Evaluation of Zippy's Friends for improving socio-emotional and academic outcomes in 6-7 year olds: Protocol for a cluster randomised controlled trial and process evaluation

Published in:
International Journal of Educational Research

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Title: Evaluation of Zippy's Friends for improving socio-emotional and academic outcomes in six to seven year olds: Protocol for a cluster randomised controlled trial and process evaluation

Author names and affiliations

Seaneen Sloan[1]   email: seaneen.sloan@ucd.ie

Aideen Gildea[2]   email: a.gildea@qub.ac.uk

Sarah Miller[2]   email: s.j.miller@qub.ac.uk

Lizzie Poulton[3]   email: lizzie.poulton@partnershipforchildren.org.uk

Caroline Egar[3]   email: caroline.egar@partnershipforchildren.org.uk

Allen Thurston[2]   email: a.thurston@qub.ac.uk

[1]School of Education, University College Dublin, Dublin, Ireland

[2]School of Social Sciences, Education and Social Work, Queen's University Belfast, Northern Ireland, UK

[3]Partnership for Children, Kingston-upon-Thames, England, UK

Corresponding author: Seaneen Sloan, School of Education, University College Dublin, Belfield, Dublin 4. Tel: 00353 (1) 716 7987, email: seaneen.sloan@ucd.ie

Abstract

Zippy's Friends is a social and emotional learning programme currently delivered in UK primary schools as well as across 30 countries worldwide. It aims to develop coping and social skills among five to seven year old pupils. It typically runs over the course of one school year, and is designed to be delivered by the class teacher on a whole-class basis. This protocol details the methodology underpinning the first large-scale, rigorous evaluation of the programme in a UK context, consisting of a cluster randomised controlled trial and

process evaluation. It will assess the impact of the programme on both socio-emotional and academic outcomes, and is the first evaluation to include a direct assessment of academic attainment.

1. Background

1.1. Significance

An estimated 10% of children and young people aged between five and 16 years in the UK 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 social and emotional skills for effective learning (Collaborative for Academic, Social, and Emotional Learning 2015; DoH, 2012; DfE 2015).

Within this context, schools are seen as playing a crucial part in contributing to the positive mental health of children and young people. Several reviews have concluded that universal interventions promoting social and emotional 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).

1.2 Intervention

Zippy's Friends is a universal, classroom-based social and emotional learning 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 five to seven year olds by the class teacher over the course of an academic year, through 24 weekly sessions which last 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 they become better equipped to cope with difficulties in life, and get on better with others.

Teachers undergo a one-day 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, in which teachers in the same area are 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 & Folkman, 1984; Sandler et al., 2000). In this context, coping skills are defined as the skills or strategies used to manage needs and emotions caused by stressful situations or experiences (Lazarus & 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 very 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), and it a central aim of this study is to test whether Zippy's Friends has an impact on children's academic attainment. The logic model (Figure 1) 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.

Zippy's Friends was first piloted in Denmark and Lithuania using a quasi-experimental design (Mishara & Ystgaard, 2006) in which the control group were matched to the intervention group based on gender and socioeconomic background. This study found improvements in coping ability, cooperation, empathy, assertion and self-control. Also, the problem behaviours of externalizing and hyperactivity decreased. The study, however, had major methodological limitations (i.e. no random allocation and differential measurement attrition).

The programme has since been evaluated using cluster-randomised controlled trials in Norway (Holen et al., 2012; 2013) and Ireland (Clarke et al., 2014). These trials demonstrated a positive impact of the programme on children's coping strategies (ES= −0.38; Holen et al., 2012), self-awareness (ES=0.35; Clarke et al., 2014), self-regulation (ES=0.12; Clarke et al., 2014), social skills (ES=0.12; Clarke et al., 2014), reduced impact

of mental health difficulties on the classroom (ES= −0.15; Holen et al., 2012) and reduced bullying (ref). The study by Holen et al. (2013) also reported a positive impact on classroom climate (ES=0.61) and pupil academic skills (ES=0.42), although this finding was based on teacher report rather than direct measures of attainment, and thus is subject to bias.

The current trial represents an opportunity to add to this emerging evidence base via a fully powered and methodologically rigorous test of *Zippy's Friends* using multiple and reliable data sources of social and emotional functioning. Further, it will be the first large-scale evaluation of the programme in a UK context, and is the first study of its kind to examine the direct impact of the programme on standardised academic test scores.

2. Research Plan

The evaluation will consist of two components: an impact evaluation and an implementation and process evaluation. The aim of the impact evaluation is to assess whether *Zippy's Friends* can improve a number of specific outcomes for Year 2 pupils who are aged six to seven years. The aim of the process evaluation is to better understand the factors associated with implementation fidelity and delivery.

2.1. Research questions

The following research questions will be addressed:

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 specific subgroups, i.e. boys or girls, those eligible for Free School Meals and those with English as an Additional Language?

Implementation and process evaluation:

3. How is the programme perceived by schools (in terms of engagement of pupils, teachers and local coordinators), 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 (Figure 1) explain any link between the programme and academic attainment?

2.2. Design

The evaluation is being conducted in primary schools across England during the 2016/17 school year. The impact evaluation consists of a Cluster Randomised Controlled Trial (CRCT) with two levels (pupils clustered within schools) and involving two arms (an intervention and control group).

2.3. Participants

The target cohort is Year 2 pupils (aged six to seven years), and their teachers, from state-funded primary schools within five Local Authorities (LAs) in England. All state-funded Infant and Primary schools based within the five LAs are eligible to take part in the trial if they:

- Have not delivered *Zippy's Friends* before;

- Are prepared to allow one hour per week in the Year 2 timetable for *Zippy's Friends*;

- Are prepared to release Year 2 teachers for training and support sessions.

Headteachers in schools who express interest in participating in the trial are asked to read and sign a Memorandum of Understanding (MoU) which explains what participation in the trial would involve for the school. Schools returning a signed MoU are then sent the pre-test assessments which must be completed and returned to the evaluation team in order to be eligible to go forward for randomization.

All pupils in Year 2 during the programme delivery year (2016/17) attending schools that are taking part in the trial are eligible to participate. Parental opt-out consent will be sought for

pupil participation in data collection. Information and opt-out consent forms will be sent home via schools between April and June 2016, prior to randomisation. Schools will inform the evaluators of any opt-out parental consent forms received; these pupils will still receive the programme (if the school is allocated to the intervention group), however will not take part in evaluation data collection. During whole class testing, any opted-out pupils will be provided with other activities to complete by their teacher.

2.4. Randomisation

Schools will be randomly allocated to either an intervention arm (which will receive the programme) or a control arm (which will continue with standard practice) in July 2016. A cluster-randomized design is used because the programme is designed to be delivered on a whole-class basis. Allocation is at the level of the school to minimize the risk of contamination between trial arms.

The evaluation team will conduct random allocation in July 2016 and will communicate the outcome of this to schools by the end of the school year. Schools must have returned a signed MoU, gained parental (opt-out) consent, completed and returned the pre-test measures (described in section 2.5) and provided the evaluation team with data on participating pupils (i.e. Unique Pupil Number, first name, last name and date of birth, for linking to the National Pupil Database) to be eligible to be included in the random allocation process.

Schools will be allocated to groups using the program Minim[1] to create groups that are balanced across pre-specified characteristics. This method of allocation is known as minimization and is a widely accepted alternative to simple or stratified randomization (Altman & Bland, 2005). For this trial, minimization will be used to ensure that equal numbers of schools within each of the five local authority area are allocated to each group, to avoid the scenario whereby one area has very few intervention schools and another has many.

---

[1] http://www-users.york.ac.uk/~mb55/guide/minim.htm

School-level attainment (high, medium and low) will also be used to ensure there are a similar proportion of high, medium and low attaining schools in both groups. School-level attainment is based on percentage of pupils obtaining level 4 or above in reading, writing and maths at Key Stage 2, split into three equally sized categories using internally-derived cut-offs. This will involve taking published figures on the percentage of pupils attaining level 4 or above in reading, writing and maths at Key Stage 2 for each school recruited to the trial, and identifying cut off points such that three equal groups are created representing the top, middle and bottom third attaining schools within the sample.

Schools allocated to the intervention arm will begin delivering the programme in September/October 2016, following one day of teacher training. Schools allocated to the control group will not receive the programme during the evaluation period but will proceed with 'business as usual'. In addition, control schools will receive a payment of £1000 at the end of the trial, and can choose to use this to purchase the programme following the trial period.

2.5. Outcome Measures

Outcomes will be measures at pre-test (April-June 2016) and post-test (June 2017). Selection of outcomes was guided by the logic model (Figure 1) which was designed by the evaluation team, the programme delivery team and the Education Endowment Foundation during initial set up meetings.

2.5.1. Primary outcomes

The two primary outcomes are 1) academic attainment (reading) and 2) emotional self-regulation.

1) Attainment in reading will be measured at pre-test using the Hodder Oral Reading Test Form A (Hodder Education) which will administered by Year 1 class teachers (or teaching

assistants) during the pre-test period and prior to randomisation. Attainment in reading will be measured again at post-test (June 2017) using the paper version of the Hodder Group Reading Test (Hodder Education). This will be administered on a whole-class basis under test conditions, independently by fieldworkers (blind to group allocation) who will be trained, overseen and monitored by the evaluation team. Completed test papers will be returned to Hodder for marking, which again will be blind to group allocation.

2) Emotional self-regulation will be measured at pre-test using teacher-report on the 'self-regulation' subscale of the Emotional Literacy Checklist (Faupel, 2003; GL Assessment). Pupil self-report at pre-test will not be collected due to the age pupil age at pre-test (five to six years) and the lack of reliable self-report measures for children of this age. However self-report be collected at post-test (when pupils are aged six to seven years), using the Child Anger Management Scale (CAMS; Zeman et al., 2001). This measure was selected over the Emotional Literacy Checklist as the self-regulation subscale in pupil version demonstrates poor reliability when used as a single subscale (Cronbach's alpha=0.52; Faupel, 2003). The Child Anger Management Scale contains 11 items covering three areas: inhibition of anger expression (four items, e.g. 'I hold my anger in'); coping, or anger control (four items, e.g. 'I try to calmly deal with what is making me feel mad'), and dysregulation of anger expression (three items, e.g. 'I say mean things to others when I'm mad'). It 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 (Zeman et al. 2001; Zeman et al., 2002; Suveg & Zeman, 2004; Sim & Zeman, 2006; Suveg et al., 2009). The scale was recently used by the evaluators in an evaluation of another social and emotional learning programme (Roots of Empathy) in a UK setting, in which it demonstrated good reliability (Cronbach's alpha=0.69-0.77) with pupils aged eight years and over. This measure will be included within a questionnaire that will be administered on a whole-class basis by the fieldworkers.

2.5.2. Secondary outcomes

The two secondary outcomes are 1) self-regulated learning and 2) social skills.

1) Self-regulated learning will be measured using teacher report at both pre-and post-test on the 'attention/persistence' and the 'emotional control' subscales (12 items in total) of the Learning Behaviour Scale (McDermott et al., 2001). Pupil report on this outcome will not be measured at either time point due to the lack of existing self-report measures available for use with pupils aged five to seven years.

2) Social skills will be measured at post-test using the social skills subscale of the Social Skills Improvement System rating scale (SSIS; Gresham & Elliott, 2008). This measure was chosen as it is well established in the academic literature, is standardised for use with children and young people aged three to 18 years, includes both teacher- and pupil-report options, and can be completed by pupils aged eight years and over. At pre-test, due to pupil age, social skills will be measured using teacher-report only. At post-test, social skills will be measured using pupil self-report only.

2.6. Covariates

Schools will be asked to provide the following pupil-level data (for those with parental consent): Unique Pupil Number (UPN), first name, surname and date of birth (in order to match to data held in the National Pupil Database) as well as English as an Additional Language (EAL) status. In September 2016, the following pupil-level data will be requested from the NPD: gender and free school meal eligibility (FSMever) for subgroup analysis.

2.7. Sample size calculations

No study has assessed the impact of *Zippy's Friends* on a direct measure of academic attainment. A meta-analysis of the impact of social and emotional learning programmes on academic performance found an effect size of g=+0.27 (Durlak et al., 2011) and effect sizes estimating the impact of *Zippy's Friends* on social and emotional outcomes have ranged between 0.12 and 0.38 (Clarke et al., 2014; Holen et al., 2012; 2013).

At the evaluation design stage, it was estimated that 70 schools with an average of 37 pupils per school (approx. 2590 pupils in total) would be sufficient to detect an effect size of 0.20 in an intention-to-treat analysis. This was calculated using Optimal Design (version 3.01) and is based on a two level cluster design with an estimated pre-post correlation (R^2) of 0.50, an intracluster correlation coefficient (ρ) of 0.16 for the primary outcome, and standard power and alpha thresholds of 0.80 and 0.05 respectively. However, during the course of the school recruitment stage of the project, a higher number of schools than anticipated expressed interest in taking part in the trial. An additional 15-20 schools would power the trial to detect an effect size of 0.18 (using the same assumptions as above), thus it was agreed with the funder that up to 90 schools could be recruited.

2.8. Analysis plan

Data will be input into SPSS (version 22), blind to allocation. Data management, checking and analysis will be conducted in Stata version 14 (Stata Corporation, College Station, Tx, US). The main stages involved in the analysis are described below, and include data cleaning procedures, methods for handling missing data and establishing baseline equivalence.

2.8.1. Quality checking, data cleaning and deriving variables

- Data quality checking and cleaning, including inspection of descriptive statistics, normal plots, histograms and scatterplots to identify potentially erroneous or implausible data;
- Derivation of scale variables following inspection of the psychometric properties of each scale through factor analysis and reliability analysis (Cronbach's alpha).

Identical quality checking, cleaning and variable derivation procedures will be used for pre-test and post-test data.

2.8.2. Methods for handing missing data

Missing post-test data for reading attainment (primary outcome) may occur if pupils are absent from school on the day of test administration. This will be minimised by a follow-up visit to any school with pupil absences. Missing data may also occur if a pupil leaves the school completely before the post-tests are administered, if the child does not assent to participate on the day of testing, or if parents withdraw consent.

For the three other outcomes measured using multi-item scales (i.e. emotional regulation, self-regulated learning and social skills), there will be two types of missing data: (1) complete missing data (i.e. due to pupil absence from school during data collection) and (2) partial missing data where the pupil has completed some but not all items of the measure. Again, complete missing data will be minimised by a second visit to the school to obtain data for any pupils who were absent on the main day of testing. Partial missing data will be minimised during administration, as the fieldworker will ask pupils to check over their questionnaires for any items that may be have accidently missed out. The extent of missing data within each scale will be checked and cross-referenced with paper questionnaires to check for data entry errors. For each scale, partial missing data will be dealt with by taking the following steps:

- Participants missing 75% or more items will be identified with a derived variable indicating 'high level of missing data for X' where X is the specific scale affected;

- For all participants, a total scale that takes account of missing data will be generated by calculating total sum of all observed scores plus the sum of missing scores with missing scores replaced with the mean of observed scales, i.e.

$$\text{Scale score} = \Sigma Io + \left(Nm \times (\Sigma Io \div No)\right)$$

Where:

      Io = all observed items

      No = number of observed items

      Nm = number of missing items

Hence it will be possible to derive a score for all pupils who have completed some of the questionnaire. Analyses will be completed with all pupils, and repeated excluding those with a high level of missing data. If removal of those with high levels of missing data changes the results in a way that would affect the interpretation, then both sets of results will be reported; otherwise only the results with all included pupils will be reported.

For all other variables, the proportion of and reason for missing data will be assessed and reported. The proportion of each outcome lost to follow-up in the control and intervention group will be examined. If missing data are less than five percent then a complete case analysis will be undertaken.

Multiple imputation will be conducted as a sensitivity analysis if there is a high level of missing data. The pattern of missing data will be explored by comparing the proportion of missing data in each of the control and intervention groups in addition to exploring how missingness is related to the outcomes in question. If the data can be assumed to be MCAR then imputation is not required. If the data are assumed to be MAR, this assumption renders the missing mechanism ignorable, simplifying the imputation step whilst ensuring correct inference. The imputation model will impute data separately for the control and intervention groups and will include all relevant variables and auxiliary variables involved in the analysis and sampling design. The imputation will be performed using chained equations which 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 will be conducted in order to lessen the simulation (Monte Carlo) error. The analysis using the imputed datasets will then be compared to the complete case analysis.

2.8.3. Establishing baseline equivalence

The extent to which the control and intervention groups are balanced will be assessed by comparing:

- school characteristics: school type, locality (urban/rural), Ofsted rating, school size (based on total number of pupils on the register),
- school-level pupil characteristics: proportion attaining level 4 or above at Key Stage 2 in reading, writing and maths (in 2015), proportion with English as an additional language, proportion FSM eligible and proportion with statements of special educational needs or education and health care plans;
- sample pupil characteristics: gender, FSM eligibility and EAL;
- the primary and secondary outcomes measured at pre-test.

Linear and logistic multilevel regression models will be used to account for the clustered nature of the data.

2.8.4. Main analysis

The main effects of the intervention on each of the outcomes will be analysed on an intention-to-treat basis (meaning that participants will be analysed in the group to which they were randomised, irrespective of whether or not they received (or completed) the programme). The clustered nature of the data will be accounted for using multi-level modelling (with pupils (level one) clustered within schools (level two)).

The primary analyses will estimate the impact of the programme on the two primary outcomes (reading attainment and emotional self-regulation) and the two secondary outcomes (self-regulated learning and social skills).

Each model will include the relevant outcome measure as dependent variable, and the following independent variables:

- a binary dummy variable representing school allocation to the intervention (coded '1') or control group (coded '0');

- the relevant associated pre-test measure of reading ability or emotional self-regulation;

- the variables used in randomisation to create balanced groups (school-level prior attainment and LA);

The focus for the analysis will be the estimated coefficient representing the difference in mean outcome scores between the intervention and control groups.

Effect size (Hedges' g) will be calculated as the standardised mean difference between the control and intervention groups, using the pooled standard deviation. The pooled standard deviation will be calculated using the formula:

$$s = \frac{\sqrt{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}}{n_1 + n_2 - 2}$$

Hedges' g will then be calculated as:

$$g = \frac{\text{coefficient}}{\text{pooled standard deviation}}$$

2.8.5. Subgroup analysis

Subgroup analysis will involve extending the main effects models for each outcome to include an interaction term based on pupil characteristic of interest. The following groups of pupils will be examined: those eligible for FSM, those with EAL and boys compared to girls.

2.8.6. Exploratory analysis

As a test of the logic model (Figure 1), exploratory analyses will examine the extent to which any effect of the programme on academic attainment is mediated through socio-emotional attributes. Mediation analysis (Baron & Kenny, 1986) assumes that the intervention influences the mediating variables (in this case, self-regulation) and through this influence on the mediator produces the effect on the outcome (in this case, attainment in reading).

Therefore mediation analyses will be performed if (through the main analysis), the programme is found to have an effect on the proposed mediators (self-regulation in the domains of emotions and learning behaviour) as well as reading attainment. If this is the case, multilevel linear regression (as per the main analysis, see section 2.8.4) will be repeated with and without the presence of mediator variables, and the effect of the intervention on the academic attainment before and after adjustment for mediating variables will be examined.

Exploratory analysis will also examine any differential impact of the programme on the primary and secondary outcomes depending on variations in implementation fidelity (described in section 2.9).

2.9. Implementation and process evaluation methods

The implementation and process evaluation will be undertaken alongside the CRCT, with the purpose of exploring four main programme implementation and fidelity issues:

- adherence to programme manual;

- levels of exposure to the programme (dosage);

- the quality of the programme delivery; and

- teacher and pupil engagement with the programme.

At pre-test (and prior to randomisation) all schools will complete a short (one page) pro forma detailing PSHE delivery and any other social and emotional learning programmes used in Key Stage 1, in order to establish baseline equivalence. This pro forma will be completed by the headteacher or a member of staff with oversight responsibility for Key Stage 1 and will be returned along with MoU.

The evaluator will attend a teacher training session (scheduled for September 2016) and follow-up session (which are due to take place in January, March and June 2017), in order to understand best practice in delivery of *Zippy's Friends*. During the course of the delivery

year, data will be collected using a mixed methods approach, including ongoing implementation logs and an implementation survey, both completed by teachers delivering the programme, fidelity ratings completed by the local coordinator, interviews and observations in a sample of schools in the intervention group.

Implementation logs: The evaluation team in collaboration with the programme delivery team will agree definitions of low, medium and high fidelity. These will be informed by:

- Programme delivery logs completed by teachers throughout the course of the school year. Teachers will keep a record each session in terms of duration and whether any aspects of the lesson plans were not covered;

- Teacher attendance at both initial training and follow-up sessions;

- Contact visits and support given by local coordinator.

Online survey: will be administered with Year 2 teachers at the beginning of the summer term of the programme delivery year. The survey will collect data on perceived ease of implementation and fidelity to the programme schedule.

Interviews and observations: A number of schools will be selected and recruited for process evaluation visits by the evaluator, to include a mix of urban/rural location, size and level of deprivation (school FSM%). Within these schools, interviews (with teachers and the school senior management team) will focus on perceptions of the programmes impact; what has worked well or less well in terms of delivery, the perceived 'costs' (time, resource, adverse consequences) or the value added to the school (in terms of wider benefits, e.g. to school ethos, teacher-pupil relations), and links between the programme content and the general curriculum. Focus groups will be used with pupils to explore the perceived impact of the programme based on the mechanisms proposed in the logic model (Figure 1), i.e. social and coping skills, classroom environment, focus on learning and communication skills.

Parents will be interviewed either via telephone or face-to-face on school premises, depending on individual participant preference.

With consent, interviews and focus groups will be tape-recorded and transcribed verbatim and field notes written up from the lesson observations.

Control group activity: Further, as the National Curriculum framework states that all schools should make provision for Personal, Social, Health and Economic (PSHE) education, it will be important to gather information on 'business as usual' in relation to pupil socio-emotional wellbeing in the control group. Thus, control group activity will be captured through the baseline (pre-randomisation) pro forma on standard practice in relation to PSHE and social and emotional learning. At the end of the school year, Year 2 teachers in the control group will also be asked to complete a questionnaire on actual practice in terms of PSHE and SEL during the school year.

3. Personnel

3.1. Programme delivery team

The programme delivery team consists of Caroline Egar (Programme Director) and Lizzie Poulton (Programme Manager), based at Partnership for Children (PfC), a UK charity which support the delivery of Zippy's Friends nationally and internationally. They are responsible for recruitment of schools through local authority (LA) areas and are supported by a local coordinator within each area. The local coordinator is generally a professional working in a school or with a remit for working in schools (e.g. educational psychologist). For the trial, local coordinators 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.

3.2. Evaluation team

The evaluation team was appointed by the Education Endowment Foundation as independent evaluators (in that they have no role in the development or delivery of the programme being evaluated) and have responsibility for the design, management, analysis and reporting of the evaluation. The evaluation is based at and managed from the Social Sciences, Education and Social Work at Queen's University Belfast (QUB), and involves Aideen Gildea, Dr Sarah Miller and Professor Allen Thurston. The lead evaluator, Dr Seaneen Sloan, is based at the School of Education, University College Dublin and is a visiting research fellow at Queen's University Belfast.

4. Timeline

Key evaluation activities, dates and responsibilities are outlined in Table 1.

Table 1: Evaluation timeline

Activity	Dates	Responsibility
Evaluation set up (ethical approval; design of memorandum of understanding and consent forms; selection and piloting of measures)	January – March 2016	Evaluation team
Recruitment of local authorities, local coordinators and schools	January – June 2016	Programme delivery team
Administration of pre-test	April – June 2016	Evaluation team
Randomisation and notification of outcome	July 2016	Evaluation team
Application to link to National Pupil Database	September 2016	Evaluation team
Teacher training	September 2016	Programme delivery team
Programme delivery	October 2016 – June 2017	Intervention schools (supported by programme delivery team and local coordinators)
Follow up training sessions (x3)	January 2017 March 2017 June 2017	Programme delivery team and local coordinators
Process evaluation – interviews and observations	January – May 2017	Evaluation team
Process evaluation – implementation survey	April – May 2017	Evaluation team
Post-test data collection	June 2017	Evaluation team
Data input	July – August 2017	Evaluation team

Data analysis and write up	August – November 2017	Evaluation team
First draft of final report	November 2017	Evaluation team

5. Ethics

The trial was approved by the Research Ethics Committee School of Education, QUB in January 2016. For the trial, school-level consent to participate will be sought from school headteachers who will all be asked to sign the MoU. Informed consent for pupil participation in trial data collection will be sought from parents or guardians on an opt-out basis. For the process evaluation, informed opt-in consent will be obtained for interviews with school personnel and parents. Parental opt-in consent and pupil verbal assent will be sought for any pupils taking part in focus group discussions.

6. Publication plan

The main report will be authored by the evaluation team and will report findings from the main analysis and the FSM subgroup analysis. This will be submitted to the Education Endowment Foundation in November 2017 and will be openly available on the funder's website in early 2018 following peer review. In addition, it is anticipated that four papers based on data from the trial will be published in peer-reviewed journals.

REFERENCES

Adi, Y., Killoran, A., Janmohamed, K. & 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. & Bland, J.M. (2005) Treatment allocation by minimisation, *British Medical Journal,* 330, 843.

Baron, R.M. & Kenny, D.A. (1986) The moderator mediator variable distinction in social psychological research. Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.

Browne, G., Gafni, A., Roberts, J., Byrne, C. & Majumdar, B. (2004) Effective/efficient mental health programmes for school-age children: a synthesis of reviews. *Social Science & Medicine*, 58(7), 1367-1384.

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 on 10th February 2016.

Chien , N.V., Harbin, V., Goldhagen, S., Lippman, L. & 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. & 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), *786-798*

Clarke, A.M. Morreale, S., Field, C.-A., Hussein, Y. & 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.

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

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

Durlak, J., Weissberg, R.P., Dymnicki, A.B., Taylor, R.D. & 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), 405-432.

Faupel, A. (2003). *Emotional Literacy: Assessment and Intervention*. UK: GL Assessment.

Green, H, McGinnity, A., Meltzer, H., Ford, T. & 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, 27-52.

Gresham, F.M. & Elliott, S.N. (2008) *Social Skills Improvement System Rating Scales Manual.* Minneapolis, MN: NCS Pearson.

Gutman, L.M. & Schoon, I. (2014) Preventative interventions for children and adolescents: a review of meta-analytic evidence. *European Psychologist*, 20, 231-241.

Holen, S., Waaktaar, T., Lervåg, A. & 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), 657-677.

Holen, S., Waaktaar, T., Lervåg, A. & 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), 420-444.

Knapp, M., McDaid, D. & Parsonage, M. (2011) Mental health promotion and mental illness prevention: the economic case. London: PSSRU.

Lazarus, R.S., & Folkman, S. (1984). *Stress, appraisal, and coping.* New York: Springer.

McDermott, P. A., Green, L. F., Francis, J. M., & Stott, D. H. (2001). *Learning Behaviors Scale.* Philadelphia: Edumetric and Clinical Science.

Mishara, B.L. & Ystgaard, M. (2006) Effectiveness of a mental health promotion program to improve coping skills in young children: Zippy's Friends. *Early Childhood Research Quarterly,* 21,110-123.

Polanczk, G.V., Salum, G.A., Sugaya, L.S., Caye, A. & 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, 345-365.

Qualter, P., Gardner, K.J., Pope, D.J., Hutchinson, J.M. & 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), 83-91.

Sandler, I.N., Tein, J.-Y., Mehta, P., Wolchik, S. & Ayers, T. (2000) Coping efficacy and psychological problems of children of divorce. *Journal of Clinical Psychology*, 29, 336-347.

Segal, J. (1983) Translating stress and coping research into public information and education. In M. Kessler & S. E. Goldston (Eds.), *A decade of progress in primary prevention*. Lebanon, New Hampshire: University Press of New England.

Tennant, R., Goens, C., Barlow, J., Day, C. & 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), 25-32.

Figure 1: Logic model outlining how Zippy's Friends is proposed to impact on socio-emotional and academic outcomes for participating pupils