additional language).

There were some important differences in prior attainment between the pupils in Zippy's Friends schools and those in comparison schools, with the former starting with slightly lower reading scores. The pupils were similar in terms of their emotional self-regulation at the beginning of the trial.

Finally, the post-intervention measure of emotional self-regulation had some weaknesses for use in this trial: it was not designed for children quite as young as those in the trial, and it was designed to measure anger management, which is a narrower outcome than the intervention aims to affect.

Additional findings

There was no evidence to suggest an effect of the programme on the primary outcomes of reading attainment or emotional self-regulation. There was an improvement in self-regulated learning, but this was measured by teachers, and so is not as reliable as improvement on an objective measure. There was no evidence to suggest an impact on social skills.



There did not appear to be any evidence to suggest that the programme has differential effects for pupils eligible for FSM pupils, or that outcomes were better where the programme was delivered more completely, however these results need to be interpreted with caution due to smaller sample sizes for the former and incomplete data on programme implementation in the case of the latter.

The process evaluation found that the programme was very well received by teachers: 100% of those surveyed agreed the programme was beneficial for the children and 100% of those interviewed said they had noticed benefits for the children in their class.

Cost

Zippy's Friends costs approximately £10 per pupil per year over three years, based on an average class size of 27 pupils. This includes the cost of training for one teacher and one full set of teacher and pupil resources, but does not include the cost of any supply teacher cover, teacher travel expenses, any additional materials needed for activities, or the cost of an optional class trip.

Table 1: Summary of impact on primary outcome

Outcome (Group)	Effect size (95% confidence interval)	Estimated months' progress	p-value	No. of pupils	EEF security rating	EEF cost rating
Reading attainment	-0.02 (-0.08, 0.05)	0	0.54	3312		£££££
Emotional self-regulation	-0.02 (-0.09, 0.06)	0	0.65	2810		£££££
Reading attainment (FSM)	0.03 (-0.16, 0.21)	0	0.61	467	N/A	£££££
Emotional self-regulation (FSM)	-0.12 (-0.32, 0.08)	-2	0.21	294	N/A	£££££

Introduction

Intervention

Zippy's Friends is a universal, classroom-based social and emotional learning (SEL) programme that aims to develop children's repertoire of coping skills and their ability to adapt those coping skills to various situations. It is designed to be delivered to children aged five to seven by the class teacher over the course of an academic year through 24 weekly sessions, each lasting around 45 minutes. The 24 sessions are divided into six modules which focus on specific themes: feelings, communication, making and breaking relationships, conflict resolution, dealing with change and loss, and general coping skills. Weekly sessions are centred around a set of illustrated stories about a group of children, their families, friends, and Zippy—a pet stick insect. Pupils also actively participate in activities such as role-playing, group discussions, drawing and crafts, and playing games. Through listening to the stories and joining in activities, it is proposed that children learn how to find their own solutions to problems and, in turn, become better equipped to cope with difficulties in life and get on better with others.

Teachers undergo one day of training and are provided with a set of teaching and children's materials. Three follow-up, twilight support sessions are held after modules two, four, and six which involve teachers from the same area being brought together to discuss experiences, address any questions, and look at the forthcoming modules.

Zippy's Friends is underpinned by conceptual frameworks of coping in which individual attributes such as coping skills can reduce the impact of environmental stressors (Segal, 1983; Lazarus and Folkman, 1984; Sandler *et al.*, 2000). In this context, we define 'coping skills' as the skills or strategies used to manage needs and emotions caused by stressful situations or experiences (Lazarus and Folkman, 1984). The programme recognises that a variety of coping strategies, including problem-focused coping and emotion-focused coping, can be helpful, and that good coping skills can vary between individuals and situations. Thus, it aims to develop children's repertoire of coping skills and their ability to adapt those coping skills to various situations. Further, it integrates problem-solving skills, social skills, and emotional literacy as skills that may facilitate adaptive coping behaviour. Such competencies may be linked to academic engagement and achievement (Durlak *et al.*, 2011; Qualter *et al.*, 2012).

The central aim of this study is to test whether Zippy's Friends has an impact on children's academic attainment. The logic model (Appendix C)—which was developed during project set-up meetings between the programme delivery team, the evaluation team, and the funder—depicts the proposed pathway through which it is theorised that the programme will impact on academic attainment. It is proposed that the programme will have an indirect impact on academic attainment through a pathway that involves increased emotional self-regulation (the control and inhibition of emotions and emotional reactions) through which an improved focus on learning (self-regulated learning) is achieved (although it is acknowledged that impacts on academic attainment may not be achieved within a year).

Background evidence

An estimated 10% of children and young people aged between 5 and 16 years in the U.K. have a clinically diagnosed mental disorder (Green *et al.*, 2005), with an estimated prevalence of 13.4% worldwide (Polanczyk *et al.*, 2015). There is a substantial body of evidence that links early socio-emotional development to later academic performance and a number of key health, social, and economic outcomes such as stress and mental health, workplace readiness, and adult wellbeing (Chien *et al.*, 2012; Durlak *et al.*, 2011; Guzman *et al.*, 2014). Prevention and early intervention in children's emotional wellbeing through universal, school-based programmes may therefore constitute an effective—as well as cost-effective—option for improving a range of short- and long-term outcomes in the social, emotional, academic, and economic domains (Allen, 2011; Knapp, 2011).

Educators are increasingly recognising the importance of SEL skills for effective learning (Collaborative for Academic, Social, and Emotional Learning 2015; DoH, 2012; DfE, 2015). Within this context, schools play a crucial part in contributing to the positive mental health of children and young people. Several reviews have concluded that universal interventions promoting SEL competencies can have significant positive outcomes for children's mental health, wellbeing and educational attainment (Browne et al., 2004; Adi et al., 2007; Tennant et al., 2007; Greenberg, 2010; Durlak et al., 2011; Clarke et al., 2015). A recent review (Clarke et al., 2015) identified 39 school-based social and emotional skills interventions in use in the U.K. and a survey of 880 U.K. schools (conducted in 2016) found that 41% were delivering dedicated 'character education' lessons with their pupils (Marshall et al., 2017).

Zippy's Friends has been evaluated before; a summary table of the relevant studies and impacts found can be consulted in Appendix D. Zippy's Friends was first piloted in Denmark and Lithuania, using a quasi-experimental design (Mishara and Ystgaard, 2006) with pupils aged six to eight years, in which the control group were matched to the intervention group based on age, gender, and socio-economic background. This study found improvements in pupil-reported self-control, and in several teacher-reported outcomes—including use of coping strategies, co-operation, empathy, assertion, and self-control. Also, the problematic behaviours of externalizing and hyperactivity decreased. The study, however, had major methodological limitations including no random allocation, no description of how the control and intervention groups were matched, a lack of reporting of baseline equivalence on sample characteristics or outcomes at pre-test, a lack of blinding of outcome assessors (for the Denmark study), and no adjustment for clustering in analysis. These limitations are likely to have biased the findings towards more positive effects.

More recently, Zippy's Friends has been evaluated using cluster randomised controlled trials in Norway (Holen et al., 2012; 2013), Ireland (Clarke et al., 2014) and the Netherlands (Hegeman, 2015). Holen et al. (2012) found a positive impact of the programme at immediate post-test in terms of reducing self-reported negative coping strategies among seven- and eight-year-olds, increasing parent-reported active coping strategies, and self-awareness. There was no effect in terms of mental health outcomes, as measured by both parent and teacher reports on the Strengths and Difficulties Questionnaire (SDQ), although teachers reported reduced impact of mental health difficulties on children and their environment. Teachers also reported improvements to aspects of the broader class climate (Holen et al., 2013) in terms of reduction in bullying, improved social integration, and improved academic skills, although these findings were based on teacher reports and thus are subject to bias. Clarke et al. (2014) tested the programme with 766 children in 48 schools in disadvantaged areas in Ireland and found a significant impact on several (teacher-reported) emotional literacy subscales, including social skills, self-regulation, motivation, and self-awareness. Finally, the RCT by Hegeman (2015), involving 1,177 six- to seven-year-olds from 29 primary schools in the Netherlands, found a positive effect on self-reported emotion recognition, self-awareness, and adaptive coping, as well as parent-reported social and emotional skills and externalising behaviour. This study found no effect on teacher-reported outcomes of problem behaviour, emotional literacy (including social skills, empathy, motivation, and self-regulation), or attainment in literacy and numeracy.

The current evaluation was conducted as an efficacy trial and represents the first large-scale evaluation of the programme in a U.K. context and is the first study to examine the direct impact of the programme on standardised academic test scores.

Evaluation objectives

The primary aim of the evaluation was to determine the impact of Zippy's Friends on academic attainment and emotional self-regulation. Theoretically (and as proposed in the logic model, Appendix C), the programme may impact on social and emotional competencies that may facilitate improvements to pupil learning behaviour (such as concentration and perseverance with school work or co-operation with class activities) and the broader learning environment, which may then lead to improvements in

academic attainment. The evaluation addressed the following research questions, as per the evaluation protocol:¹

Impact evaluation:

1. What is the overall effectiveness of the programme on:
 - pupil reading attainment and emotional self-regulation (primary outcomes)?
 - self-regulated learning and social skills (secondary outcomes)?
2. Is there a differential impact of the programme for children eligible for free school meals?

Implementation and process evaluation:

3. How is the programme perceived by schools (in terms of engagement of pupils, teachers, and local co-ordinators), and what are the barriers and facilitators to implementation?
4. Is any variability in implementation associated with variability in outcomes?
5. Do the proposed mechanisms as depicted by the logic model (Appendix C) explain any link between the programme and academic attainment?

The first and second research questions are addressed through the main and subgroup analyses which are presented within the Impact Evaluation section. The third research question is addressed throughout the Process Evaluation section. It was intended to address the fourth research question through on-treatment analysis, however (and as described within the Impact evaluation section), this analysis was limited due to the extent of missing implementation data. The fifth research question is addressed within the Process Evaluation (Outcomes) section.

Ethical review

The trial was approved by the Research Ethics Committee, School of Education, Queen's University Belfast (QUB) in January 2016. School-level consent to participate in the trial was sought from headteachers who were asked to sign a memorandum of understanding (MoU, Appendix D). Informed consent for pupil participation in trial data collection was sought from parents or guardians on an opt-out basis (Appendix E). Schools informed the evaluators of any opt-out parental consent forms received. Twenty-four parental opt-out consent forms were returned in total. These pupils still received the programme (if the school was allocated to the intervention group), however did not take part in evaluation data collection. During whole-class testing, any opted-out pupils were provided with other activities to complete by their teacher.

Project team

The programme delivery team consisted of Caroline Egar (Programme Director) and Lizzie Poulton (Programme Manager), based at Partnership for Children (PfC), a U.K. charity which supports the delivery of Zippy's Friends nationally and internationally. They were responsible for the recruitment of schools through local authority areas and were supported by a local co-ordinator within each area. The local co-ordinator was generally a professional working in a school or with a remit for working in schools (for example, an educational psychologist). For the trial, local co-ordinators were recruited for the purpose of engaging with schools in the local area, gauging interest in the programme and—once the intervention commences—monitoring fidelity of delivery and providing ongoing support.

The evaluation team were responsible for the design, management, analysis, and reporting of the evaluation. The evaluation is based at the Centre for Evidence and Social Innovation (CESI) at QUB and includes Dr Seaneen Sloan, Dr Aideen Gildea, Dr Sarah Miller, and Professor Allen Thurston.

Trial registration

The trial was prospectively registered with the ISRCTN Registry (ref: ISRCTN82558815).

Methods

Trial design

The evaluation was conducted with Year 2 classes in primary schools in five local authorities in England during the 2016/2017 school year. The impact evaluation was designed as a cluster randomised controlled trial involving two arms (an intervention and control group). The intervention arm received the programme and the control arm was asked to continue as normal and did not receive the programme.

The unit of randomisation was the school. This was considered appropriate over randomisation at the individual pupil level because the intervention is designed to be delivered on a whole-class basis. Further, randomisation within schools (at the class or year-group level) would have increased the risk of contamination between trial arms.

Schools allocated to the control group received a payment of £1,000 at the end of the trial, with the option to purchase the programme if they wished.

There was one change made to the original trial design published in the protocol. This was in relation to the measure of emotional self-regulation used at post-test—the Children's Anger Management Scale (CAMS). It was originally intended to use the pupil-report version of the pre-test measure (that is, the self-regulation subscale of the Emotional Literacy Checklist), which was based on teacher-report at pre-test. However, scrutiny of the literature on the psychometric properties highlighted that the pupil-report version of the self-regulation subscale has poor reliability (Cronbach's alpha = 0.52; Faupel, 2003), and thus the CAMS was selected as an alternative which has shown good reliability when used with young children.

Participant selection

Zippy's Friends is designed to be delivered with pupils aged five to seven years, and the target cohort was identified as Year 2 pupils (aged six to seven years).

State funded infant and primary schools in five local authority areas (Nottingham, Lincolnshire, Dudley, Croydon, and Cheshire) were eligible to participate. These local authorities were selected as there was sufficient interest from schools in these areas during scoping work by the programme delivery organisation during the trial set-up period. Schools were eligible to take part in the trial if they:

- had not delivered Zippy's Friends before;²
- were prepared to allow one hour per week in the Year 2 timetable for Zippy's Friends; and
- were prepared to release Year 2 teachers for training and support sessions.

Recruitment of local authorities and schools took place between January and May 2016 and was led by the programme developer. Within each local authority, a local co-ordinator was employed to assist with recruiting schools to the study and then to monitor fidelity and provide ongoing support to schools in the delivery of the programme. The local co-ordinator was generally a professional working in a school or with a remit for working in schools.

Headteachers in schools who expressed interest in participating in the trial were given the MoU (Appendix D), which explained what participation in the trial would involve for the school. They were asked to read, sign, and return this to the local co-ordinator who (through the programme developer) informed the evaluation team when a signed MoU was received.

²

The evaluator then sent each school a 'pre-test pack' containing the following:

- parental information sheets and opt-out consent forms (Appendix E);
- pre-test assessments; and
- an encrypted spreadsheet for returning pupil details; this included details required for linking to the National Pupil Database (UPN, pupil first name, last name, and date of birth), as well as EAL status.

The target cohort within each school was pupils due to enter Year 2 in the 2016/2017 school year. All pupils in the target cohort were eligible to participate in the trial, provided their parent had not opted out.

Schools were informed that, in order to be eligible to go forward for randomisation (and thus be included in the trial) they needed to complete the following steps:

1. ensure parental information sheets and opt-out consent forms were sent home to parents of the target year group;
2. ensure the pre-test assessments were completed for the target year group and returned to the evaluation team; and
3. ensure that the relevant pupil details had been securely shared with the evaluator.

Outcomes measures

Table 2 summarises the pre- and post-test outcomes, measures, informant, and use in previous evaluations.

Table 2: Summary of outcome measures at pre- and post-test

Outcome	Measure	Informant	Use in previous evaluations
Pre-test			
Reading	Hodder Oral Reading Test	Pupil	No
Emotional self-regulation	Emotional Literacy Checklist – self-regulation subscale	Teacher	Yes
Self-regulated learning	Learning Behaviour Scale	Teacher	No
Social skills	Social Skills Information System	Teacher	No
Post-test			
Reading	Hodder Group Reading Test	Pupil	No
Emotional self-regulation	Children's Anger Management Scale (CAMS)	Pupil	No
Self-regulated learning	Learning Behaviour Scale	Teacher	No
Social skills	Social Skills Information System	Pupil	No

Primary outcomes

The two primary outcomes were reading attainment and emotional self-regulation.

Reading attainment was measured at post-test (June 2017) using the paper version of the Hodder Group Reading Test (HGRT; Vincent and Crumpler, 2007), which assesses reading comprehension at word, sentence, and text levels. This was administered on a whole-class basis under test conditions by independent fieldworkers who were blind to school allocation status and trained, overseen, and monitored by the evaluation team. Completed test papers were packaged and returned by the

fieldworker on the day of testing directly to Hodder for marking, which again was blind to school allocation status.

At pre-test, reading attainment was measured using the Hodder Oral Reading Test (HORT; Vincent and Crumpler, 2006) which contains three separate tests measuring word reading, sentence reading, and reading speed. This was administered individually with each pupil by Year 1 class teachers (or teaching assistants) using standardised instructions. Pre-test data collection was completed between April and June 2016, prior to randomisation.

With a preference to measure emotional self-regulation based on pupil self-report, available instruments were reviewed for suitability for use with children aged six to seven years. The Children's Anger Management Scale (CAMS; Zeman *et al.*, 2002) was selected as a self-report measure of emotional self-regulation suitable for use at post-test. The CAMS was administered on a whole-class basis by fieldworkers (blind to allocation) on the day of HGRT administration. On completion of testing, questionnaires were returned by the fieldworker to the evaluation team for processing.

The CAMS contains 11 items covering three areas: inhibition of anger expression (four items, for example, 'I hold my anger in'); coping, or anger control (four items, such as, 'I try to calmly deal with what is making me feel mad'), and dysregulation of anger expression (three items, for example, 'I say mean things to others when I'm mad'). Each item is rated on a three-point scale (1, 'hardly ever'; 2, 'sometimes'; and 3, 'often'), and total scale scores can range from 11 to 33 (with higher scores reflecting greater emotional self-regulation). The CAMS has been used with children aged six years and over and has demonstrated good internal consistency (Cronbach's alpha = 0.62–0.77) as well as convergent and discriminate validity with measures of emotion, psychopathology, and social functioning (*et al.*, 2002; Suveg and Zeman, 2004; Suveg *et al.*, 2009). The scale was found to have acceptable reliability in this sample ($\alpha = 0.77$).

It was not feasible to measure self-reported emotional self-regulation at pre-test (end of Year 1) due to the lack of reliable self-report measures for this age group. Instead, teacher-report on the 'self-regulation' subscale of the Emotional Literacy Checklist (Faupeil, 2003; GL Assessment) was used, and was found to have good reliability in this sample ($\alpha = 0.82$). This contains four items which assess teacher perception of a child's ability to regulate their emotions, for example, 'loses temper when loses at a game or in a competition'. Each item is rated on a four-point scale (0, 'very true'; 1, 'somewhat true'; 2, 'not really true'; and 3, 'not at all true'). Total scores on this scale can range from 0 to 12, with higher scores reflecting greater emotional self-regulation.

Secondary outcomes

The two secondary outcomes were self-regulated learning and social skills.

Again, due to the lack of reliable self-report measures available for this age group, self-regulated learning was measured at both pre- and post-test using teacher report on two subscales ('attention/persistence' and the 'emotional control') of the Learning Behaviour Scale (McDermott *et al.*, 2001). These subscales contain 12 items in total (for example, 'doesn't stick to tasks'), each rated on a three-point scale (1, 'doesn't apply'; 2, 'sometimes applies'; and 3, 'always applies'). Each item was reverse scored and then summed to create a total scale score (with higher scores reflecting better self-regulated learning behaviours).

Social skills were measured at pre- and post-test using the social skills subscale of the Social Skills Improvement System rating scale (SSIS; Gresham and Elliott, 2008). Social skills were measured via teacher-report at pre-test, and then self-report at post-test (again, due to lack of reliable self-report measures for this age group). The SSIS contains 46 items reflecting the dimensions of social skills, including communication, co-operation, assertion, responsibility, empathy, engagement and self-control. Items are rated on a four-point scale (0, 'never'; 1, 'seldom'; 2, 'often'; and 3, 'always').

At post-test, the measure of social skills was administered on a whole-class basis by fieldworkers (blind to allocation) on the same day of the CAMS. Self-regulated learning was measured by teacher report. On completion of testing, both measures were returned by the fieldworker to the evaluation team for processing.

Sample size

The initial sample size calculation estimated that 70 schools with an average of 37 pupils per school (within the selected year group) would be sufficient to detect an effect size of 0.24 in an intention-to-treat analysis. This was calculated using Optimal Design (version 3.01) and based on a two-level cluster design with a conservative estimate of the pre-post correlation (r) of 0.60 ($R^2 = 0.36$) and the intra-cluster correlation coefficient (ρ) of 0.16 for the primary outcome (reading), and standard power and alpha thresholds of 0.80 and 0.05 respectively.

During the set-up and recruitment phase of the project, a high number of schools expressed interest in taking part in the trial, and with agreement from the funder, up to 90 schools were permitted to take part in the trial, which (using the same assumptions as above) would power the trial to detect an effect size of 0.18. Based on Department for Education statistics indicating that the average proportion of pupils eligible for and claiming FSM is 16.5% (DfE, 2015), it was estimated that there would be an average of six pupils eligible for FSM (based on an average year group of 37 pupils). Using the same assumptions as above, the minimum detectable effect size for FSM pupils was 0.33.

A total of 83 schools were recruited and eligible to be randomised. At the randomisation stage, MDES was recalculated for the increased number of schools recruited, and the larger size of schools recruited meant that the average number of pupils per year group increased to 45. Based on these assumptions, the trial was designed with an 80% power to detect an effect size of 0.22 (and 0.30 for FSM pupils; see Table 4).

Randomisation

The evaluation team conducted random allocation in July 2016 and communicated the outcome of this to schools before the end of the school year. As described above, schools needed to have gained parental (opt-out) consent for Year 2 pupils, returned the completed pre-test measures, and provided the evaluator with data on participating pupils to be eligible to be included in the random allocation process.

Schools were allocated to groups using the program Minim (Evans et al., 2004) to create groups that are balanced across pre-specified characteristics (see Appendix F). This method of allocation is known as 'minimization' and is a widely accepted alternative to simple or stratified randomization (Altman and Bland, 2005). Minimization used two characteristics: local authority and school-level attainment. This was used to ensure that schools allocated to each trial arm were broadly equivalent in terms of school-level attainment, and also to ensure that equal numbers of schools within each of the five local authority areas were allocated to each group (to avoid the scenario whereby one area has very few intervention schools and another has many).

School-level attainment was also used to ensure a similar proportion of high, medium and low attaining schools in the two treatment groups. School-level attainment was based on percentage of pupils obtaining level 4 or above in reading, writing, and maths at Key Stage 2. This data was collected from DfE school performance tables published in December 2015 (using the variable PTREADWRITTAMATX). KS2 data was unavailable for four schools (three of these were infant schools, and one was a newly established primary school that had yet to have pupils sit KS2 tests). The corresponding Key Stage 1 results obtained directly from schools were used in these four cases.

School-level attainment was split into three equally sized categories using internally-derived cut-offs. This involved identifying cut off points such that three equal groups were created representing the top, middle, and bottom third attaining schools within the sample. Cut-offs for the percentage of pupils obtaining level 4 or above in reading, writing, and maths at KS2 derived for the current sample were as follows:

- low attainment: $\leq 76\%$;
- medium attainment: $\geq 77\%$ and $<86\%$; and
- high attainment: $\geq 86\%$.

A total of 83 schools were entered into the random allocation process. Of these, 41 were allocated to the control group and 42 were allocated to the intervention group.

Analysis

Analysis was conducted in Stata 14 (StataCorp, 2015). The extent to which the control and intervention groups were balanced at baseline was assessed by comparing school-level characteristics (school type, location, Ofsted rating, school size and attainment, proportion of EAL pupils, proportion of FSM pupils, and proportion with statements of special educational needs or education and health care plans). Pupils in the intervention and control group were also compared on the basis of gender, FSM, EAL, and pre-test scores on the primary and secondary outcomes. Linear and logistic multilevel regression models were used to assess for imbalance between the groups, accounting for the clustered nature of the data. Means and standard deviations of these characteristics are presented in tabular format, disaggregated by control and intervention group. Imbalance at baseline attainment is reported as an effect size.

The main effects of the intervention were estimated on an intention-to-treat basis, using multilevel modelling to take account of the clustered nature of the data (see Appendix G for analysis code). A simple analysis was conducted for each of the primary and secondary outcomes, with the raw post-test score as the dependent variable and two independent variables—a dummy variable representing treatment allocation (intervention group coded '1' and control group coded '0'), and the standardised pre-test score for that outcome.

The coefficient associated with the dummy variable for treatment allocation, representing the difference between the intervention and the control group (controlling for pre-test score), was then used to estimate the effect size (Hedges' *g*) of the programme for each outcome. All effect size calculations used the total variance from multilevel models (without covariates) as the denominator.

One pre-specified subgroup analysis was conducted to determine whether the programme worked differently for FSM pupils. This involved repeating the main analyses for children who were identified as 'ever eligible for FSM' (using the EVERFSM_ALL variable obtained from the NPD).

The proportion of missing outcome data was examined (by treatment allocation) and found to be over 5% in both the intervention and control group for all outcomes at post-test (12% for reading attainment; 26% for emotional self-regulation; 17% for self-regulated learning; and 39% for social skills—see Table 9 for further details in outcome attrition). Multiple imputation was conducted as a sensitivity analysis. The imputation model imputed data separately for the control and intervention groups and included all variables used in the analysis and the two variables used in randomisation (school-level attainment and area). The imputation was performed using chained equations; this fills in missing values in multiple variables iteratively by using a sequence of univariate imputation methods with fully conditional specification of prediction equations. This method accommodates arbitrary missing-value patterns. Twenty imputations were conducted in order to lessen the simulation ('Monte Carlo') error. The analysis using the imputed datasets was then compared to the complete case analysis (reported in the results section).

The on-treatment analysis involved repeating the primary and secondary outcome analyses using a dosage variable (see Implementation and Process Evaluation below) in place of the intervention group variable.

Implementation and process evaluation

The implementation and process evaluation (IPE) was undertaken alongside the trial to support understanding of the trial findings, to understand control group activity, and to explore four main programme implementation and fidelity issues:

- adherence to the programme manual;
- levels of exposure to the programme (dosage);
- quality of the programme delivery; and
- teacher and pupil engagement with the programme.

Data collection took place between May 2016 and July 2017 using the data collection methods described below. For the most part, IPE data was collected directly by the evaluator, with the exception of the implementation logs.

Observation of training and support sessions

The evaluator observed a one-day training session for teachers delivered by the programme delivery team in October 2016. The evaluator also attended a sample of follow up support meetings (n = 5), convened by the local co-ordinator in each area. Field notes were made by the evaluator during all observations.

Implementation logs

Module report forms (designed by the programme delivery team with input from the evaluator) were designed to be completed by all intervention teachers with the aim of recording information on programme delivery and fidelity. This included:

- the date and duration of each session and whether each session was delivered in the regular slot (and if not, why not);
- whether the rules were included and the extent to which the activities in the folder were followed;
- the use of the Inclusion Supplement and the Activities for Tablets supplement; and
- the teacher's perception of children's enjoyment of the session.

These forms were supplied to teachers by the programme developer to complete after delivery of each of the 24 sessions. Teachers were asked to return module reports to their local co-ordinator at the support sessions, and copies of these were then provided to the evaluator.

Module reports were also intended as the basis for measuring dosage, with each school being assigned a score computed from the number of sessions delivered and the extent to which the activities outlined in the manual were followed. Scoring was allocated as follows:

1. session not delivered (0);
2. session delivered but activities not followed (1);
3. session delivered and some activities followed (2); and
4. session delivered and all activities followed (3).

Thus, across the 24 sessions, scores could range from 0–72, with higher scores reflecting a greater dosage. In practice, many teachers did not complete and return the module report forms, generally due to lack of time (discussed further under Process Evaluation findings). For example, 59% of teachers (46 of 78) in the intervention group returned the report for Module 1, however this figure fell progressively during the course of the year: by Module 6, only 13% of teachers (10 of 78) returned the forms. Over a third of teachers (36%; n = 28) returned no forms at all, and only three returned all six. An implication of this is that the estimate of dosage derived from the module reports is subject to bias.

Observations of programme delivery

The evaluator observed a Zippy's Friends session during site visits to six schools—selected to include a mix of urban/rural location, size, and level of deprivation.

Interviews and focus groups

Interviews and focus groups were undertaken by the evaluator with a range of relevant stakeholders (described below). All such meetings were audio-recorded (with participant consent) and transcribed verbatim. Qualitative data was analysed using thematic analysis (Braun and Clarke, 2006).

Local co-ordinators

Semi-structured interviews (two per co-ordinator) with each of the five local co-ordinators took place towards the beginning and end the programme delivery period. The majority were telephone interviews; these focused on co-ordinators' experiences of supporting schools with the delivery of the programme and any issues that had arisen within schools (and how these issues were dealt with).

Teachers

A total of 16 teachers participated in semi-structured interviews or a focus groups with the evaluator. These took place over the course of the intervention delivery period—some during site visits, others by telephone. Data was collected on teacher perceptions of the programme's impact: what was working well (or less well) in terms of delivery, the perceived 'costs' (time, resource, adverse consequences), the value added to the school (in terms of wider benefits such as school ethos or teacher-pupil relations), and links between the programme content and the general curriculum.

Headteachers

Semi-structured telephone interviews took place with two headteachers at the end of the intervention delivery period.

Parents

There were semi-structured telephone interviews with eight parents from six schools at the end of the intervention delivery period. These focused on parental awareness of the programme and the perceived impact that the programme was having on their child.

Pupils

Seven focus groups took place with a total of 48 pupils from six schools over the course of the intervention period during site visits to schools. During these, pupils were asked to describe Zippy's Friends and what they had been learning during sessions. They were asked what they liked or disliked about the programme and, if it had helped them, they were asked for examples of how it had helped.

Implementation survey

An online survey was administered to all intervention teachers ($n = 78$) towards the end of the programme implementation (with the first invitation issued at the end of May 2017, and reminders sent in June and July). This contained 32 statements (see Appendix H) which participants responded to on a six-point scale (ranging from 'strongly agree' through to 'strongly disagree') with the aim of collecting information on teachers' perceptions of the programme, the training and support received, the materials and resources, and the benefits associated with the programme. Teachers were also asked to report on how much time was spent each week preparing for and delivering sessions, and about additional resource requirements. The survey response rate was 59% (46 out of 78 teachers).

Control group activity

A baseline (pre-randomisation) questionnaire was used to capture data on standard practice in relation to social and emotional learning in participating schools. At the end of the school year, Year 2 teachers in the control group were also asked to complete a questionnaire on actual practice in terms of PSHE (personal, social, health, and economic education) and SEL during the school year. This included information on how much time per week was spent on SEL activities with their class, and details of any specific SEL programmes or interventions used. This was completed by 50 teachers in 36 of the 41 control group schools. Thus, it was possible to explore any change in practice in the control group following group allocation.

Costs

Cost data was collected from the programme developer. Specifically, this included the cost of training and programme materials for delivery of Zippy's Friends with one class for one school year. Cost per pupil was estimated by dividing the cost per class by the average number of pupils per class in the study ($n = 27$) in line with EEF guidance. Cost per pupil was estimated over a three-year period.

Additional data was collected from intervention schools via the survey administered at the end of the school year and completed by 59% of intervention group teachers. Here, teachers were asked to estimate the amount of time they spent preparing for and delivering the programme each week, and were also asked to describe any additional resources or materials needed to deliver the programme, over and above those provided by the programme developer.

Timeline

Table 3: Timeline

Date	Activity
Preparation for programme delivery and evaluation set-up	
January–May 2016	Recruitment of schools (PFC)
January 2016	Ethical approval (CESI)
July 2016	Randomisation of schools (CESI)
September–October 2017	Teachers trained (PFC)
Impact evaluation	
April–July 2016	Pre-testing (CESI)
September–June 2017	Intervention delivery (PFC)
June 2017	Post-testing (CESI)
July–October 2017	Data entry, cleaning, and analysis (CESI)
Process evaluation	
May–June 2016	Standard practice pro-forma (CESI)
September 2016	Teacher training observed (CESI)
October 2016–June 2017	Observations of programme delivery (CESI)
November 2016–July 2017	Interviews with local co-ordinators and intervention teachers (CESI)
January–June 2017	Focus groups with pupils in intervention schools (CESI)
October 2016–June 2017	Programme delivery logs (PFC)
May–July 2017	Administration of survey (CESI)
Report	
November 2017	Final report (CESI)

Impact evaluation

Participants

Figure 1: Participant flow diagram

Note: HGRT = Hodder Group Reading Test (primary outcome); CAMS = Children's Anger Management Scale (primary outcome).

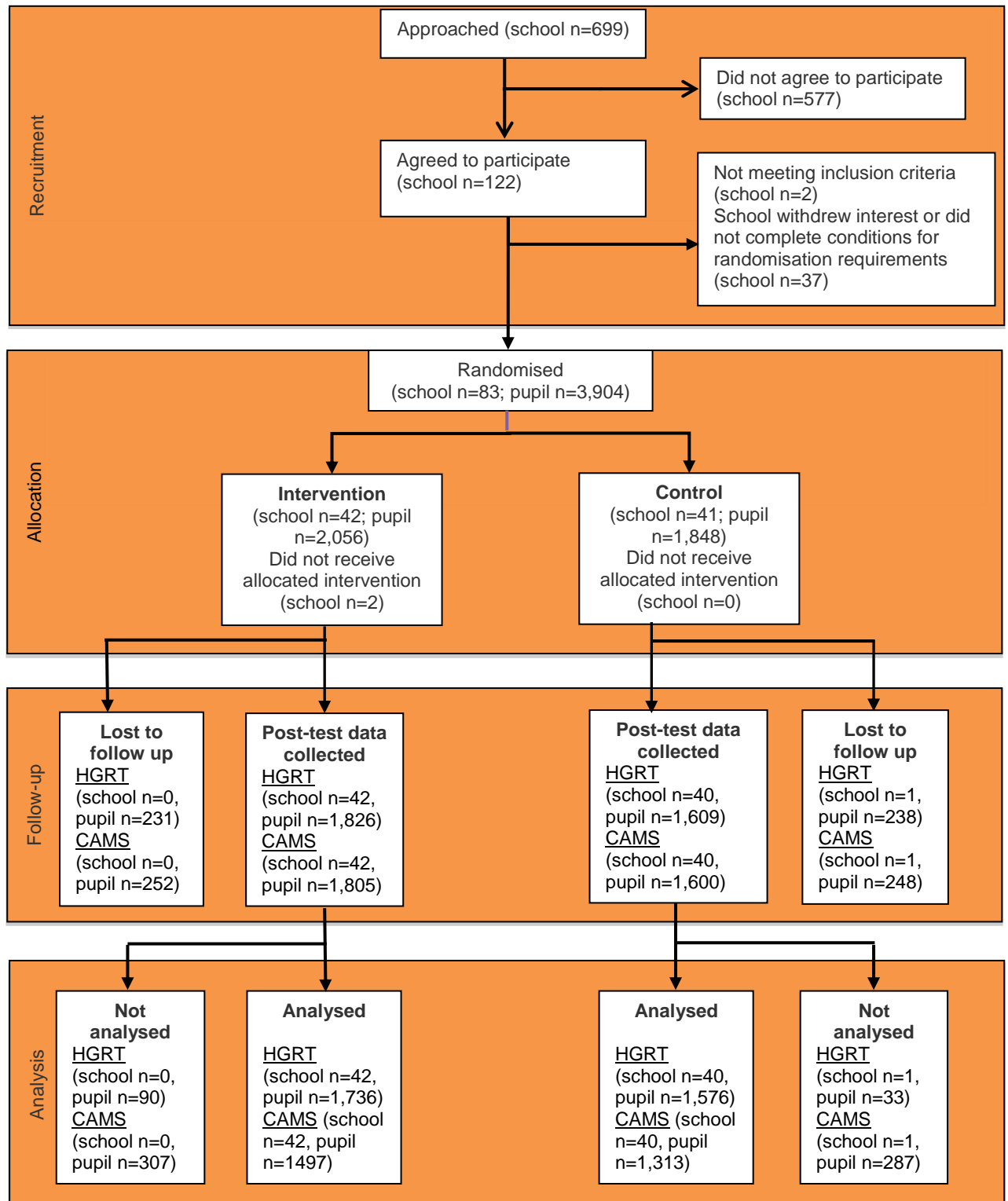


Figure 1 displays the flow of schools and pupils through the trial and for both primary outcomes. Pupils who were not analysed are those who were tested at post-test but not able to be included in the main analysis because of missing data at pre-test for that outcome. For example, of the 1,826 pupils in the intervention group who completed the reading test at post-test, 90 of these pupils were not included in analysis because of missing pre-test reading scores.

Attrition, measured as the drop-out from the initial sample at randomisation to the point of analysis, was 15% for reading and 28% for emotional self-regulation.

Table 4: Minimum detectable effect size at different stages

Stage	N [schools] (n=intervention; n=control)	Correlation between pre- test (+other covariates) and post-test	ICC	Blocking/ stratification or pair matching	Power	Alpha	Minimum detectable effect size (MDES)
Protocol	70 (35; 35)	0.60	0.16	Minimisation (on local authority and KS2 attainment)	80%	0.05	0.24 ^a
Randomisation	83 (42; 41)	0.60	0.16	Minimisation (on local authority and KS2 attainment)	80%	0.05	0.22
Analysis	82 (42; 40)	0.74	0.07	Minimisation (on local authority and KS2 attainment)	80%	0.05	0.15

^a MDES at the protocol stage here differs to published protocol (where it was 0.20) due to an error made at the protocol stage, where the estimated pre-post test correlation of 0.60 was entered as a variance

At the analysis stage, the minimum detectable effect size was calculated using the final ICC (0.07) and the pre-post correlation (0.74) with 42 schools and an average of 41 pupils per school. The trial achieved power to detect an effect size of 0.15 (see Table 4).

Eighty-three schools were recruited to the study and randomly allocated to the intervention group (n = 42) or the control group (n = 41). After randomisation, two schools in the intervention group withdrew and did not implement Zippy's Friends, however both remained in the trial on an intention-to-treat basis. One of these schools cited not having time to fit in the large number of initiatives in place in the school at the time; the other was unable to make the training. One school from the control group withdrew from the trial and did not take part in post-test data collection. Attrition at the pupil level varied across outcomes. For reading, 15% of pupils were not included in the main analysis due to missing data at post-test (n = 469; this includes 61 pupils from the one school that withdrew) or pre-test (n = 123). There was no evidence of differential attrition between the groups (in the intervention group it was 15.6%; control 14.7%; p = 0.61). For emotional self-regulation, attrition was higher at 28%. This includes participants who did not complete some or all of the measure at post-test (n = 960) or pre-test (n = 134). There was no evidence of differential attrition between the groups (intervention: 27.2%; control 29.0%; p=0.93).

School and pupil characteristics

On a range of school-level characteristics, the intervention and control groups were broadly similar (see Table 5), with no significant differences (95% confidence) between the group on school type, school location, Ofsted rating, or school-level attainment. Approximately two thirds of schools in both the control and intervention groups were local authority maintained. The majority, overall, were based in urban settings. In terms of Ofsted rating, about a third of schools in both groups were rated 'good'.

Approximately a third of schools in each group fell into the 'high', 'medium' and 'low' categories for school-level attainment (which was expected as this was a factor used in minimisation).

Intervention group schools were slightly larger in terms of total number of pupils enrolled, and had a slightly larger percentage of FSM pupils and pupils with SEN or EHC plans, however none of these differences were statistically significant (see Table 5). Analysis of baseline school-level characteristics suggests that randomisation was successful in creating two groups that were broadly equivalent.

Of these schools, 82 took part in post-testing, with one control school withdrawing after randomisation.

Table 5: Baseline comparison of school-level characteristics

Variable	Intervention group		Control group		Test statistic (p-value)	Effect size (g)
	n/N*	%	n/N*	%		
School-level (categorical)						
School type						
Academy (Converter and Sponsor-led)	14/42	33	14/41	34	Chi ² = 0.01 (p = 0.94)	0.02
Local authority maintained schools	28/42	67	27/41	66		
Academy Converter	10/42	24	11/41	27		
Academy Sponsor-led	4/42	10	3/41	7		
Community School	16/42	38	17/41	41		
Foundation School	1/42	2	3/41	7		
Voluntary Aided School	8/42	19	4/41	10		
Voluntary Controlled School	3/42	7	3/41	7		
Location					Chi ² = 0.91 (p = 0.34)	0.21
Urban	37/42	88	33/41	80		
Rural	5/42	12	8/41	20		
Ofsted rating					Chi ² = 0.72 (p = 0.70)	0.17
Outstanding	6/42	14	9/41	22		
Good	28/42	67	27/41	66		
Requires improvement	5/42	12	4/41	10		
No rating	3/42	7	1/41	2		
Attainment (% attaining level 4 or above in reading, writing and maths at KS2)					Chi ² = 0.10 (p = 0.95)	< 0.001
High (≥ 86%)	15/42	36	13/41	32		
Medium (≥ 77% and < 86%)	13/42	31	14/41	34		
Low (≤ 76%)	14/42	33	14/41	34		
School-level (continuous)	n (missing)	Mean	n (missing)	Mean	Test statistic (p-value)	Effect size (g)
School size						
Total number of pupils enrolled	42 (0)	327	41 (0)	319	t = -0.20 (p = 0.84)	-0.04
% FSM	42 (0)	18%	41 (0)	15%	t = -1.14 (p = 0.26)	-0.25
% EAL	42 (0)	17%	41 (0)	14%	t = -0.74 (p = 0.46)	-0.16
% SEN	29 (13)	2%	29 (12)	2%	t = -0.31 (p = 0.76)	-0.08

* n/N refers to the number of schools within each category (n) out of the total number of schools in each group (N).

FSM = eligible for free school meals; EAL = English as an additional language; SEN = pupils with special educational needs or EHC (education, health and care) plans.

The trial involved 3,904 Year 2 pupils—2,056 (53%) in the intervention group and 1,848 (47%) in the control group. The intervention and control groups were compared using multilevel linear and logistic regression modelling on a number of pupil-level sample characteristics (see Table 6), including gender, FSM eligibility, and EAL status. In both groups, approximately half of pupils were male. The percentage of FSM pupils was slightly higher in the intervention group (17%) compared to the control group (13%), however this difference was not statistically significant. There was also a higher proportion of missing data for this variable within the intervention group, however this was not statistically significant (p = 0.14). The percentage of the sample with EAL was also slightly higher in the intervention group (21%) compared to the control group (16%), but again, this was not a statistically significant difference.

Pre-test scores on each of the four outcomes (reading, emotional self-regulation, self-regulated learning, and social skills) were compared between the intervention and control groups. Differences between the two groups were small and not statistically significant (Table 6), however the effect sizes for the difference between the groups for reading attainment and self-regulated learning suggested there were some differences between the groups, with the intervention group scoring lower on both reading attainment and self-regulated learning.

Table 6: Baseline comparison of pupil-level characteristics

Variable	Intervention group		Control group		Test statistic (p-value)	Effect size (g)
	n/N (missing)	Percent	n/N (missing)	Percent		
Pupil-level (categorical)						
Male	1,029/2,056 (53)	50	964/1,848 (21)	52	B = -0.06 (p = 0.40)	-0.05
Eligible for FSM	350/2,056 (75)	17	245/1,848 (39)	13	B = 0.29 (p = 0.19)	-0.11
English as an Additional language	433/2,056 (4)	21	289/1,848 (3)	16	B = 0.50 (p = 0.22)	-0.14
Pupil-level (continuous)						
	n (missing)	Mean (SD)	n (missing)	Mean (SD)	Test statistic (p-value)	Effect size (g)
Reading attainment (pre-test)	1,936 (120)	68.67 (31.54)	1,803 (45)	71.20 (31.21)	B = -1.84 (p = 0.36)	0.08
Emotional self-regulation (pre-test)	1,998 (58)	9.52 (3.19)	1,772 (76)	9.54 (3.21)	B = -0.01 (p = 0.97)	0.01
Self-regulated learning (pre-test)	1,961 (95)	31.72 (5.06)	1,747 (101)	32.05 (4.72)	B = -0.31 (p = 0.42)	0.07
Social skills (pre-test)	1,851 (205)	104.01 (24.33)	1,664 (184)	104.69 (23.01)	B = -0.73 (p = 0.79)	0.03

Outcomes and analysis

The main analysis of the primary and secondary outcomes was conducted with two independent variables: the pre-test measure of the outcome and group allocation (Table 7). There was no evidence to suggest that the intervention group had better outcomes at post-test for the primary outcomes of reading attainment ($g = -0.02$, $p = 0.54$) and emotional self-regulation ($g = -0.02$, $p = 0.65$). Full details of the multilevel models (including the standard deviation used to compute the effect size for each outcome) are presented in Appendix H.

In terms of secondary outcomes, there was a very small, statistically significant difference in self-regulated learning at post-test ($g = 0.08$, $p = 0.03$) in favour of the intervention group, however, as this outcome was based on teacher report, it is subject to potential biases. There was no evidence to suggest an impact on social skills ($g = -0.07$, $p = 0.15$).

Table 7: Main analysis using only pre-test scores as a covariate

Outcome	Raw means				Effect size		p-value
	Intervention group		Control group		n in model (intervention; control)	Hedges g (95% CI)	
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)			
Primary outcomes							
Reading attainment	1,735 (321)	29.6 (29.2, 30.0)	1,577 (271)	30.2 (29.9, 30.6)	3,312 (1,735; 1,577)	-0.02 (-0.08, 0.05)	0.54
Emotional self-regulation	1,497 (559)	30.5 (30.2, 31.0)	1,313 (535)	31.0 (30.6, 31.4)	2,810 (1,497, 1,313)	-0.02 (-0.09, 0.06)	0.65
Secondary outcomes							
Self-regulated learning	1,621 (435)	33.1 (32.9, 33.3)	1,466 (382)	32.5 (32.3, 32.7)	3087 (1,621, 1,466)	0.08 (0.01, 0.15)	0.03
Social skills	1,116 (940)	153.9 (152.7, 155.2)	1,039 (809)	156.2 (155.0, 157.4)	2155 (1,116; 1,039)	-0.07 (-0.15, 0.02)	0.15

Subgroup analysis

Pre-specified subgroup analysis was carried out on pupils ever eligible for FSM (see Table 8). This involved repeating the main analyses described above for only this subgroup of pupils. There was no evidence to suggest that the programme benefits children eligible for FSM, with no effect on one primary outcome (reading attainment [$g=0.03$, $p=0.61$] or the secondary outcomes (self-regulated learning [$g=0.10$, $p=0.10$] and social skills [$g= -0.15$, $p=0.16$]), and a small negative effect on emotional self-regulation [$g= -0.12$, $p=0.21$]).

Table 8: Main analysis (using only pre-test scores as a covariate) for pupils eligible for FSM

Outcome	Raw means				Effect size		p-value
	Intervention group		Control group		n in model (intervention; control)	Hedges g (95% CI)	
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)			
Primary outcomes							
Reading attainment	284 (66)	27.3 (26.4, 28.3)	197 (48)	27.3 (26.1, 28.5)	481 (284; 197)	0.03 (-0.16, 0.21)	0.61
Emotional self-regulation	249 (101)	29.4 (28.5, 30.4)	161 (84)	30.6 (29.3, 31.9)	410 (249, 161)	-0.12 (-0.32, 0.08)	0.21
Secondary outcomes							
Self-regulated learning	268 (82)	31.9 (31.3, 32.4)	199 (46)	31.4 (30.8, 32.0)	467 (268, 199)	0.10 (-0.09, 0.28)	0.10
Social skills	179 (171)	151.6 (148.2, 155.1)	115 (130)	155.3 (151.8, 158.9)	294 (179; 115)	-0.15 (-0.38, 0.09)	0.16

On-treatment analysis

The poor return rate for the module report forms resulted in a high level of missing data for the dosage variable. Scores for dosage were based on whether each session was delivered and whether all, some, or none of the activities in the manual were followed. Scores could range from 0 (which would indicate

that the programme was not delivered at all) to 72, indicating that every session was delivered, and all activities were followed.

Reports were received at least once from 51 of the 78 teachers in the intervention group. These covered a total of 707 individual sessions (out of a potential maximum of 1,872—based on 24 lessons per teacher), equivalent to a completion rate of 37.7%. Dosage was therefore calculated based on partial data, and was available for 32 of the 42 intervention schools. The two intervention schools that withdrew after randomisation (but remained in the trial on an intention-to-treat basis) received a dosage score of 0. Dosage scores for remaining schools ranged from 14 to 64 and were approximately normally distributed with a mean of 38 (SD = 16).

There was no evidence to suggest that higher dosage was associated with better reading attainment ($p = 0.25$), emotional self-regulation ($p = 0.93$), self-regulated learning ($p = 0.94$), or social skills ($p = 0.26$); however, the partial data for dosage means that these results should be interpreted with caution and considered indicative at best.

Interviews with local co-ordinators indicated that fidelity was high. In their opinion—as local co-ordinators with ongoing contact with schools—all schools continued to deliver the programme and there was very little adaptation to, or deviation from, the programme manual. Also, the survey completed by 46 intervention group teachers suggested a high level of engagement with, and acceptance of, the programme (as described in the Process Evaluation section). However, as response rates were low, the dosage score generated from module reports may not be a true reflection of children's actual exposure to the programme. Further, given that module reports were generally collected by local co-ordinators at the follow-up support sessions, missing data is likely to be associated with non-attendance at these meetings.

Sensitivity analysis: multiple imputation

The proportion of missing data for all outcomes at pre- and post-test was examined, and was found to be greater than 5% for several outcomes, including 12% and 26% for the primary outcomes of reading attainment and emotional self-regulation at post-test (see Table 9).

Table 9: Proportion of missing data for each outcome

Outcomes	Intervention missing n (%)	Control missing n (%)	Total missing n (%)
Pre-test reading (HORT)	120 (6%)	45 (2%)	165 (4%)
Post-test reading (HGRT)	231 (11%)	238 (13%)	469 (12%)
Pre-test emotional self-regulation	58 (3%)	76 (4%)	134 (3%)
Post-test emotional self-regulation	527 (26%)	481 (26%)	1,008 (26%)
Pre-test self-regulated learning	95 (5%)	101 (5%)	196 (5%)
Post-test self-regulated learning	362 (18%)	293 (16%)	655 (17%)
Pre-test social skills	205 (10%)	184 (10%)	389 (10%)
Post-test social skills	817 (40%)	701 (38%)	1,518 (39%)

The proportion of missing data was particularly high for emotional self-regulation at post-test (26%), and social skills at post-test (39%). For both of these self-report scales, 13% of missing scores were a result of children not being present on the day of testing—in most cases, because they had left the school ($n = 244$) or were absent on the day of testing ($n=139$). The remaining missing scores were due to item-level missing data, which meant that scale scores were not calculated. The proportion of the sample with some missing data within scales is presented in Table 10.

Table 10: Proportion of missing data within scales

Scale	Missing all items	Missing 1 item	Missing 2 items	Missing 3 items	Missing 4+ items
Pre-test	n (% of 3,904)				
Emotional self-regulation (4 items)	106 (3%)	27 (0.7%)	2 (0.05%)	-	-
Self-regulated learning (12 items)	120 (3%)	48 (1%)	3 (0.07%)	2 (0.05%)	2 (0.05%)
Social skills (46 items)	103 (3%)	169 (4%)	57 (1%)	14 (0.4%)	46 (1%)
Post-test	n (% of 3,904)				
Emotional self-regulation (11 items)	499 (13%)	337 (9%)	112 (3%)	34 (1%)	26 (1%)
Self-regulated learning (12 items)	544 (14%)	103 (3%)	7 (0.2%)	1 (0.03%)	-
Social skills (46 items)	490 (13%)	480 (12%)	272 (7%)	96 (2%)	180 (5%)

The extent of missing outcome data at pre- and post-test was compared by group allocation using multilevel logistic regression modelling. There were no significant differences in the proportion of missing data between the intervention and control groups on any of the outcome variables.

Missingness on each of the post-test measures was also associated with scores on the respective pre-test measures for each outcome (determined using multilevel modelling with the pre-test variable as the dependent variable and a dummy variable representing missingness at post-test as the independent variable). Missingness at post-test across the full sample was associated with poorer performance on the pre-test measure for reading ($p < 0.001$), emotional self-regulation ($p < 0.001$), self-regulated learning ($p < 0.001$) and social skills ($p < 0.001$). This may be because those who were lower-achieving, or who had greater social and emotional difficulties, may have been more likely to be absent or have moved schools, or less likely to comply with data collection at post-test. The implications of this are that available data at post-test may be biased in favour of pupils with less severe academic, social, or emotional difficulties.

Pupil characteristics (gender, FSM eligibility, and EAL status) were also examined in relation to missingness on each outcome at post-test. Gender was not associated within missingness on any outcome. FSM eligibility was associated with missingness at post-test for reading ($p = 0.01$), emotional self-regulation ($p = 0.02$), and social skills ($p = 0.001$). Pupils with EAL were more likely to have missing data for reading ($p = 0.03$), emotional self-regulation ($p = 0.03$) and self-regulated learning ($p = 0.001$).

The data was reanalysed using multiple imputation. The imputation model imputed the data separately for the intervention and control groups, and included all variables used in analysis and the two variables used in randomisation (school level attainment and area). Other available variables (gender, FSM, and EAL) were considered as auxiliary variables, but were not correlated with missingness on any of the outcome variables, or scores on the outcome variables. Data was presumed to be missing at random (MAR) rather than missing completely at random (MCAR) or missing not at random (MNAR). This assumption renders the missing mechanism ignorable, simplifying the imputation step while ensuring correct inference (Enders, 2010). The imputation was performed using chained equations; this fills in missing values in multiple variables iteratively using a sequence of univariate imputation methods with fully conditional specification of prediction equations. This method accommodates arbitrary missing-value patterns. Twenty imputations were conducted in order to lessen the simulation ('Monte Carlo') error. Two hundred iterations (with a burn-in of 10) were used, and estimates were combined using Rubin's pooling rules. Table 11 compares the main complete cases analysis with the results of the analysis using the imputed datasets.

Table 11: Main analysis using both complete cases analysis and multiple imputation

Outcome	Raw means				Effect size		p-value
	Intervention group		Control group		n in model (intervention; control)	Hedges g (95% CI)	
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)			
Primary outcomes							
Reading attainment							
Complete case analysis	1,735 (321)	29.6 (29.2, 30.0)	1,577 (271)	30.2 (29.9, 30.6)	3312 (1735; 1577)	-0.02 (-0.08, 0.05)	0.54
Multiple imputation	2,056 (0)	29.2 (28.8, 29.5)	1,848 (0)	29.9 (29.6, 30.3)	3904 (2056; 1848)	-0.02 (-0.08, 0.04)	0.37
Emotional self-regulation							
Complete case analysis	1,497 (559)	30.5 (30.2, 30.9)	1,313 (535)	31.0 (30.6, 31.4)	2810 (1497, 1313)	-0.02 (-0.09, 0.06)	0.65
Multiple imputation	2,056 (0)	30.5 (30.1, 31.9)	1,848 (0)	31.0 (30.5, 31.4)	3904 (2056; 1848)	-0.02 (-0.08, 0.04)	0.58
Secondary outcomes							
Self-regulated learning							
Complete case analysis	1,621 (435)	33.1 (32.9, 33.3)	1,466 (382)	32.5 (32.3, 32.7)	3,087 (1,621, 1,466)	0.08 (0.01, 0.15)	0.03
Multiple imputation	2,056 (0)	33.0 (32.8, 33.1)	1,848 (0)	32.4 (32.2, 32.6)	3,904 (2,056; 1,848)	0.09 (0.02, 0.15)	0.01
Social skills							
Complete case analysis	1,116 (940)	153.9 (152.7, 155.2)	1,039 (809)	156.2 (155.0, 157.4)	2,155 (1,116; 1,039)	-0.07 (-0.15, 0.02)	0.15
Multiple imputation	2,056 (0)	153.2 (152.0, 154.3)	1,848 (0)	155.8 (154.5, 157.0)	3,904 (2,056; 1,848)	-0.06 (-0.12, 0.01)	0.10

Analysis using multiple imputation elicited broadly the same magnitude of effects as the complete case analysis, therefore there is no evidence to suggest that the main findings would have been different had there been no missing data, under the assumption that data was missing at random.

Cost

In the first year, Zippy's Friends costs approximately £512 for a school to deliver with one class. This estimate includes initial training for one teacher (one day) and the purchase of programme materials (teacher and pupil resources, as described in Table 12) for one class. These costs were covered by the EEF for schools participating in the trial that were allocated to the intervention group.

Table 12: Breakdown of costs associated with delivering Zippy's Friends (first year)

Item	Detail	Cost per class
Training	£2,000 for one-day training for up to 25 teachers (excluding VAT).	£80

Materials	One full set of teacher and pupil resources (including £32 for postage and packaging; excluding VAT). This includes teacher's folder, inclusion supplement, 8 posters, 60 puppets, 30 crowns, 30 children's certificates, 1 class certificate, 1 school certificate, 30 parents' guides, 80 Zippy Tips cards and 30 <i>My Zippy</i> books.	£357
	Story booklets (excluding VAT).	£75
Total		£512

Assuming an average of 27 pupils per class (and one class per school), the average cost of the programme is approximately £19 per pupil in the first year. This figure, however, does not include indirect costs associated with:

- supply cover for teachers to attend one day of training and attendance at the three support sessions (if during the school day);
- teacher time preparing for, and delivering, sessions;
- any materials over and above those provided in the resource pack; or
- the cost of the class trip to the graveyard (for example, transport hire and additional DBS-checked adults in line with adult: child ratio for school trips). (See page 37 for further details.)

Teachers (n = 41) reported spending an average of 70 minutes per week (SD = 20) on Zippy's Friends during school time, of which, 45–60 minutes related to session delivery and 10–25 minutes to preparation. Time taken to attend the support sessions varied; these meetings lasted one to two hours, and teachers generally needed to travel to another venue (usually another school) within the local authority.

In the online survey, teachers reported that additional materials used in activities were minimal. The majority of teachers who responded to the survey reported a small amount of photocopying or printing of activity sheets each week, and the need for artwork materials.

Schools will likely need to buy additional pupil materials each year at a cost of £173 per class (including postage and packaging and excluding VAT). This means that in subsequent years, the cost per pupil is £6.41. Over three years, the cost of delivering Zippy's Friends is low, at approximately £10 per pupil (Table 13).

Table 13: Cost of Zippy's Friends over three years

Number of years using the programme	Cumulative cost per pupil (£)	Average cost per pupil per year (£)
1 year	19	19
2 years	25	13
3 years	31	10

Process evaluation

This section presents findings from the implementation and process evaluation of Zippy's Friends. Participants included teachers, headteachers, pupils, local co-ordinators, and parents involved in the project. Data was collected through interviews with a sample of teachers, headteachers, and parents, interviews with each of the local co-ordinators, focus groups with pupils, observations of programme delivery, an implementation survey, and module report forms. Findings from these components have been integrated and are organised under the main headings of Implementation, Fidelity, Outcomes, Formative Findings and Control Group Activity.

Implementation

Training and support for delivery

Year 2 teachers received one day of training, delivered directly by Partnership for Children (PFC), at the beginning of the 2016/2017 school year and in advance of programme delivery. Training was delivered for groups of teachers within each local authority. At this stage, teachers were provided with a set of core materials needed for delivering the programme. Overall, teachers were positive about the training they received, with 85% (of the 42 implementation survey respondents) agreeing or strongly agreeing with the statement 'I am confident in the delivery of the programme following the training I have received'. This is also reflected in the interview data: some teachers made reference to the group-based nature of the training and how practical aspects of programme delivery were demonstrated:

We had a pretend Zippy's Friends lesson that they [PFC] ran; [...] we were the pupils and that kind of gave us an insight into how to run the lessons, which was really probably the most, for me, important part of the training.

The group-based nature of the training was cited as an advantage, with many teachers valuing the opportunity to meet staff from other schools and discuss the programme:

I think it really helped, and talking to lots of others was great support as well.

Ongoing support for teachers in the delivery of the programme was provided by the local co-ordinators through group-based support sessions and observation visits. There was a high level of satisfaction with the level of support received—93% of survey respondents reported that there was sufficient support for the ongoing delivery of the programme. The group-based support sessions took place three times during the course of the school year and brought together intervention group teachers within each local authority to discuss experiences delivering the programme (including what has worked well and what was found to be challenging about delivery), to guide and support teachers in terms of implementing the programme with fidelity, to address any questions, and to plan for forthcoming modules. Teachers reported that they valued the opportunity to share experiences and good practice, and gain further ideas for activities and use of resources. For example, a local co-ordinator observed:

I think the main benefit [of the support sessions] is coming together, is in sharing the experience, what others have perceived as positive and areas for development identified, and we have been able to address many of the areas for development just through discussion. An example of that, it was reported by the teachers at the twilight session, that repeating the rules [in] every session had become quite stagnant and become kind of a box ticking exercise, rather than anything meaningful. So in twilight meeting two, we actually brainstormed all the different ways in which the teacher could deliver those rules in a more engaging way and a more lively way to keep the children on their toes with the rules, but also to make sure that that period of time was meaningful, purposeful, part of the programme.

While the majority of teachers in most local authorities managed to attend at least one of these meetings, overall, attendance varied. On average, across all five areas, 58% of teachers attended the first support meeting (with attendance ranging from 33%–80%). At the second meeting, average attendance was 59% (ranging from 53%–64%), and at the third session, 68% (ranging from 32%–92%). There appeared to be an overall sense from the teachers interviewed that while these sessions were useful for sharing experiences with other teachers, they were not entirely necessary for delivering the programme. One local co-ordinator reflected that the comprehensive and structured nature of the programme materials may have meant that many teachers did not feel the need to attend support sessions:

I think the reason why attendance at the meetings has been varied, is some teachers felt the additional meetings could be seen as a bolt-on rather than integral to the process. I think that the provision is comprehensive enough that a teacher can just pick it up and, 24 weeks later, they are going to have finished that programme, and they have stuck pretty religiously to that [...] But I don't think they have needed the support meetings in order to successfully implement the programme.

Local co-ordinators also conducted an observation of one full session in all the intervention schools, shortly after the initial training. This served as an opportunity to give teachers feedback, answer any questions, and identify or address any further training or support needs. Interviews with local co-ordinators during this time indicated there were no issues with the initial implementation of the programme. In interviews, teachers reported that the visit from the local co-ordinator was very supportive:

Yeah, and we were observed teaching it [by a local co-ordinator], and we were given feedback about how well we were doing, and [they] gave us feedback about how we could improve it or adapt it. That was really supportive.

Barriers to delivery

The main barriers to delivery were identified as a lack of support from senior management and time constraints. These two issues appeared to be conflated at times, the former creating difficulties in terms of ensuring sessions were timetabled appropriately—that enough time was given to the programme within the school day.

There was some variation in the extent to which teachers perceived the school leadership team supported the programme. The majority of survey respondents (64%) agreed or strongly agreed that 'senior management in the school are keen to see the programme work', however, 17% disagreed with this statement. This variation is also reflected in the interviews with teachers. Many teachers reported being supported by school leadership, in particular when leadership recognised the need for a SEL programme and were involved in the decision to take Zippy's Friends on:

It was the headteacher's decision, actually, to be part of the programme, because lots of the children in the class have issues. She [headteacher] is very supportive and lets me [go] away to all the meetings and stuff.

One teacher wanted school leadership to be more enthusiastic about the programme; in another school, support from leadership was felt to be lacking:

The only sort of support we have had is from [the local co-ordinator]; we have had no sort of support from inside school to deliver it—we pretty much ran with it ourselves.

Although teachers were, on the whole, very positive about Zippy's Friends, a number of practical barriers to delivery were reported, and these were generally around time constraints related to preparation and delivery. At times, the volume of activities to fit into a 45-minute session was considered

challenging, particularly when sessions involved preparing children's resources. Some teachers split these sessions into smaller blocks so they could be completed.

Some teachers interviewed noted that leadership support was needed to ensure timetabling issues were addressed, and to ensure teachers were released from normal duties so they could complete paperwork (such as report forms and questionnaires). It must be noted, however, that much of the paperwork involved related to the evaluation rather than to the intervention itself.

The awareness and understanding of school leadership in relation to the programme was cited by a local co-ordinator as a fundamental aspect of successful delivery, particularly in terms of authorising school trips and ensuring teachers had time to attend meetings:

I also think that an important element to this is the headteacher's awareness and understanding of the programme and the commitment associated to the programme, because in dealing with class teachers, inevitably sometimes they haven't got the authority and the responsibility to make decisions around trips, around things such as releasing themselves to attend meetings [...] So I think that I would potentially, [when] identifying that schools are interested, have some kind of twilight meeting for headteachers, to talk about what the programme is going to entail, and the school's commitment.

Teacher buy-in

An overwhelming majority of survey respondents (95%) felt that a SEL programme (although not specifically Zippy's Friends) was necessary, with a similarly high proportion of respondents reporting (towards the end of the intervention period) that they were 'engaged and excited with the programme' (95%) and 'enthusiastic that my school has adopted this programme' (98%). The commitment of the school staff involved with implementing the programme emerged from the interviews. For the majority of teachers and headteachers interviewed, Zippy's Friends was seen as an important addition to the school in terms of addressing pupils' social and emotional development. Many teachers cited the programme as being an important resource for them in helping to address increasing unmet needs within the class:

Yes. I think [a programme like this is necessary] in particular settings, definitely. Just because of the needs of our children, it's worth dedicating that time in our timetable to that particular aspect.

Really before Zippy we were just firefighting with some of the kids. This has been a really useful thing now for us to have.

There is a rising need in this school and in particular my class this year. I feel it has just come at the right time.

Some schools expressed the view that they chose to participate in the evaluation of Zippy's Friends because they felt the materials would be a key way of developing pupils' literacy levels as well as supporting their social and emotional development. This is reflected in the following comment:

So it was two birds, one stone. We had to see if we could actually put anything in place that would significantly impact the children's emotional resilience and have a positive outcome for their literacy levels at the same time.

Pupil engagement

Overall, in interviews, teachers reported that, in general, children had responded positively to the programme and looked forward to the weekly Zippy's Friends sessions; that children responded well to

the stories and to the activities—using their 'Zippy' pencils, mystery box, role-play making puppets, and so on.

I've really enjoyed it, the children really enjoy it as well and they really look forward to the sessions and they're able to take what they've learnt and use them in different contexts as well, so like in the playground, [they can] sort out their friendships; in class, if we're reading a story they can relate it to issues that might have been in Zippy's as well.

I like it, yes, it is nice—it is always a really rewarding lesson. And they [pupils] really, really enjoy it. Sometimes, it is hard to get them actually back into it, because they are all laughing so much, and they just find it really funny to act out real-life situations, they are laughing, they just find it funny.

One aspect that teachers believed contributed to the high level of enjoyment among pupils was the more relaxed nature of the lessons, and the ambiance this created in the class.

The high level of engagement from pupils is also reflected in the module reports, which were received from 51 out of 78 teachers and covered a total of 707 sessions. Pupil enjoyment was rated by teachers on a five-point scale: 1, 'not at all'; 2, 'a bit'; 3, 'moderately'; 4, 'very much'; and 5, 'extremely'. Teachers most often rated pupil enjoyment of sessions as 'very much' (52%) or 'extremely' (24%). Teachers rated 17% of sessions as being enjoyed 'moderately' by pupils, and only a small number were enjoyed 'a bit' (6%) or 'not at all' (0.3%). Mean scores for pupil enjoyment were generated for each module, and appeared to be highest early on in the programme, with the highest scores received for modules 1 and 3 (Table 14), although the fall in return rate over time makes it difficult to identify any meaningful trend. Module 5 was the least enjoyed (mean = 3.6), and comments from teachers in the module reports indicate that in some cases, children got upset due to the nature of the story and discussion (focused on death and loss).

Table 14: Mean pupil enjoyment score, by module (teacher-rated)

Module	Number of schools (out of 42)	Pupil enjoyment (across 4 sessions)	
		Mean (SD)	Min, max
1	31	4.05 (0.51)	3.00, 5.00
2	28	3.94 (0.50)	2.75, 5.00
3	27	4.05 (0.58)	2.50, 5.00
4	18	3.83 (0.65)	2.25, 5.00
5	13	3.61 (0.65)	2.33, 4.67
6	10	3.75 (0.81)	2.00, 5.00

The overwhelming majority of children from across the focus groups felt that all children should have Zippy's Friends. As illustrated below by a representative sample of comments, the children were keen to convey the reasons why children would benefit from getting the programme:

We learn how to keep safe and when you are older you will not split up with your friends.

I think everybody should get Zippy's Friends because we've got the opportunity to do it and it helps us all to stay friends so it would be unfair if they didn't get the chance.

Programme content and resources

All survey respondents agreed that the programme materials and resources were useful. The majority of references to the Zippy's Friends manual and teaching materials made by teachers were very positive, with reference to the clear structure, detailed instructions, and ease of use:

So, in terms of implementation, there's four sessions in a module which is relatively easy to fit within a half-term. Plus, all the resources are so self-explanatory and so in-depth that you know exactly what you're doing. Even if I picked up that resource ten minutes before, nine times out of ten I'll probably be able to follow it through and run with it without being panicked that I wouldn't be able to.

I think what helps is it is literally ready to go, there are resources there, it doesn't take too much time to prepare for, you just put all the resources up on the board ready [...] and it is very easy to follow.

Teachers also reported that pupils reacted positively to the content and resources:

The resources are fantastic. I mean, they're there, they're totally manageable, they're attractive, the kids have got to know Zippy, they love their books, they have their special pencil; from the day we started they have loved doing that.

They [pupils] really enjoyed the stories and they really could relate to the characters, and empathise and talk about them.

However, a few teachers felt that elements of the programme could be repetitive at times and that sometimes the children would get bored with this. This is reflected in the comments below:

Yeah, they [the lessons] were very scripted at the beginning. And, I think the books as well sometimes became a little bit repetitive. When they filled in their evaluation [activity completed by pupils after each session as part of the programme], it kind of felt like sometimes they were doing the same thing.

So, a different type of activity might have engaged them a little bit more. Varying those activities a wee bit might have just kept the students engaged a wee bit, give them a wee bit of variety.

Sessions are designed to last around 45 minutes, and the manual suggests timings for each activity within the session. The module reports indicated that sessions lasted an average of 45 minutes (SD = 7.3), and 24% of sessions lasted less time than this. Teachers are expected to read through the plans before each session to see what they need to prepare in advance. Some of this preparation includes having to photocopy activity sheets for some of the session or make larger versions of them to show the whole class.

When teachers were asked in interviews about their perceptions of the content of each session, responses were generally positive:

The content's been good. Some of the activities were either a lot longer than planned or they're a lot quicker but most of them have been really good and they really enjoy ones where they get to role play and talk as well and draw.

I think [the content] is excellent. I think the repetition of it is lovely and I think for this age group it's absolutely ideal.

However, one teacher reported that it could be difficult to 'pad out' the session for the full duration, and others reported that some sessions involved too much sitting and listening:

But sometimes there are lessons where it's a bit of carpet stuff then we go back to the table, do a little bit in the books, come back or they might stay at the table and we talk. But I think it is a long time for them in Year 2 to stay engaged, so the lessons where there is a lot of sitting and listening, I find I'm having to find ways of doing little brain breaks or, you know, bringing other little activities that I can think of into it.

The module report forms indicate that teachers, on the whole, used all of the activities (81% of sessions delivered), with very few using none of the activities (1% of sessions). One issue, which arose in interviews and was reflected in the module reports, was that there were topics on which the class had lots to share, and in these instances, a choice had to be made between completing all the activities, or allowing discussions with the children to continue:

Yes, sometimes it is a bit rushed and they [children] have all got their hand up so you want to listen to every single one, because they want to share, but sometimes you can't and you have got to just say stop, so sometimes time can be a little bit restricted.

In addition to each session, there is an Inclusion Supplement that incorporates additional activities to enhance and reinforce the original session plans. These activities are designed particularly for children with additional support needs. The majority of teachers interviewed stated they did not use any of the extra activities due to time constraints, and this is reflected in the module reports (where 3% of sessions used this supplement). Some teachers alluded that in the future, when they were more familiar with the programme, they felt these would be useful to use with small groups or one-to-one with pupils with special needs:

I never do [use the Inclusion Supplement]. I don't know whether we would ever get on to them, because, especially my class, they all want to share. So maybe if I had a quieter class that don't all want to share their information, they would get on to that. But for me, it is always a rush. This is a pity because some of our children would benefit. I am hoping to look closer at these next year.

Teachers in general felt that the programme was well suited to Year 2, and that resources and activities were age-appropriate and child-friendly:

I think it is really important for this age group, because I think aspects of SEAL can be a little bit boring and a bit repetitive, whereas [with] Zippy's Friends it is always something new, a new lesson every single time, and it always links to what has happened on the playground. Some of the children just don't know how to deal with conflict and it is just teaching those skills that maybe their parents don't have time to or don't think it is necessary to teach.

Yes, I think it is very age-appropriate, it relates to all, they can relate to it, it is things that they recognise have happened to themselves. So far—I know we are coming up to the bereavement part, which I know that some of them haven't [experienced]—but so far all of them have been in situations that they can link to.

It has actually been very much pitched to their level. You could probably almost even bring it into Year 1, Zippy's Friends, or as a crossover unit between, start it halfway through Year 1 and roll it into Year 2. That might benefit them more.

Fidelity

Fidelity of implementation was examined quantitatively through the module report forms and qualitatively through interviews with teachers and local co-ordinators. As described earlier, there was a poor return rate for the forms. Reports were received from 51 out of 78 teachers in the intervention group. These covered a total of 707 individual sessions (representing 36% of all sessions delivered, based on 78 teachers delivering 24 lessons each).

Analysis of the reports received indicated that the majority of sessions (76%) lasted the recommended 45 minutes (with the remaining sessions lasting less than 45 minutes). Zippy's Friends was delivered in a regular slot each week in 82% of cases, suggesting high priority given to the programme by schools. Reasons given for changing the time slot were to facilitate lesson observation by the local co-ordinator, to accommodate other activities in school (such as Christmas plays, school trips, assessments, parent meetings, and sports day), and, to a lesser extent, staff illness or unavailability.

One of the key components of the programme is the 'rules', which are read out at the beginning of each session, and module reports indicated that this happened in 94% of sessions. Table 15 presents a breakdown of the extent to which the activities for each session were followed. All activities in the manual were followed for 81% of sessions, and some activities were followed for 16% of sessions. Only very rarely (1% of sessions) did teachers report not using any of the activities, and where this was the case, the reasons given were time constraints, opting to spend longer on other parts of the session (such as the discussion), that some parts of activities seemed repetitive or not needed, that pupils found some activities difficult to understand, and that teachers perceived children as self-conscious during the acting activities. Of note within Table 15 is module 5 session 3, in which no activities were followed in almost half (43%) of module report forms returned. This session involved a class trip to a graveyard, and comments within the module report forms indicate that this was perceived by some schools to be impractical or inappropriate.

On the whole, analysis of available module reports indicates a high level of fidelity; however, it is acknowledged that the received module reports reflect only 36% of total sessions delivered. It is possible that fidelity was poorer for sessions where no module report forms were returned, however it is worth noting that survey responses (from 46 intervention group teachers) indicated a high level of engagement with the programme among schools.

Table 15: Use of programme activities during each of the 24 sessions

Module (session)	Module reports returned	Activities followed		
		All n (%)	Some n (%)	None n (%)
1 (1)	48	44 (92%)	4 (8%)	-
1 (2)	47	42 (89%)	5 (11%)	-
1 (3)	47	42 (89%)	5 (11%)	-
1 (4)	47	43 (91.5%)	4 (8.5%)	-
2 (1)	44	37 (84%)	7 (16%)	-
2 (2)	43	35 (81%)	8 (19%)	-
2 (3)	42	38 (90.5%)	4 (9.5)	-
2 (4)	40	36 (90%)	4 (10%)	-
3 (1)	39	31 (79.5%)	8 (20.5%)	-
3 (2)	37	28 (76%)	9 (24%)	-
3 (3)	37	29 (78%)	8 (22%)	-
3 (4)	33	28 (85%)	5 (15%)	-
4 (1)	24	19 (79%)	5 (21%)	-
4 (2)	22	15 (68%)	7 (32%)	-
4 (3)	20	14 (70%)	6 (30%)	-
4 (4)	20	12 (60%)	8 (40%)	-
5 (1)	15	11 (73%)	3 (20%)	1 (7%)
5 (2)	16	12 (75%)	4 (25%)	-
5 (3)	14	7 (50%)	1 (7%)	6 (43%)
5 (4)	13	10 (77%)	3 (23%)	-
6 (1)	12	9 (75%)	3 (25%)	-
6 (2)	11	10 (91%)	1 (9%)	-
6 (3)	10	9 (90%)	1 (10%)	-
6 (4)	10	10 (100%)	-	-

Challenges to delivery

Although teachers were on the whole very positive about Zippy's Friends, a number of practical barriers to delivery were reported, and these were generally around time constraints related to preparation and delivery. At times, the volume of activities to fit into a 45-minute session was considered challenging, particularly when sessions involved preparing children's resources.

While the majority of teachers were able to fit in the programme with ease throughout the year, many noted that the pressure of KS1 tests in the final school term made it challenging to deliver the final module, not least because of the demanding subject matter—dealing with death:

We struggle a bit more around [...] SATs time, because the timetable then gets a little bit more packed with exam revision, testing, mock tests, so at that point it is quite a hard thing to keep going regularly. And, at that point also coming in with a module on death is quite... I think you really want it to be like, less regimented but unfortunately it came with a difficult module at a difficult time, by starting it at the start of Year 2.

The content of the programme was well received by schools taking part in the process evaluation, with the exception of module 5, which got mixed feedback. Module 5 involves stories and discussions on change and loss as part of life, coping with death, a visit to a graveyard, and learning from change and loss. While most teachers interviewed cited the learning from this module as crucial in terms of helping children to become better equipped to cope with difficult situations, a minority of teachers struggled with this content and found it challenging to implement in terms of dealing with children's emotional reactions:

To be dealing with tears and things, yes we really want them to be happy and enjoy coming into school, and do their best for the tests. But, unfortunately when we are trying to teach them something about death and loss, so that was the bit I found a bit hard to do.

The visit to the graveyard was also challenging, mainly due to logistical aspects such as transport, time, health and safety, having enough members of staff involved to satisfy staff:pupil ratio. Discussions at the final support meeting, and interviews with local co-ordinators at the end of the programme, indicated that few schools completed this visit. There was indication, through module report forms and interviews, that teachers perceived the graveyard visit as not being a core component of the programme, however teachers strove to find other activities that could be used in place of this visit, for example, using photos of a graveyard in class or visiting a memorial bench in the school grounds.

Teachers reported that at times the discussions or activities meant that issues were raised that may have benefited from further follow-up or input, but that there may not be the time to do so:

There's one little boy in particular that really opens up about things that are happening at home and he's one that you wouldn't expect. Sometimes it would be great to have more time after the lesson to follow this up but it is straight back to other lessons.

Outcomes

Perceived benefits of the programme

All teachers who responded to the survey agreed (45% strongly agreed) that the programme is very beneficial for the children. On a similar note, all teachers who were interviewed said they had noticed the benefits of Zippy's Friends for children in their class. The main changes noted by the teachers included improved problem-solving, being more self-reliant when dealing with, for example, difficulties in the playground, and using more varied vocabulary when discussing how they were feeling. This is reflected in the following comments made by teachers in the module report forms:

Children are able to reflect with adult support if something was a good rule or choosing a solution. Children feel more confident about talking in circle than in first few sessions.

The children are beginning to solve their problems using the golden rules solution.

Ability to talk through problems and think of solutions.

The noise levels are down in the class. Behaviour is better in class.

Perceived benefits of the programme, as reported by teachers through the process evaluation, were contradictory to the findings of the impact evaluation (which found no impact on emotional self-regulation and social skills, and a very small impact on self-regulated learning). Potential reasons for this are discussed in the Conclusion section of this report.

Teachers interviewed were generally unsure about whether the programme would impact on the academic attainment of their class, and there was a mixed reaction on this topic within the survey where

almost two thirds of teachers (68%) responded 'no' to the question 'Do you think Zippy's Friends has an effect on the children's academic achievement?' (with the remaining 32% responding 'yes').

Teachers recalled certain components within the programme that they perceived were helping to develop skills with the pupils. First, the rules that are read at the start of every lesson were found to be applied more widely in the class/school/lunch time in many of the schools.

The rules... now they know them really well and use them in lots of other situations throughout the day.

[Children are showing] more mature behaviour when sorting out problems in the playground.

Children support each other better in the playground at breaks/lunchtimes.

Second, teachers commented on the benefits of Zippy's Friends for pupil's emotional literacy and vocabulary, and how, in some cases, this had helped pupils in terms of understanding their own feelings and the feelings and behaviours of others:

The children who found it more difficult previously are able to discuss feelings and deal with them more appropriately as Zippy's Friends has provided them with a structure.

Unless the children have got the language to convey how they are feeling, they can't get it across to you, so at least it has broadened their vocabulary.

Extremely impressed in how the children are articulating their feelings anxious, annoyed.

A little boy in my class, going back a couple of months ago, he was telling on children at lunch time when they hadn't done anything, and we kind of sat there talking about why he did it and he said, 'Well, actually I felt really jealous of the two [pupils] because they have lots of friends, and I felt jealous because it makes me feel sad inside that I haven't got the friends that they have got.' But, words like jealously and comparing the sort of friendship circles to him and that sort of emotions, would never have really come from a child like him, which is wonderful to see actually, that he can put his thoughts down to jealousy, and wish he had a bigger set of friends like the other children had.

One teacher commented how developing emotional literacy may be especially advantageous for children from more vulnerable or deprived backgrounds (in contrast to the impact evaluation findings which suggest no additional benefit of the programme for FSM pupils on any outcome):

I think some of the children in my class have very stable, very lovely home lives where actually things like feelings are talked about on a daily basis, and actually that all feelings are good feelings. I think some of the children from more problematic home lives actually need to know that their feelings are valued, and that to feel angry is a perfectly good emotion to have. I think for those sorts of children from more deprived home lives, it has been great for them. More than some of our children from very privileged home lives, actually they probably didn't get as much out of this than some of the other children.

One teacher commented that the programme may flag up social and emotional issues, for example, relating to the difficulty some children have in identify feelings, and that this may be where the Inclusion Supplement would be useful:

I think the challenges have been how to think about those children, because when we've come out of those sessions and they've still not been able to give an example of when they felt a certain way or something's happened, well, I don't like to leave that. If I've identified a need then let's do something about it. So, it's about scooping those children up after the event to try

and address that, and having the time and the facilities and the resources to do so, but the small-group supplement does support them with that.

There was also evidence that the programme was beneficial for teacher-pupil relations. In the survey, 98% of teachers agreed that they were more aware of children's feelings, and 90% agreed that 'my relationship with the children has improved'. This is also reflected in some of the interviews with teachers:

I know more about them, what makes them upset, what worries them outside of school. And, if the child is coming in late or upset, you can almost help talk to them about what has happened at home, and they are more likely to share that sort of information with you, because they are so used to doing it on a weekly basis.

And, you never have time without this [Zippy's Friends] to discuss things like that [feelings], it is always maths, English, go and sit down, answer the questions. So, it actually gives them a chance to talk about it if they want to, and share things, and I learnt things about the children that I didn't necessarily know before.

Almost all survey respondents (95%) agreed with the statement 'the programme has made a difference to the atmosphere in the classroom', and in interviews, teachers also felt that the programme had a beneficial impact on the classroom environment.

It's just nice, it's a really informal, friendly, welcoming environment with the children and they do enjoy it [...] they know they can go into that lesson feeling secure and safe.

And they are more willing to come to school.

At the beginning of the year they used to like bicker and be like, 'Miss they're doing this' and just tell on everything whereas now they try and solve the problems themselves and they're a lot calmer as well which is good.

When asked what they had learned from the programme, the majority of the children in the focus groups made reference to learning about the importance of maintaining and creating friendships, and treating others well. Contrary to the logic model proposed for this study, there was no suggestion (based on pupil focus groups) that Zippy's Friends was impacting on academic attainment:

It helps me to make new friends and stay friends and to not bully other children and make them feel sad.

I like Zippy's Friends because it tells you, like, how to make new friends and how to keep them, and how to not fall out.

[Zippy's Friends has taught me] to always care about people when they're upset. If someone fell down I would go and help them.

Zippy's Friends taught me when someone is very upset you could think of a good solution to tell them to make them feel better.

All parents interviewed were aware that their child's class was delivering Zippy's Friends, however there did not appear to be a great deal of awareness on the content or intended outcomes of the programme. All parents knew about the character, Zippy, as a result of their children talking to them about it at home, or from reading about it in some of the homework activities. Some of the parents commented that the programme had helped their children to deal better with problems:

There's one of the cards that she got from school, you know, how to deal with a problem that you don't know how to resolve and, you know, think, calm down, think of the best solution and, you know, she still got that stuck on her wall in her bedroom.

She used to come home and say, "Oh, we had Zippy today and we did this, and we did that" [...]. And, I noticed like suddenly throughout the year, I kind of noticed the change in kind of how she responded to things, and stuff like that. And, so that was... Yeah, it was good. I remember a few times when I said, "Oh wow! That's impressive!" (laughter).

Formative findings

Overall, the structure and content of the programme was highly acceptable to schools. Only 12% of survey respondents agreed that 'the programme misses important elements that it should have covered'. One recommendation, which came up in some interviews with teachers and a local co-ordinator, was that school leadership should be provided with more information about the programme and the time commitment involved for class teachers with the aim of ensuring that class teachers are fully supported in delivering the programme.

Linked to the first point, a second recommendation arising from the process evaluation was that a whole-school approach to delivery (at least ensuring an awareness of programme throughout the whole school) would be beneficial:

One of the main areas for development has been that schools have felt that they are doing this in a silo really, in a discreet year group, and a lot of schools have lacked that whole-school awareness of what is happening in Year 2. I think once again that comes back to the headteacher or a member of senior leadership really taking responsibility and ownership for this.

Several participants suggested that it would be important for both teaching and non-teaching staff in the school to receive some training on the programme. In particular, lunchtime staff were mentioned on several occasions as staff who tend to interact with pupils during unstructured periods (like break and lunch time where issues between children can come to the fore). Ensuring these members of staff were aware of, and able to, use some of the strategies introduced by the programme, may be beneficial. Several teachers felt that the lack of consolidation of the learning from the programme across the school would impact on its effectiveness:

I'd like our lunchtime supervisors to have some of those little mini cards and things like that so that they can use that. We have peer mediators in the school which are the older children and I think it would be a really good tool for them to use, you know.

Unfortunately, because it is not reinforced as a whole while at school with lunchtime staff etc., I don't think it is as effective as it could be.

You need assemblies, all the children to know about it. You need staff meetings, all the staff to know, and not just as a one off, I think you need it reinforced once every few weeks to be honest.

Other suggestions for improvements from teachers were around introducing more variety into the stories, which they reported became a bit repetitive, reducing the amount of sitting down, 'circle-time'-based work within a session and building in more role play and multi-sensory activities that involve moving around.

Control group activity

Randomisation took place at the school level, minimising the risk of contamination. There was no evidence to suggest that schools in the control group delivered Zippy's Friends.

Prior to randomisation, all schools were asked to provide information on standard practice in relation to social and emotional learning. The vast majority of schools (that were subsequently allocated to the control group; $n = 41$) reported that SEL was covered 'a little' (64%) or 'very much' (31%). These percentages are comparable to those reported by schools subsequently allocated to the intervention group, of which 61% reported covering SEL 'a little' and 29% 'very much'. Only 5% of control group schools and 10% of intervention group schools reported that there was no SEL provision in place at this stage of the project.

Just over a third of control group schools ($n = 15$; 37%) reported having a SEL policy in place, compared to 41% ($n = 17$) of intervention group schools. SEL activities were covered across other parts of the curriculum for 87% of control group schools and 89% of intervention group schools. Fifty-nine percent of control group schools had a timetabled slot for SEL compared to 47% of intervention groups schools ($p = 0.31$). In control group schools, the timetabled SEL slot lasted an average of 37 minutes ($SD = 24$) per week, which was not significantly different to the timetabled slot in the intervention group, which lasted an average of 41 minutes ($SD = 18$; $p = 0.55$). A specific SEL programme was used in 37% ($n = 15$) of control group schools and 41% ($n = 17$) of intervention group schools ($p = 0.79$). The most frequently mentioned programme was SEAL ($n = 10$ schools in each group). The other programmes mentioned in the control group were Philosophy for Children, R-time, SUMO, Time to Talk and Bear Buddies. The intervention group mentioned the programmes Jigsaw, KIVA, SUMO and other, local schemes of work.

At the end of the intervention period, teachers in the control group were asked to complete a survey on actual practice in relation to SEL during this year. This asked about any interventions, programmes, or strategies that any or all pupils in their class would have been exposed to during the trial period. Fifty (out of a potential 71) teachers responded. Of these, 64% reported that their school had a SEL policy, which was similar to the proportion in the intervention group (69%) at the end of the intervention period.

Teachers in the control group reported spending an average of 55 minutes per week on SEL activities with the whole class, compared to 65 minutes for intervention group teachers ($t = -1.70$, $df = 90$, $p = 0.09$). This suggests that the control arm was, in fact, very active in the delivery of social and emotional learning, constituting a threat to the internal validity of the trial. While it is highly unlikely that children in the control arm would have had any exposure to Zippy's Friends (as access to programme training and materials is controlled by the programme delivery team), there are a number of other social and emotional learning programmes and interventions available. Teachers in the control group reported using such programmes with their class during the trial period (see Table 16). Notably, Social and Emotional Aspects of Learning (SEAL) was used by 62% of teachers in the control group (compared to 37% in the intervention group; $p = 0.01$). Also notable is that 54% ($n = 27$) teachers in the control group (compared to 33% in the intervention group) reported that individual pupils from their class attended a nurture group ($p = 0.04$). Schools that signed up to participate in this project generally did so in order to address perceived needs within their school, and this data would suggest that schools allocated to the control condition actively sought out other SEL programmes to implement in order to address these needs. Thus the counterfactual is no longer 'business as usual' or 'no treatment'; instead, it seems likely that the counterfactual represents the delivery of other SEL programmes, in which case we cannot confidently conclude that Zippy's Friends is not effective—only that it is not more effective than other programmes.

Table 16: Comparison of other social and emotional learning programmes used by teachers at the end of the trial

Programme	Intervention group n (%)	Control group n (%)	p-value
SEAL	17 (37%)	31 (62%)	0.01
Targeted Mental Health for Schools	8 (17%)	9 (18%)	0.94
National Healthy Schools Programme	11 (24%)	17 (34%)	0.28
Circle Time	10 (22%)	9 (18%)	0.65
Circle of Friends	10 (22%)	17 (34%)	0.18
Nurture Groups	15 (33%)	27 (54%)	0.04
Inclusion Development Programme	2 (4%)	2 (4%)	-
Achievement for All	3 (7%)	4 (8%)	0.78
Place2Be	4 (9%)	3 (6%)	0.61
Seasons for Growth	1 (2%)	0	-
Friends Forever	1 (2%)	0	-
Fun Friends	2 (4%)	1 (2%)	-
Behaviour for Learning	5 (11%)	5 (10%)	0.89
PATHS	1 (2%)	0	-
Roots of Empathy	1 (2%)	1 (2%)	-
Second Step	1 (2%)	0	-
Increasing Learning Power	1 (2%)	2 (4%)	-
Incredible Years	1 (2%)	0	-
Jigsaw	4 (9%)	1 (2%)	0.14
KIVA	4 (9%)	2 (4%)	0.34
RTime	4 (9%)	0	-
Good Behaviour Game	1 (2%)	2 (4%)	-
SUMO	2 (4%)	4 (8%)	0.46

Tested using chi-square (p-value not reported where all cells had a minimum expected count <2).

Conclusion

Key conclusions

1. The project found no evidence that Zippy's Friends improved reading outcomes or emotional self-regulation for Year 2 pupils. These findings have low to moderate and low security, respectively.
2. There was no effect on reading attainment for pupils who had ever been eligible for free school meals. There was a small negative impact on emotional self-regulation for these pupils, but this result has lower security than the overall findings because of the smaller number of pupils.
3. Children receiving Zippy's Friends made small improvements in teacher-reported self-regulated learning, compared to other pupils.
4. The programme was very well received by teachers: all survey respondents felt Zippy's Friends had benefits for the children involved.
5. Evidence suggests that the impact of social and emotional learning programmes can take time to feed into academic outcomes. It is therefore recommended that academic outcomes are followed up at a later time point.

Interpretation

The primary aims of this evaluation were to assess the effect of Zippy's Friends, a universal social and emotional learning (SEL) programme, on pupil academic attainment (reading) and emotional self-regulation, as the two primary outcomes. The impact on self-regulated learning and social skills (secondary outcomes) was also assessed. The control and intervention groups were balanced on a range of school-level and pupil-level characteristics.

There was no evidence of an impact of Zippy's Friends on reading attainment, which was measured using a standardised attainment test (the Hodder Group Reading Test), administered and marked blind to group allocation. While pupils eligible for free school meals made small gains in reading attainment, the wide confidence interval associated with the estimate of the effect indicates no impact on reading attainment for these pupils. SEL programmes may be especially beneficial for disadvantaged children in particular, perhaps due to socio-economic gaps in childhood socio-emotional development (Dearden *et al.*, 2011). Other trials of Zippy's Friends have either not explored disadvantaged pupils as a specific subgroup (for example, Holen *et al.*, 2012), or drew their sample exclusively from schools in areas of disadvantage (such as Clarke *et al.*, 2014).

Other trials of SEL programmes have reported mixed findings in relation to reading attainment, with effect sizes from RCTs ranging from -0.14 (Hanson *et al.*, 2011) to +0.73 (Flay *et al.*, 2006). Meta-analytic studies of RCTs of SEL programmes on academic attainment have produced mean effect sizes of +0.20 (Corcoran *et al.*, 2018) and +0.27 (Durlak *et al.*, 2011). In the current trial, there was some evidence of a modest effect ($g = 0.05$, $p = 0.03$) on children's learning behaviour (including concentration and perseverance with school work, and cooperation in class activities), however this was measured based on teacher-report and is therefore subject to bias. Nevertheless, a tentative suggestion here is that the programme may impact on children's focus on learning in class, which, in the longer term, may benefit academic achievement.

Surprisingly, there was no evidence of an impact of the programme on emotional self-regulation or social skills—for the full sample or for the subgroup of FSM pupils. Emotional self-regulation was measured via self-report using the Children's Anger Management Scale, while social skills were measured using the self-report version of the Social Skills Improvement System (SSIS). It is possible that the young age of the children in this study precluded reliable self-reporting of both of these outcomes, as neither scale has been demonstrated to have been reliably and validly used when administered on a group (self-completion) basis with the same age group as in this sample. It is

important to note that the majority of control schools reported using SEL programmes or strategies with their class during the intervention year, and control group teachers reported spending approximately 55 minutes per week specifically on SEL activities. This is, perhaps, unsurprising, given that the control group had all expressed interest in delivering Zippy's Friends (and thus may be schools in which SEL is a priority). However, this may have contributed to attenuating programme effects on both academic and social and emotional outcomes.

The findings from the process evaluation indicate that Zippy's Friends was well received by schools and appeared to be implemented with high fidelity. Most teachers reported that Year 2 is a good year group to receive the programme, however issues with Key Stage 1 tests meant that towards the end of the year, some lessons may have been skipped or postponed. Some teachers suggested that beginning the programme mid-way through Year 1 and continuing it through the first half of Year 2 might work well in addressing this issue. The main barriers around implementation were time constraints and perceived lack of awareness from school leadership and the rest of the school. Teachers perceived the programme to benefit students in relation to their social and emotional outcomes, although this contradicted the results of the impact evaluation, which failed to find an impact on either social skills or emotional self-regulation. Potential explanations for such contradictory findings may include a lack of effectiveness of the programme (or lack of superiority over 'business as usual'), the high level of SEL activity implemented in the control group, the timing of outcome measurement which (immediately at the end of programme delivery) may have been too soon to detect change in outcomes, or problems with outcome measures.

Cost per pupil was low (averaging approximately £10 per pupil over three years), although this is likely to be an underestimate of the full cost of delivering the programme if attendance at support meetings and completing the graveyard visit is accounted for. Apart from attending the training (and support sessions), the time required by teachers to prepare for each session, and the additional resources required to deliver the programme, were minimal.

Limitations

It is important to note the potential limitations of this trial. The control group reported using a range of SEL programmes and strategies, and spent almost as long as the intervention group on SEL-related activities each week, which may have impacted on the ability of the study to detect any impact of the programme. Indeed, while it would appear that 95% of control schools had some form of SEL provision prior to randomisation, there appeared to be more reported use of SEL programmes at the end of the intervention period, perhaps indicative of the 'John Henry effect' (compensation rivalry).

There was a large amount of missing data on the actual implementation of the programme; this limits the extent to which judgements on fidelity can be made.

Only two schools in the intervention group did not go on to deliver the programme after allocation (however remained in the trial for evaluation purposes), indicating a high level of commitment and engagement among schools. Only one of the original 83 schools did not participate in post-test data collection. There were high levels of measurement attrition for some outcomes (as high as 28% for one of the primary outcomes, emotional self-regulation). While pupil self-report was preferable, the young age of the sample meant that identifying suitable and validated self-report measures of the outcomes was challenging. The extent of item-level missing data may reflect difficulties in collecting self-report data from pupils at the early end of primary-level education.

Future research and publications

Further analysis will be conducted on two additional subgroups—by (child) gender and English as an additional language status—and this will be published separately. Future research should consider longer-term follow up to determine whether any impact of the programme is seen in later years in terms

of both academic and social and emotional outcomes. As the data will be archived (with the Fisher Family Trust archive, in line with EEF policy), follow-up of academic outcomes would be feasible through linkage to the NPD.

References

Adi, Y., Killoran, A., Janmohamed, K. and Stewart-Brown, S. (2007) 'Systematic review of the effectiveness of interventions to promote mental well-being in primary schools: Universal approaches which do not focus on violence or bullying', London: National Institute of Clinical Excellence.

Allen, G. (2011) 'Early Intervention: The Next Steps. An Independent Report to Her Majesty's Government', London: Cabinet Office.

Altman, D. G. and Bland, J. M. (2005) 'Treatment allocation by minimisation', *British Medical Journal*, 330, p. 843.

Browne, G., Gafni, A., Roberts, J., Byrne, C. and Majumdar, B. (2004) 'Effective/efficient mental health programmes for school-age children: a synthesis of reviews', *Social Science and Medicine*, 58 (7), pp. 1,367–384.

CASEL (Collaborative for Academic, Social and Emotional Learning) (2015) 'Social and Emotional Learning Core Competencies', retrieved from: casel.org/social-and-emotional-learning/core-competencies [accessed 10 February 2016].

Chien, N. V., Harbin, V., Goldhagen, S., Lippman, L. and Walker, K. E. (2012) 'Encouraging the development of key life skills in elementary school-age children: a literature review and recommendations to the Tauck Family Foundation', Child Trends Working Paper, Publication 2012-28.

Clarke, A. M., Bunting, B. and Barry, M. M (2014) 'Evaluating the implementation of a school-based emotional wellbeing programme: a cluster randomised trial of Zippy's Friends for children in disadvantaged primary schools', *Health Education Research*, 29 (5), pp. 786–798.

Clarke, A. M., Morreale, S., Field, C.-A., Hussein, Y. and Barry, M. M. (2015) 'What works in enhancing social and emotional skills development during childhood and adolescence? A review of the evidence on the effectiveness of school-based and out-of-school programmes in the UK', report to the Early Intervention Foundation, England.

Corcoran, R. P., Cheung, A., Kim, E. and Xie, C. (2018) 'Effective universal school-based social and emotional learning programs for improving academic achievement: A systematic review and meta-analysis of 50 years of research', *Educational Research Review*, doi: 10.1016/j.edurev.12.991.

Dearden, L., Sibieta, L., Sylva, K. (2011) 'The socio-economic gradient in early child outcomes: Evidence from the Millennium Cohort Study', IFS Working Paper no. 11, 03, available at: <https://www.econstor.eu/bitstream/10419/47494/1/654793913.pdf>

Department of Health (2012) 'No health without mental health: Implementation framework', London: DoH.

Department for Education (2015) 'Mental health and behaviour in schools, Departmental advice for school staff', London: DfE.

Department for Education (2015) 'Schools, Pupils and their Characteristics: January 2015 – National Tables' (Table 3b), available at: <https://www.gov.uk/government/statistics/schools-pupils-and-their-characteristics-january-2015> [accessed 12 May 2017].

Durlak, J., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D. and Schellinger, K. B. (2011) 'The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions', *Child Development*, 82 (1), pp. 405–432.

- Education Endowment Foundation (2015) 'Intra-cluster correlation coefficients', available at: educationendowmentfoundation.org.uk/public/files/Evaluation/Writing_a_Protocol/ICC_2015.pdf
- Education Endowment Foundation (2017) 'Teaching and Learning Toolkit | Social and emotional learning', available at: educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/social-and-emotional-learning/
- Enders, C. K. (2010) *Applied missing data analysis*, New York, NY: Guildford Press.
- Evans, S., Royston, P. and Day, S. (2004) 'Minim: Allocation by minimisation in clinical trials', available at <https://www-users.york.ac.uk/~mb55/guide/minim.htm> [accessed 1 June 2016].
- Faupel, A. (2003) 'Emotional Literacy: Assessment and Intervention', UK: GL Assessment.
- Flay, B. R., Acock, A., Vuchinich, S. and Beets, M. (2006) 'Progress report of the randomised trial of Positive Action in Hawaii: End of third year of intervention', Positive Action.
- Green, H., McGinnity, A., Meltzer, H., Ford, T. and Goodman, R. (2005) 'Mental health of children and young people in Great Britain, 2004', London: HMSO.
- Greenberg, M. T. (2010) 'School-based prevention: current status and future challenges', *Effective Education*, 2, pp. 27–52.
- Gresham, F. M. and Elliott, S. N. (2008) *Social Skills Improvement System Rating Scales Manual*, Minneapolis, MN: NCS Pearson.
- Gutman, L. M. and Schoon, I. (2014) 'Preventative interventions for children and adolescents: a review of meta-analytic evidence', *European Psychologist*, 20, pp. 231–241.
- Hanson, T., Izu, J. A., Petrosino, A., DeLong-Cotty, B. and Zheng, H. (2011) 'A randomised experimental evaluation of the tribes learning communities prevention programme', available at: <https://www.ncjrs.gov/pdffiles1/nij/grants/237958.pdf>
- Holen, S., Waaktaar, T., Lervåg, A. and Ystgaard, M. (2012) 'The effectiveness of a universal school-based programme on coping and mental health: A randomised, controlled study of Zippy's Friends', *Educational Psychology*, 32 (5), pp. 657–677.
- Holen, S., Waaktaar, T., Lervåg, A. and Ystgaard, M. (2013) 'Implementing a universal stress management program for young school children: are there classroom climate or academic effects?', *Scandinavian Journal of Educational Research*, 57 (4), pp. 420–444.
- Humphrey, N., Hennessey, A., Wigelsworth, M., Lendrum, A., Pert, K., Joyce, C., Stephens, E., Wo, L., Squires, G., Woods, K., Calam, R., Harrison, M. and Turner, A. (2015) 'Promoting Alternative Thinking Strategies (PATHS): Evaluation report and executive summary', Education Endowment Foundation.
- Knapp, M., McDaid, D. and Parsonage, M. (2011) 'Mental health promotion and mental illness prevention: the economic case', London: PSSRU.
- Lazarus, R. S., and Folkman, S. (1984) *Stress, appraisal, and coping*, New York: Springer.
- Marshall, L., Rooney, K., Dunatchik, A. and Smith, N. (2017) 'Developing character skills in schools: Quantitative survey', London: NatCen Social Research.
- McDermott, P. A., Green, L. F., Francis, J. M., and Stott, D. H. (2001) *Learning Behaviors Scale*, Philadelphia: Edumetric and Clinical Science.

- Mishara, B. L. and Ystgaard, M. (2006) 'Effectiveness of a mental health promotion program to improve coping skills in young children: Zippy's Friends', *Early Childhood Research Quarterly*, 21, pp. 110–123.
- Polanczk, G. V., Salum, G. A., Sugaya, L. S., Caye, A. and Rohde, L. A. (2015) 'A meta-analysis of the worldwide prevalence of mental health disorders in children and adolescents', *Journal of Child Psychology and Psychiatry*, 56, pp. 345–365.
- Qualter, P., Gardner, K. J., Pope, D. J., Hutchinson, J. M. and Whiteley, H. E. (2012) 'Ability emotional intelligence, trait emotional intelligence, and academic success in British secondary schools: A 5 year longitudinal study', *Learning and Individual Differences*, 22 (1), pp. 83–91.
- Sandler, I. N., Tein, J.-Y., Mehta, P., Wolchik, S. and Ayers, T. (2000) 'Coping efficacy and psychological problems of children of divorce', *Journal of Clinical Psychology*, 29, pp. 336–347.
- Segal, J. (1983) 'Translating stress and coping research into public information and education', in M. Kessler and S. E. Goldston (eds), *A decade of progress in primary prevention*, Lebanon, New Hampshire: University Press of New England.
- Suveg, C., Hoffman, B., Zeman, J. L., & Thomassin, K. (2009). Common and specific emotion- related predictors of anxious and depressive symptoms in youth. *Child Psychiatry and Human Development*, 40, 223–239.
- Suveg, C. & Zeman, J. L. (2004). Emotion regulation in children with anxiety disorders. *Child Clinical and Adolescent Psychology*, 33(4), 750–759.
- Taylor, R. D., Oberle, E, Durlak, J. A. and Weissberg, R. P. (2017) 'Promoting Positive Youth Development Through School-Based Social and Emotional Learning Interventions: A Meta-Analysis of Follow-Up Effects', *Child Development*, 88 (4), pp. 1,156–171.
- Tennant, R., Goens, C., Barlow, J., Day, C. and Stewart-Brown, S. (2007) 'A systematic review of reviews of interventions to promote mental health and prevent mental health problems in children and young people', *Journal of Public Mental Health*, 6 (1), pp. 25–32.
- Vincent, D. and Crumpler, M. (2006) *Hodder Oral Reading Tests Manual* (2nd ed.), Hodder Education.
- Vincent, D. and Crumpler, M. (2007) *Hodder Group Reading Tests Manual* (2nd ed.), Hodder Education.
- Zeman J., Shipman K. & Suveg C. (2002). Anger and sadness regulation: Predictions to internalizing and externalizing symptoms in children. *Journal of Clinical Child & Adolescent Psychology*, 31, 393-398.

Appendix A: EEF cost rating

Cost ratings are based on the approximate cost per pupil per year of implementing the intervention over three years. More information about the EEF's approach to cost evaluation can be found [here](#). Cost ratings are awarded as follows:

Cost rating	Description
£ £ £ £ £	<i>Very low</i> : less than £80 per pupil per year.
£ £ £ £ £	<i>Low</i> : up to about £200 per pupil per year.
£ £ £ £ £	<i>Moderate</i> : up to about £700 per pupil per year.
£ £ £ £ £	<i>High</i> : up to £1,200 per pupil per year.
£ £ £ £ £	<i>Very high</i> : over £1,200 per pupil per year.

Appendix B: Security classification of trial findings

OUTCOME: Reading (HODDER)

Rating	Criteria for rating			Initial score	Adjust	Final score
	Design	Power	Attrition			
5	Well conducted experimental design with appropriate analysis	MDES < 0.2	0-10%			
4	Fair and clear quasi-experimental design for comparison (e.g. RDD) with appropriate analysis, or experimental design with minor concerns about validity	MDES < 0.3	11-20%	4	Adjustment for Balance [-1]	
3	Well-matched comparison (using propensity score matching, or similar) or experimental design with moderate concerns about validity	MDES < 0.4	21-30%			
2	Weakly matched comparison or experimental design with major flaws	MDES < 0.5	31-40%		Adjustment for threats to internal validity [-1]	2
1	Comparison group with poor or no matching (E.g. volunteer versus others)	MDES < 0.6	41-50%			
0	No comparator	MDES > 0.6	over 50%			

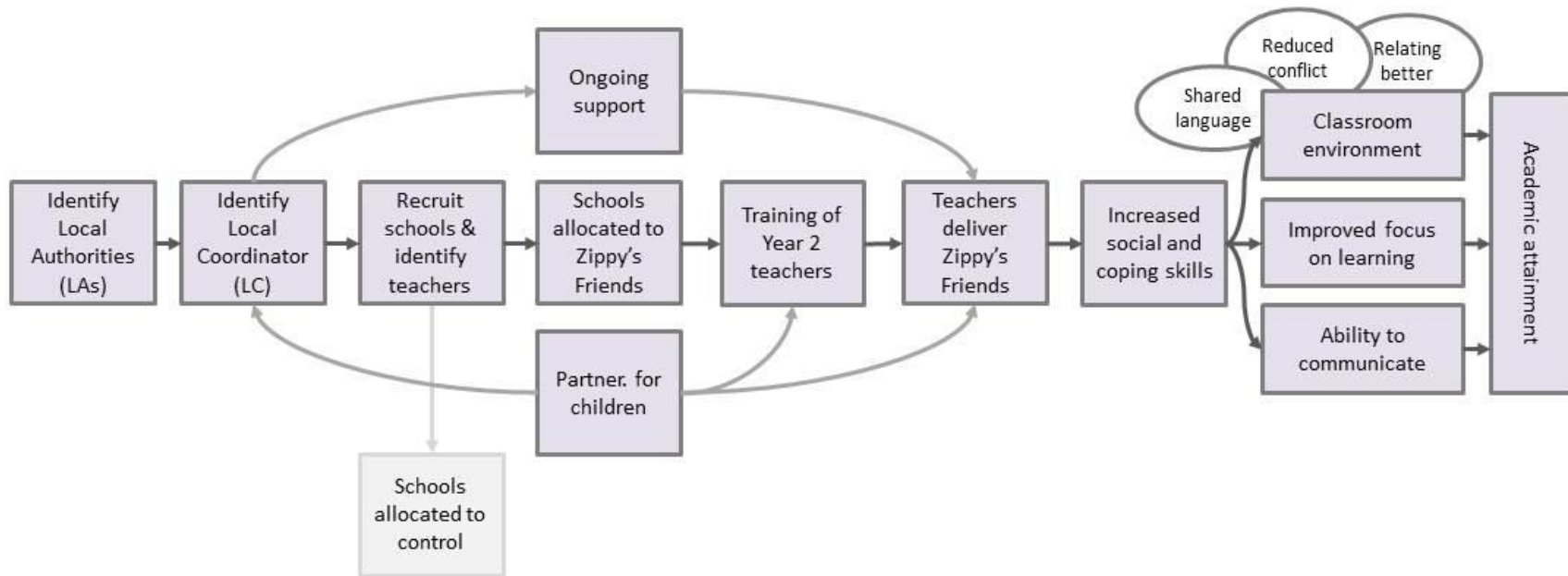
- **Initial padlock score:** lowest of the three ratings for design, power and attrition = [4] padlocks due to MDES=0.24 and Attrition rate of 15%
- **Reason for adjustment for balance** (if made): [-1] padlock due to an imbalance of 0.08
- **Reason for adjustment for threats to validity:** [-1] padlock due to evidence of compensatory rivalry in the control group
- **Final padlock score:** 3 padlocks

OUTCOME: Emotional Self-Regulation (CAMS)

Rating	Criteria for rating			Initial score	Adjust	Final score
	Design	Power	Attrition			
5	Well conducted experimental design with appropriate analysis	MDES < 0.2	0-10%		Adjustment for Balance [0]	
4	Fair and clear quasi-experimental design for comparison (e.g. RDD) with appropriate analysis, or experimental design with minor concerns about validity	MDES < 0.3	11-20%			
3	Well-matched comparison (using propensity score matching, or similar) or experimental design with moderate concerns about validity	MDES < 0.4	21-30%	3		Adjustment for threats to internal validity [-2]
2	Weakly matched comparison or experimental design with major flaws	MDES < 0.5	31-40%			
1	Comparison group with poor or no matching (E.g. volunteer versus others)	MDES < 0.6	41-50%			1
0	No comparator	MDES > 0.6	over 50%			

- **Initial padlock score:** lowest of the three ratings for design, power and attrition = [3] padlocks due to 28% attrition
- **Reason for adjustment for balance** (if made): No adjustment necessary due to good balance
- **Reason for adjustment for threats to validity:** [-2] padlock due to evidence of compensatory rivalry in the control group, and suitability of measure for this age group.
- **Final padlock score:** 2 padlocks

Appendix C: Theory of change



Appendix D: Summary of previous evaluations of Zippy's Friends

Table D1: Summary of previous evaluations of Zippy's Friends

Author	Country	Sample age	Sample size	Study design	Measure	Outcome (effect size)
Mishara & Ystgaard, 2006	Lithuania	6-7 years	418 pupils; 17 schools	QED	Social Skills Questionnaire, Teacher Form (Gresham & Elliott, 1990)	Cooperation (0.12 [*])
						Assertion (0.48 ^{***})
						Self-control (0.58 ^{***})
					Social Skills Questionnaire, Student Form (Gresham & Elliott, 1990)	Cooperation
						Assertion
	Teacher problem behaviour scales	Self-control (0.44 [*])				
		Empathy (0.44 [*])				
		Externalising (-0.17 ^{***})				
	Denmark	7-8 years	432 pupils; 18 schools	QED	Social Skills Questionnaire, Teacher Form (Gresham & Elliott, 1990)	Internalising
						Hyperactivity (-0.17 ^{***})
Number of strategies (0.47 ^{***})						
Social Skills Questionnaire, Student Form (Gresham & Elliott, 1990)					Number of strategies x helpfulness [*]	
					Cooperation	
Holen et al., 2012, 2013	Norway	7-8 years	766 pupils; 48 schools	RCT	Kidcope (Spirito et al., 1998), self-report	Cooperation
						Assertion
						Self-control
					Kidcope (Spirito et al., 1998), parent-report	Cooperation
						Assertion
Kidcope (Spirito et al., 1998), self-report	Self-control					
	Empathy					
	Active/emotional regulation					
	Oppositional coping (-0.38 [*])					
Kidcope (Spirito et al., 1998), parent-report	Withdrawal					
	Emotional regulation					
	Active (0.19 ^{**})					
	Withdrawal/ oppositional					
Kidcope (Spirito et al., 1998), self-report	Self-awareness (0.35 ^{**})					
	Conduct problems					

					Strengths and Difficulties Questionnaire (Goodman, 1997), teacher-report	Emotional symptoms
						Hyperactivity
						Peer relationships
						Prosocial behaviour
						Impact of mental health difficulties (-0.15*)
					Strengths and Difficulties Questionnaire (Goodman, 1997), parent-report	Conduct problems
						Emotional symptoms
Hyperactivity						
<i>Need to add measures</i>	Peer relationships					
	Prosocial behaviour					
	Reduction in bullying (0.55**)					
<i>Need to add measures</i>	Social integration (0.33**)					
<i>Need to add measures</i>	Academic skills (0.42**)					
Clarke et al., 2014	Ireland	7-8 years	766 pupils; 44 schools	RCT	Emotional Literacy Checklist (Faupel, 2003), teacher-report	Social skills (0.12****)
						Self-regulation (0.12**)
						Motivation (0.13****)
						Self-awareness (0.35****)
						Conduct problems (0.14*)
					Strengths and Difficulties Questionnaire (Goodman, 1997), teacher-report	Emotional symptoms (-0.11)
						Hyperactivity (-0.04)
						Peer relationships (0.09)
						Prosocial behaviour (0.08)
						Self-awareness (0.31**)
Hegeman, 2015	Netherlands	6-7 years	1177 pupils; 29 schools	RCT	Levels of Emotional Awareness for Children (Bajgar & Lane, 2004), pupil-report	Awareness of others (0.26**)
						Total awareness (0.27)
						Adaptive coping (0.17****)
					Social Information Processing Scale (Orobio et al., 2004), pupil-report	Hostile attribution (-0.04)
						Self-awareness (0.04)
					Emotional Literacy Checklist (Faupel, 2003), teacher-report	Self-regulation (-0.08)
						Motivation (0.21)
						Empathy (-0.01)
						Social skills (0.03)
						Self-awareness (1.98)
Emotional Literacy Checklist (Faupel, 2003), parent-report	Self-regulation (0.26)					
	Motivation (0.12)					

					Empathy (0.04)
					Social skills (-0.02)
				Preschool and Kindergarten Behaviour Scales (Merrel, 1996), teacher-report	Egocentric/explosive behaviour (0.02)
					Inattention/hyperactivity (-0.04)
					Antisocial/aggressive behaviour (0.05)
					Socially withdrawn (0.25)
					Anxiety/somatic problems (0.12)
					Externalising behaviour (-0.55)
					Internalising behaviour (0.19)
					Social cooperation (0.14)
					Social interaction (-0.1)
					Social independence (-0.07)
				Preschool and Kindergarten Behaviour Scales (Merrel, 1996), parent-report	Egocentric/explosive behaviour (0.18)
					Inattention/hyperactivity (-0.37*)
					Antisocial/aggressive behaviour (-0.37*)
					Socially withdrawn (-0.07)
					Anxiety/somatic problems (0.01)
					Externalising behaviour (-0.31*)
				Internalising behaviour (-0.03)	
				Head Start Competence Scale (Domitrovich, 2001), teacher-report	Emotional regulation (0.13)
				Head Start Competence Scale (Domitrovich, 2001), parent-report	Emotional regulation (0.23)

*p<0.05; **p<0.01; ***p<0.001

Appendix E: Memorandum of Understanding for schools



Memorandum of Understanding *Zippy's Friends* Evaluation

School Name: _____ School URN: _____

This is a memorandum of understanding between Partnership for Children (the programme deliverer), the Centre for Effective Education at Queen's University Belfast (the evaluator) and the school.

Thank you for your interest in this evaluation of *Zippy's Friends*, which is being carried out by the Centre for Effective Education (CEE) at Queen's University Belfast, with funding from the Education Endowment Foundation (EEF). This document provides you with details about the evaluation and what participation will involve for your school.

Overview of the programme

Zippy's Friends is a universal programme that aims to develop children's repertoire of coping skills and their ability to adapt those coping skills to various situations. Further, it integrates problem solving skills, social skills and emotional literacy as skills that may facilitate adaptive coping behaviour. It is delivered to Year 2 pupils by their class teacher over the course of the school year, through 24 weekly sessions which last around 45 minutes. The 24 sessions are divided into six modules, each of which is centred around a set of illustrated stories about a group of children, their families, friends and Zippy, a pet stick insect. The modules focus on particular themes: feelings, communication, making and breaking relationships, conflict resolution, dealing with change and loss and general coping skills. Teachers are provided with a class set of teaching and children's materials and receive training and support, which is organised by Partnership for Children and a local coordinator. *Zippy's Friends* has already been taught to over one million children around the world, and research in other countries have found that children enjoy the sessions and show significant improvements in their social and emotional skills. This will be the first large-scale evaluation of the programme in the UK.

The evaluation

Approximately 70 schools will participate in this evaluation. The evaluation involves a randomised controlled trial, meaning that half of participating schools will be randomly allocated to deliver the programme (the 'intervention group') and the other half will not deliver the programme and will instead continue as they normally would (the 'control group'). Random allocation is the best way of separating out the effect of the programme from other factors which impact on children's learning and development. To ensure that this allocation is genuinely random, it is important that schools commit to the trial before it

takes place. Schools in the control group will receive a payment of £1,000 at the end of the trial period, which can be used to purchase *Zippy's Friends*, if they wish, for the 2017/18 school year.

Schools who sign up to the trial will be asked to do the following:

1. Parental consent for pupil participation in the evaluation

We would like you to send home the enclosed parental information and **opt-out** consent forms to parents of pupils who will be in a Year 2 class in the 2016/17 school year. This explains to parents what the evaluation involves and gives them an opportunity to opt their child out of the evaluation by simply signing and returning the opt-out consent form to their class teacher (although any child who is opted out will still participate in *Zippy's Friends* lessons; they will simply not be included in any data collection for the trial).

2. Pupil details and baselines assessments (April – June 2016)

The school will be asked to provide the CEE evaluation team with the following details for all participating Year 2 pupils: UPN, first name, surname, date of birth, and whether English is an Additional Language. The CEE evaluators will provide you with information on how to safely send this information through the University secure network. This information will enable the research team to link to the National Pupil Database records where we can obtain data on pupil gender and Free School Meal eligibility. We will use this information to find out whether the programme works differently for certain groups of pupils, such as boys and girls, those entitled to Free School Meals and those with English as an Additional Language.

We are interested in testing the impact of *Zippy's Friends* on pupil socio-emotional and academic outcomes. It is therefore important that we assess both of these outcomes before the programme begins and again when the programme is finished. The baseline assessments are to be completed during the last term of Year 1 (April - June 2016). There are two assessments per pupil (and QUB will post the school enough copies of each):

- Hodder Oral Reading Test, which takes about 3-4 minutes per pupil to administer and can be administered by the class teacher or teaching assistants;
- Teacher questionnaire on pupil social and emotional wellbeing, which takes about 10-15 minutes per pupil and should be completed by the class teacher (based on a class of 30, this would take between 5 and 7 hours in total between April and June 2016).

3. Random allocation of schools to deliver Zippy's Friends (end of June 2016)

The CEE evaluators will randomly allocate participating schools to either the **intervention group** (which will receive *Zippy's Friends* programme) or the **control group** (which will not receive the programme and will continue as normal). To be eligible to be entered into the randomisation process, schools need to complete actions 1 and 2 (above) and return this information to the evaluators, along with the signed Memorandum of Understanding, prior to randomisation.

4. Teacher training (September – October 2016)

Year 2 teachers in the schools allocated to the intervention group will attend one whole day's training in September/October 2016, delivered by Partnership for Children. Schools must meet the cost of supply cover for that day, but both the training and the teaching materials which will be supplied to teachers will be free of charge.

5. *Delivery of Zippy's Friends (October 2016 onwards)*

Year 2 teachers in schools allocated to the intervention group will deliver *Zippy's Friends* to their class over the course of the school year. This involves delivering 24 weekly sessions which each last around 45 minutes.

6. *On-going teacher support*

Partnership for Children will provide on-going support for teachers throughout the duration of the programme. The local coordinator will visit each class at least once during the autumn term 2016 to observe a *Zippy's Friends* session, and will hold three follow-up meetings for all *Zippy's Friends* teachers in the area during the course of the programme: 1) at the end of Module 2 (January 2017), 2) at the end of Module 4 (end of March 2017) and 3) at the end of Module 6 (June 2017). Teachers must commit to attending all three meetings. The evaluators will attend some of these meetings/sessions in order to better understand this element of implementation.

7. *Outcome data collection*

To explore the impact of *Zippy's Friends* on outcomes for Year 2 pupils, the evaluators will collect outcome data at the end of the 2016/17 school year, **in both the intervention and control groups**. Pupils will be asked to complete a group reading test, followed by a short questionnaire that measures their social skills and emotional self-regulation. Trained fieldworkers will visit the school to administer both the test and questionnaire on a whole class basis, which should take between 1 hour and 1 hour 15 minutes. They will make contact with class teachers ahead of time to arrange a day and time that suits the school best. The evaluators will directly oversee this process and at all times will try to minimise disruption to the class. All schools will be provided with the results of the Year 2 reading tests.

During this term, teachers will also be asked to complete a short (less than 5 minutes per pupil) questionnaire on each pupil's learning behaviour in class.

8. *Teacher survey*

During the 2016/17 school year, all teachers delivering *Zippy's Friends* will be asked to complete a short online survey, to capture teachers' experiences of delivering the programme. Year 2 teachers in control schools will also be asked to fill in a questionnaire on standard practice.

9. *School interviews and observations*

Your school may be selected to take part in a small number of interviews with one of the evaluation team. We aim to carry out interviews with the school Principal, Year 2 teachers, and a selection of parents to further explore overall perceptions of *Zippy's Friends* and its impact on the class. The evaluators would also like to talk to pupils about their experiences, through focus groups, and observe a lesson in practice. If your school is selected to participate in this part of the evaluation, you will be contacted by the evaluators from January 2017 to discuss your participation. Conducting these interviews may require a short visit to the school by the research team; however the timing will be agreed with the school and teacher(s) to ensure disruption is minimised.

Will participation in the study be kept confidential?

All data collected will be treated with the strictest confidence. The responses will be collected by the research team. For the purpose of the study, information provided will be linked with the National Pupil Database (held by the Department for Education), other official records, and shared with CEE evaluation

team, the Department for Education, Education Endowment Foundation (EEF), EEF's data contractor FFT Education and in an anonymised form to the UK Data Archive. No individual school or pupil will be identified in any reports or presentations arising from the research.

Who to contact for more information

If you have any questions regarding the programme or the evaluation, please do not hesitate to contact:

Lizzie Poulton

Programme Manager

Partnership for Children
26-27 Market Place
Kingston upon Thames
Surrey
KT1 1JH
Tel: 020 8974 6004
Email: lizzie.poulton@
partnershipforchildren.org.uk

Aideen Gildea

Trial Manager

Centre for Effective Education
School of Education
Queen's University Belfast
69-71 University Street
Belfast
BT7 1HL
Tel: 02890975193
Email: a.gildea@qub.ac.uk

Dr Seaneen Sloan

Principal Investigator

Centre for Effective Education
School of Education
Queen's University Belfast
69-71 University Street
Belfast
BT7 1HL
Tel: 02890975981
Email: s.sloan@qub.ac.uk

Signatures

We, undersigned, commit to participate in the evaluation of *Zippy's Friends*, as detailed above.

FOR THE PROGRAMME DELIVERY TEAM

Partnership for Children Director:

Name:

Signature:

Date:

FOR THE EVALUATOR

Principal Investigator:

Name: Dr Seaneen Sloan

Signature:



Date: 24/02/2016

FOR THE SCHOOL

Headteacher:

Name:

Signature:

Date:

Appendix E: Parent information and consent form



Centre for Effective Education
School of Education
Queen's University Belfast
69-71 University Street
Belfast BT7 1HL

Dear parent or guardian,

Re: Evaluation of Zippy's Friends social and emotional learning programme

We are a research team from the Centre for Effective Education at Queen's University Belfast. We are writing to you because your child's school has signed up to be part of an evaluation of a programme called *Zippy's Friends*. Your child's class will be taking part in this evaluation, but it is up to you whether you want your child to participate. Before you make a decision, it is important for you to know why this evaluation is being carried and what it will involve. If you **do not** want your child to take part in the evaluation, we ask that you to sign the attached 'opt-out' consent form and return it to your child's teacher.

Please take your time to read the following information carefully and discuss it with others if you wish.

Why are we doing this study?

Zippy's Friends has already been taught to over one million children around the world, and research in other countries have found that children enjoy the sessions and show significant improvements in their social and emotional skills. This study will be the first large-scale evaluation of the programme in the UK. We want to find out if *Zippy's Friends* has an impact on children's learning, as well as their social and emotional development. *Zippy's Friends* is delivered by the class teacher to the whole class, and is based on a series of stories about a group of children, their families, friends and Zippy, a pet stick insect. The programme covers feelings, communication, making and breaking relationships, conflict resolution, dealing with change and loss and general coping skills. There are 24 weekly sessions which last around 45 minutes each. The evaluation is being funded by the Education Endowment Foundation and has been approved by the Research Ethics Committee at the School of Education, Queen's University Belfast. Approximately 70 primary schools across England are taking part in this evaluation. Half of these schools will deliver *Zippy's Friends* with Year 2 pupils during the 2016/17 school year. We will randomly decide which schools will deliver the programme (a bit like tossing a coin). The remaining schools will not deliver the programme in 2016/17 and will instead continue with their normal curriculum, but will have the opportunity to run *Zippy's Friends* in 2017/18.

What will happen if my child takes part?

If you consent for your child to take part, we will ask their teacher to complete a short reading test with them at the end of Year 1 (between April and June 2016), and provide us with some pupil details (i.e. their name, date of birth and first language). We will also ask their teacher to complete a short questionnaire on your child's social and emotional skills.

If your child's school is selected to deliver the programme, their class teacher will deliver *Zippy's Friends* throughout the 2016/17 school year.

At the end of Year 2, we will ask the whole class to complete a reading test (the same type of test that schools normally do at the end of each school year), and also a short questionnaire on their social and

emotional skills, e.g. about friendships and social skills. By collecting the same information from both the classes who receive *Zippy's Friends*, and classes who don't, we can see whether or not the programme has an effect on children's learning and development.

As well as this, we will also visit a small number of schools to ask some of the children about whether they enjoy *Zippy's Friends*. We would do this through a short focus group discussion with about 6-8 children. If your child's school is selected for this part of the evaluation, we will write to you again and ask for your consent for your child to be included; we will also ask your child if they are happy to talk to us.

How long will the evaluation last?

This study will begin in the summer term 2016 and will finish in July 2017, at the end of Year 2. The findings from the evaluation will be openly available (<https://educationendowmentfoundation.org.uk>) from about February 2018.

Does my child have to take part?

If your school is selected to take part in *Zippy's Friends* your child will participate in the lessons but it is up to you whether your child takes part in the evaluation. If you do NOT want your child to take part, please complete the 'opt-out consent form' attached and return it to their class teacher. Your child will also be given the opportunity to decide whether they want to take part in the evaluation, and we will explain this to your child.

You can change your mind at any time during the evaluation and decide that you do not want your child to take part any more. You do not even have to give a reason why. Your decision to take part will in no way impact on your relationship with any of the organisations involved (the school, Queen's University Belfast, Partnership for Children and the Education Endowment Foundation) now or in the future.

What will happen to the information collected as part of the evaluation?

All information which is collected during the course of the research will be kept strictly confidential. Research folders will be kept in a locked office in the Centre for Effective Education, Queen's University Belfast, at all times. Access to these folders will be restricted to the evaluation team. Any information that is stored electronically will be kept locked by password access and access restricted to the research team only.

For the purpose of the study, information provided will be linked with the National Pupil Database (held by the Department for Education), other official records, and shared with the Centre for Effective Education, the Department for Education, Education Endowment Foundation (EEF), EEF's data contractor FFT Education and in an anonymised form to the UK Data Archive. **We will not use your child's name, their teacher's name or the name of the school in any report arising from the research.**

How to contact us to find out more about the evaluation

If you would like to find out more about the evaluation (even if you decide not to take part), please contact one of the evaluators:

Dr Seaneen Sloan (Principal Investigator)
Telephone: 028 9097 5981
E-mail: s.sloan@qub.ac.uk

Aideen Gildea (Trial Manager)
Telephone: 028 9097 5193
E-mail: a.gildea@qub.ac.uk



Zippy's Friends Programme Evaluation
Parent Opt-Out Consent Form

Please return this slip to your child's class teacher if you DO NOT want your child to take part in the evaluation of Zippy's Friends.

I **do not** give my consent for my child to take part in the evaluation of Zippy's Friends

Your Name: _____

Your Child's Name: _____

Primary School Name: _____

Signed _____

Date _____

Appendix F: Randomisation procedure

Allocation was undertaken using the programme Minim. This allocation method uses minimisation to create balanced groups (see for example Altman & Bland, 2005). The following information was provided to Minim to set up the random allocation procedure:

1. The factors used in the minimisation process:
 - a. School-level attainment, based on percentage of pupils attaining Level 4 or above in Reading, Writing and Maths at Key Stage 2 in 2015. Each school was then classified as having low, medium or high levels of attainment (using internally derived cut-offs). Taking the percentage of pupils attaining Level 4 or above in Reading, Writing and Maths at Key Stage 2, the sample was split into three equal groups such that:
 1. Low attainment: $\leq 76\%$
 2. Medium attainment: $\geq 77\%$ and $< 86\%$
 3. High attainment: $\geq 86\%$
 - b. Local authority area (5 areas)
 1. Nottingham
 2. Cheshire
 3. Lincolnshire
 4. Dudley
 5. Croydon
2. List of treatment groups: Control (0) and Intervention (1), allocated on a 1:1 ratio.

The following screenshots document allocation at set up and mid-way through the process:

```

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MINIM
CURRENT STATE OF TRIAL
ZIPPYS FRIENDS
ATTAINMENT:
  Intervention  Control
  LOW          0          0
  MEDIUM      0          0
  HIGH         0          0
AREA:
  NOTTINGHAM  0          0
  CHESHIRE    0          0
  LINCOLNSHIRE 0          0
  DUDLEY      0          0
  CROYDON     0          0
Totals :      0          0
Grand Total : 0
Is that all correct ? Y
Please type Y or N
  
```



```

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MINIM
ALLOCATION OF THIS PATIENT
Patient allocated to treatment Control.
Do you have details of a patient ? Y

Please type Y or N

```

```

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MINIM
CURRENT STATE OF TRIAL
ZIPPYS FRIENDS
          Intervention      Control
ATTAINMENT:
LOW              2              2
MEDIUM          0              0
HIGH             1              2
AREA:
NOTTINGHAM      0              0
CHESHIRE        0              0
LINCOLNSHIRE    0              0
DUDLEY          0              1
CROYDON         3              3
Totals :        3              4
Grand Total :   7
Do you want to stop ?

Please type Y or N

```

```

DOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Program: MINIM
CURRENT STATE OF TRIAL
ZIPPYS FRIENDS
          Intervention      Control
ATTAINMENT:
LOW              8              9
MEDIUM          9              9
HIGH             6              7
AREA:
NOTTINGHAM      2              3
CHESHIRE        5              5
LINCOLNSHIRE    5              6
DUDLEY          5              5
CROYDON         6              6
Totals :        23             25
Grand Total :   48
Do you want to stop ?

Please type Y or N

```

Appendix G: Analysis code

**Complete Case Analysis

```
//Primary outcome 1 - reading
mixed HGRT_RawScore intervention z_HORT, || school_id:, covariance(independent) mle
variance
```

```
//Primary outcome 2 - emotional self-regulation
mixed CAMS_post intervention z_ESR_pre, || school_id:, covariance(independent) mle
variance
```

```
//Secondary outcome 1 - self-regulated learning
mixed SRL_post intervention z_SRL_pre, || school_id:, covariance(independent) mle
variance
```

```
//Secondary outcome 2 - social skills
mixed social_skills_post intervention z_social_skills_pre, || school_id:,
covariance(independent) mle variance
```

**Subgroup Analysis (FSM)

```
//Primary outcome 1 - reading
mixed HGRT_RawScore intervention z_HORT if fsm==1, || school_id:,
covariance(independent) mle variance
```

```
//Primary outcome 2 - emotional self-regulation
mixed CAMS_post intervention z_ESR_pre if fsm==1, || school_id:,
covariance(independent) mle variance
```

```
//Secondary outcome 1 - self-regulated learning
mixed SRL_post intervention z_SRL_pre if fsm==1, || school_id:,
covariance(independent) mle variance
```

```
//Secondary outcome 2 - social skills
mixed social_skills_post intervention z_social_skills_pre if fsm==1, || school_id:,
covariance(independent) mle variance
```

**On-treatment Analysis

```
//Primary outcome 1 - reading
mixed HGRT_RawScore dosage z_HORT, || school_id:, covariance(independent)
mle variance
```

```
//Primary outcome 2 - emotional self-regulation
mixed CAMS_post dosage z_ESR_pre, || school_id:, covariance(independent)
mle variance
```

```
//Secondary outcome 1 - self-regulated learning
mixed SRL_post dosage z_SRL_pre, || school_id:, covariance(independent) mle
variance
```

```
//Secondary outcome 2 - social skills
mixed social_skills_post dosage z_social_skills_pre, || school_id:,
covariance(independent) mle variance
```

Appendix H: Multilevel models

Table H1: Standard deviation used to calculate effect size for each outcome

Outcome	Pooled standard deviation
Reading	16.24
Emotional self-regulation	11.63
Self-regulated learning	8.85
Social skills	32.81

Table H2: Unadjusted and adjusted ICC for each outcome

	Unadjusted ICC	Adjusted ICC
Reading	0.07	0.09
Emotional self-regulation	0.03	0.04
Self-regulated learning	0.07	0.15
Social skills	0.06	0.07

Table H3: Reading

	Complete case analysis			Multiple imputation		
	Coef.	S.E.	p-value	Coef.	S.E.	p-value
Group (intervention)	-0.25	0.41	0.54	-0.33	0.39	0.39
Pre-test	5.98	0.09	<0.001	6.05	0.94	<0.001
Constant	29.84	0.29	<0.001	29.79	0.28	<0.001
Obs	3312			3904		
Log Likelihood	-10216.412			-		
School-level variance	2.60 (S.E.=0.54)			-		
Pupil-level variance	26.97 (S.E.=0.67)			-		

Table H4: Emotional self-regulation

	Complete case analysis			Multiple imputation		
	Coef.	S.E.	p-value	Coef.	S.E.	p-value
Group (intervention)	-0.21	0.46	0.65	-0.27	0.41	0.52
Pre-test	1.06	0.15	<0.001	1.09	0.15	<0.001
Constant	30.80	0.33	<0.001	30.83	0.30	<0.001
Obs	2810			3904		
Log Likelihood	-9660.1457			-		
School-level variance	2.40 (S.E.=0.70)			-		
Pupil-level variance	55.28 (S.E.=1.50)			-		

Table H5: Self-regulated learning

	Complete case analysis			Multiple imputation		
	Coef.	S.E.	p-value	Coef.	S.E.	p-value
Group (intervention)	0.74	0.33	0.03	0.78	0.28	0.01
Pre-test	2.40	0.06	<0.001	2.39	0.07	<0.001
Constant	32.33	0.23	<0.001	32.29	0.20	<0.001
Obs	3087			3904		
Log Likelihood	-8032.3967			-		
School-level variance	1.77 (S.E.=0.34)			-		
Pupil-level variance	10.15 (S.E.=0.26)			-		

Table H6: Social skills

	Complete case analysis			Multiple imputation		
	Coef.	S.E.	p-value	Coef.	S.E.	p-value
Group (intervention)	-2.16	1.51	0.15	-2.21	1.35	0.10
Pre-test	3.85	0.48	<0.001	4.28	0.41	<0.001
Constant	155.83	1.08	<0.001	155.67	0.95	<0.001
Obs	2155			3904		
Log Likelihood	-9547.6676			-		
School-level variance	28.08 (S.E.=7.36)			-		
Pupil-level variance	397.69 (S.E.=12.35)			-		

Appendix I: Implementation survey responses

Table I1: Responses to implementation survey completed by 46 teachers at post-test

	Strongly agree	Agree	Slightly agree	Slightly disagree	Disagree	Strongly disagree
				%		
1. I feel I am suitably qualified to deliver this programme	52	36	2	2	-	-
2. I am enthusiastic that my school has adopted this programme	52	36	10	2		
3. I don't feel this type of programme is necessary	-	5	-	7	29	59
4. The programme complements the existing curriculum/programme of work	26	55	14	-	5	-
5. The programme is suitable for all the children regardless of ability/stage of development	31	50	5	10	5	-
6. I find there is not enough time to deliver all aspects of the programme in a given day/session	2	19	19	10	38	12
7. The programme materials and resources are useful	50	48	2	-	-	-
8. I think a different type of programme may have been more suitable for the children's needs	-	2	10	21	48	19
9. I am confident in the delivery of the programme following the training I have received	60	33	2	2	2	-
10. There is not sufficient support in the ongoing delivery of the programme	-	7	-	14	33	45
11. The room where the programme is delivered is adequate for all activities	24	57	2	2	5	10
12. I find the programme content is too prescriptive	-	7	14	33	45	-
13. The programme instructions are very clear	57	41	2	-	-	-
14. The programme misses important elements that I feel it should have covered	-	2	10	10	64	14
15. I think the programme goals are realistic	26	69	-	2	2	-
16. I feel the programme is very beneficial for the children	45	36	19	-	-	-
17. Other activities and demands get in the way of delivering the programme as intended	10	12	27	12	37	2
18. I don't think the programme frequency (intensity) is sufficient to impact on the children's development	-	2	19	17	45	17
19. I think that the delivery of the programme has had a positive impact on my job/role	17	43	31	5	5	-
20. The working environment in my school is not suitable for the delivery of this programme	-	7	-	2	43	48
21. The programme engages and motivates the children	36	50	14	-	-	-
22. The senior management in the school are keen to see the programme work	17	48	19	7	7	2
23. I feel engaged and excited with the programme	33	55	7	5	-	-
24. I feel the programme is an important addition to the provision in this school	38	48	10	5	-	-

	Strongly agree	Agree	Slightly agree	Slightly disagree	Disagree	Strongly disagree
				%		
25. I am more aware of the children's feelings as a result of the programme	17	60	21	2	-	-
26. I have better engagement/collaboration with parents as a result of the programme	5	5	41	26	19	5
27. My relationship with the children has improved	7	45	38	5	2	2
28. I am more aware of listening to the children as a result of the programme	10	52	29	5	-	5
29. The programme has made a difference to the atmosphere in the classroom	7	36	52	5	-	-
30. The programme has given me a structure/ support to help children cope with difficult situations	31	55	12	-	2	-
31. I feel that my teaching has improved as a result of the programme	5	19	48	17	12	-
32. I would be happy for this school to continue with this programme in subsequent years	55	33	7	2	2	-

Appendix J: Distribution of primary outcomes

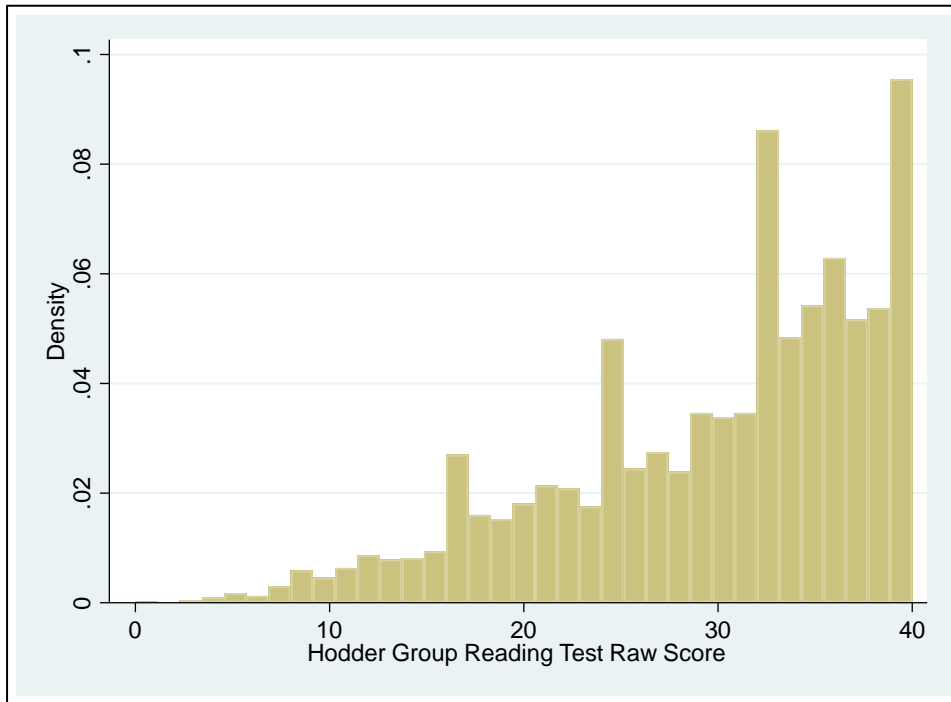


Figure J.1: Histogram of Hodder Group Reading Test (post-test) raw scores

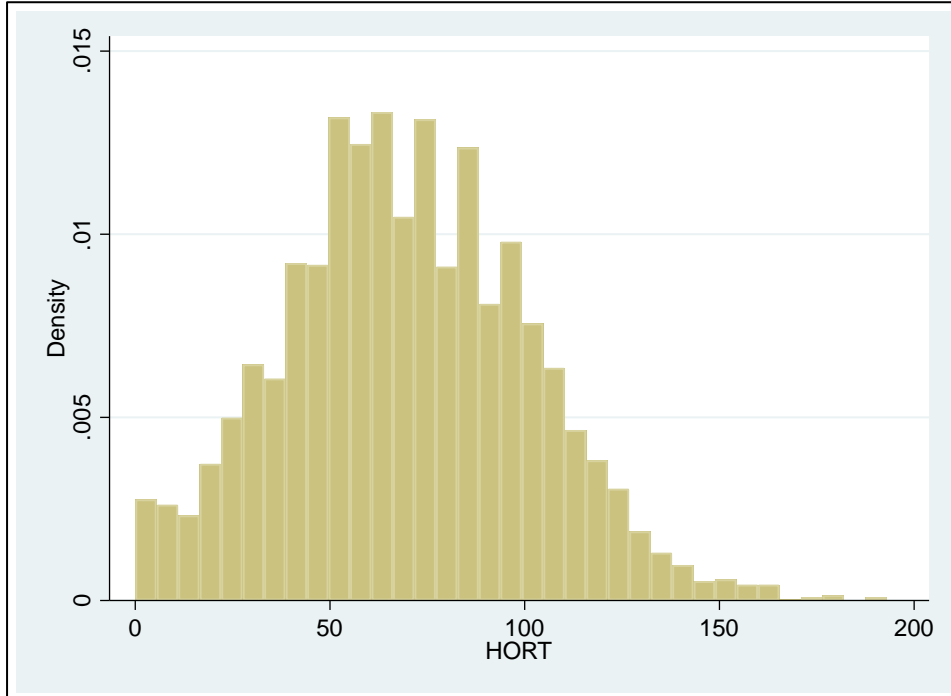


Figure J.2: Histogram of Hodder Oral Reading Test (pre-test) raw scores

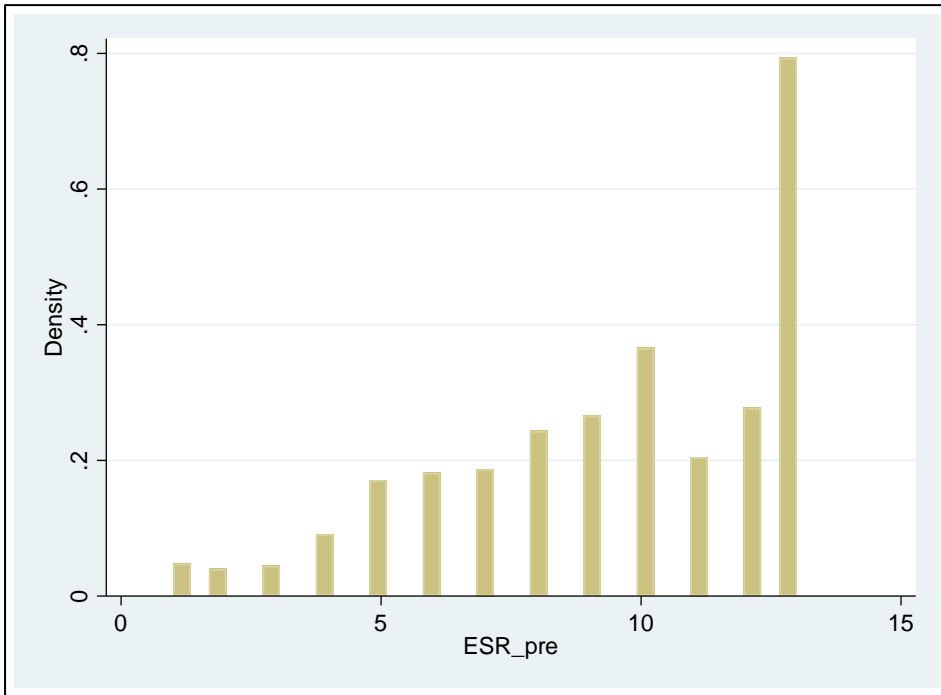


Figure J.3: Histogram of Emotional Self-Regulation (pre-test) raw scores

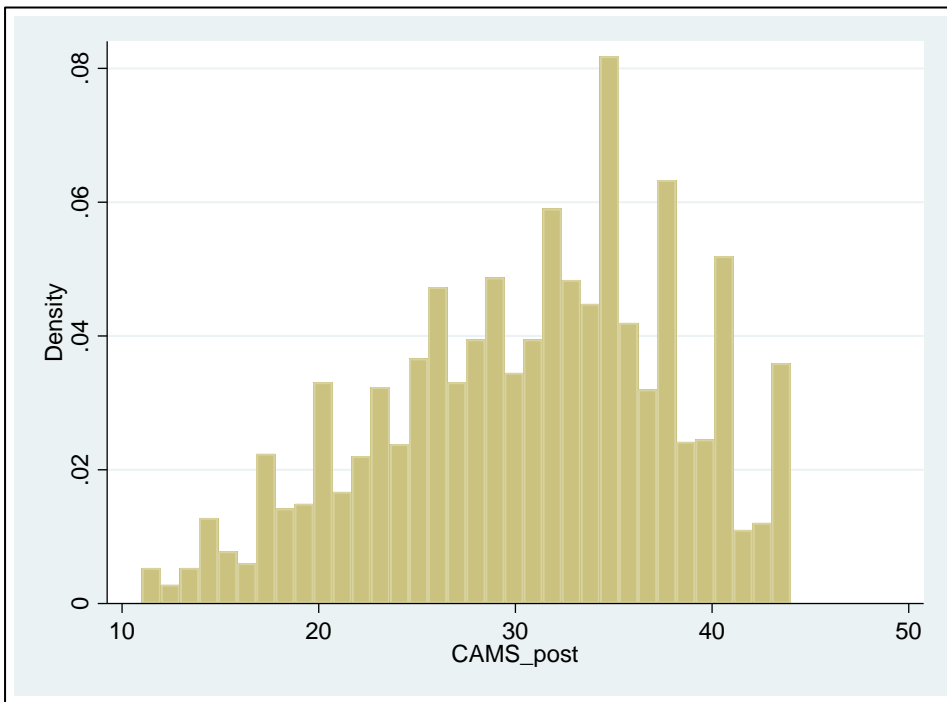


Figure J.4: Histogram of Emotional Self-Regulation (post-test) raw scores

You may re-use this document/publication (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0.

OGL This information is licensed under the Open Government Licence v3.0. To view this licence, visit <http://www.nationalarchives.gov.uk/doc/open-government-licence/>

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

This document is available for download at www.educationendowmentfoundation.org.uk



Education
Endowment
Foundation

The Education Endowment Foundation
9th Floor, Millbank Tower
21-24 Millbank
London
SW1P 4QP
www.educationendowmentfoundation.org.uk