FRIENDS for Life Study - Statistical Analysis Plan

Institute of Education, University of Manchester

TRIAL FULL TITLE	FRIENDS for Life Study
IMPLEMENTOR	Salus
EVALUATOR	University of Manchester
TRIAL REGISTRATION	ISRCTN13721202
PRINCIPAL INVESTIGATOR	Dr. Michael Wigelsworth
SAP VERSION	1
SAP VERSION DATE	18/05/2017
DATE OF EEF APPROVAL	20/07/2017
DATE OF DEVELOPER APPROVAL	23/07/2017

Brief background

This one-year project will evaluate the effectiveness of a primary-school based intervention called FRIENDS. FRIENDS is designed to promote emotional resilience in order to prevent (or stabilise) the development of negative feelings of anxiety and depression, which are recognised as amongst the most common childhood disorders. FRIENDS is delivered both as a whole class programme, and also as a 'selective' or 'small group' intervention for particular pupils requiring additional support. It is intended to address all levels of prevention, early intervention, and treatment within a school setting. FRIENDS has developed an impressive international evidence base since its official launch in 1991, and is in use in at least 20 countries (including Ireland, where it forms part of the curriculum). It is the only childhood anxiety prevention program acknowledged by the World Health Organization. The project is important because research tells us that anxiety and depression are very common in children. It is suggested that by the age of 18, 1 in 10 children will have suffered from an anxiety disorder, with many more children experiencing serious symptoms that fall below clinical criteria for diagnosis. The aim of the FRIENDS curriculum is to prevent such problems before they occur, so it is important to see if, how, and why, and for who, the FRIENDS curriculum works for in English primary schools.

Study aims

The primary aims of the proposed research are to examine the impact of the FRIENDS for Life programme on primary school children's: 1) academic attainment and 2) health related outcomes through a cluster randomised control trial. Specifically:

Universal effects

Hypothesis 1: Children in classes implementing FRIENDS will demonstrate measureable improvements in **Key Stage 2 Maths and Reading combined scores (equal weighting)** when compared to those children attending comparison classes

Hypothesis 2: Children in classes implementing FRIENDS for the duration of the programme will demonstrate measureable decreases in their **self-rated worry** when compared to those children attending comparison classes

Hypothesis 3: Effects outlined in H1 & H2, will also be present for:

H3a. **Maths scores only**
H3b. **Reading scores only**

H3c. Self-rated **Total anxiety and depression score**
H3d. Teacher rated **emotional symptoms**
H3e.Teacher rated difficulties **conduct problems**

Subgroup effects

Hypothesis 4: For the outcome variables listed in H1-3 (a-e); There will be a significant interaction between allocation to condition and:

H4a. Free school meal status (**EverFSM**)
H4b. Pupil scoring in the top 20% of self-rated worry at baseline (**'at risk'**)

Implementation

Hypothesis 5: Variation in implementation fidelity (H5a), adaptations (H5b), quality (H5c), dosage (H5d), programme differentiation (H5e), reach (H5f) and participant responsiveness (H5g) will moderate education-related outcomes (H1) in schools implementing FRIENDS

Hypothesis 6: Variation in implementation fidelity (H5a-f) will moderate health related outcomes (H2, H3) in schools implementing FRIENDS

Hypothesis 7: Variation in implementer characteristics, specifically emotional self-efficacy (H7a), teaching efficacy (H7b) and views of social and emotional learning (H7c) will be related to implementation variation (H5 a-f).

Power and sample size (PASS)

The initial aim of the study was to recruit a total of 110 classes from 77 schools (an estimate based on the number of single entry schools in the implementation area – see appendix 1). All calculations assumed*: N= 28 per cluster (Department for Education, 2015a); ICC (class level) = 0.17*[1]*, Power=0.8, Alpha=.05,*[2] *proportion of single form entry = 51%*[3]

Recruitment exceeded targets, resulting in a total of 128 classes (81 schools), resulting in a minimal detectable effect of **0.1.**

Randomisation

A cluster-randomised design was utilised, using matched pair (by type of entry form – single; double; triple; quad) randomisation at the class level. KS1 and baseline 'total anxiety and depression scores' (see Outcome Measures) were used for minimisation.

For each participating school:

- Single form entry: Randomly allocated to receive either FRIENDS, or £1000 (delivered at the end of the academic year)

[1] Estimate drawn from the largest academic ICC (KS2 results in writing) from the PATHS trial
[2] Although we would expect attrition to be low given the study design, we have included a conservative estimate drawn from previous trials (e.g. PATHS) with a loss of 9% in a C-RCT design.
[3] Based on approximate data of Kent school sizes, provided by Salus

- Double form entry: Random allocation of one class to receive FRIENDS, one class to serve as usual practice
- Triple form entry: Random allocation of either one or two classes to receive FRIENDS, with remaining classes to serve as usual practice.
- one class to receive FRIENDS, one class to
- Four form entry: Random allocation of 2 classes to receive FRIENDS, two classes to serve as usual practice.

Regarding split classes (e.g. a merged year 4/5 class), this was treated as a class of 15 year 5 pupils (assuming a class size of 30).

Randomisation procedures were conducted by the Manchester Academic Health Science Centre Clinical Trials Unit (MAHSC-CTU) to ensure an independent process, free from bias. The final allocation was **65 classes allocated to FRIENDS** & **63 classes allocated to comparison**.

Outcome measures

Primary outcome measures

The primary academic outcome measure (H1) is **KS2 Maths and reading combined (equal weighting)**, as measured by the standardised curriculum tests conducted in July 2017. Key Stage 1 scores will be used as the baseline co-variate. Key Stage 1 scores are provided by Kent County Council (KCC) (due to time constraints), and Key Stage 2 scores will be obtained from the National Pupil Database (as some minor additional cleaning of the data makes this a more preferable source in comparison to KKC).

The primary heath outcome measure (H2) is **self-rated worry** as measured by the *Penn State Worry Questionnaire for Children (PSWQ-C) (*Meyer, Miller, Metzger, & Borkovec, 1990). This paper-based measure is deployed at baseline (March 2016) and at non-academic post-test (December 2016).

Secondary outcome measures

Self-rated **Total anxiety and depression score** (H3a) is assessed by taking the total score from the *Revised Child Anxiety and Depression Scale (RCADS 30) (*Ebesutani et al., 2012). This paper-based measure is deployed at baseline (March 2016) and at non-academic post-test (December 2016).

Teacher rated **emotional symptoms** (H3b) are measured using the 5 point subscale from the teacher version of the Strengths and Difficulties Questionnaire (Goodman, 2001). This measure is deployed at online at baseline (March 2016) and at non-academic post-test (December 2016).

Teacher rated difficulties conduct problems (H3c) are measured using the 5 point subscale from the teacher version of the Strengths and Difficulties Questionnaire (Goodman, 2001). This measure is deployed at online at baseline (March 2016) and at non-academic post-test (December 2016).

Analysis protocol

Initial treatment

For any of the above outcome measures, the standard procedure to be applied is as follows:

- Data cleaning and screening ahead in preparation for analysis
    - Merging of online survey export files and hand input files.
- Basic descriptive analysis
    - Production of descriptive statistics by intervention group (e.g. means, standard deviations) and visual data displays (e.g. error bar charts) to identify key trends vis-à-vis trial Hypotheses.

Demonstration of equivalence at baseline (on the basis of primary and secondary outcomes).

Missing data

First, the extent of missing data will be established. As Educational data (H1, H3a, H3b) is drawn from the NPD, there will be minimal missing data. However, for the primary outcome data (H2, H3c-e) this is also likely to be minimal given excellent responses at baseline. Once data is available, differences between complete and missing cases will be examined to establish any pattern to the missingness. Logistic regression will be used to predict missingness, whereby each child will be coded as providing complete (0) or incomplete (1) outcome data, with other study data as explanatory variables (Pampaka, Hutcheson, & Williams, 2017).

We will also perform an analysis using complete cases and a sensitivity analysis using multiple imputation (via the REALCOM-Impute extension to MLWin) *unless* missing data is less than 5%. Accordingly, multiple imputation procedures will be carried out in REALCOM-Impute, using the missing at random assumption (Carpenter, Goldstein, & Kenward, 2011). This will enable us to include both partially and completely observed cases of all schools and pupils in the analysis, thereby reducing the bias associated with attrition. Demographic variables (e.g. gender, FSM eligibility), explanatory outcome variables (e.g. KS1 scores), and the constant will be entered as auxiliary variables and used to impute missing values. REALCOM-Impute default settings of 1000 iterations and a burn-in of 100, refresh of 10, will be used, following guidance for multi-level imputation with mixed response types (Carpenter et al., 2011).

Primary intention-to-treat (ITT) analysis (H1, H2, H3 (H3a-e))

An ITT analysis (according to intention-to-treat principles, e.g. ignoring noncompliance, protocol deviations and other events that take place after randomisation (Gupta, 2011)) will be conducted for H1, H2, H3(a-e). This will conducted through the construction of 2-level (class, pupil) hierarchical models (fixed effects at the class level, utilising robust standard errors) to account for nested nature of dataset using MLWin Version 2.36.

We will employ an Initial unconditional model to ascertain variance attributable to class and pupil followed by conditional model with treatment allocation (e.g. FRIENDS vs. comparison) included at

class level. This will be followed by a 'conditional' model to include trial allocation (FRIENDS vs. Comparison).

Subgroup (ITT) analysis (H4 (H4a-b))

The protocol for the ITT analysis will be followed. In addition, a further model for each hypothesis H4a-b) will be constructed that will include risk status as a cross-level interaction term (e.g. FSM*Allocation to condition). An intervention effect at the subgroup level will be noted if the co-efficients associated with the interaction terms noted above are statistically significant. These will subsequently be converted to Hedge's *g* as per EEF reporting standards.

Implementation analysis (H5 (H5a-g), H6, H7 (H7a-c))

FRIENDS trial group only: Scores for dosage, fidelity/adherence, quality, reach, and participant responsiveness from observations, each score for each quality indicator transformed to a z score

This will conducted through the construction of 2-level (class, pupil) hierarchical models (fixed effects at the class level, utilising robust standard errors) to account for nested nature of dataset using MLWin Version 2.36. Following from the unconditional models created in the ITT analysis, additional modes will be constructed utilising implementation variables for each hypotheses (H5a-g), (H6), (H7a-c), and each outcome (H 1, H2, H3a-e) included at class level.

Effect size calculation

In all cases, effect sizes will be reported using Hedge's g (Cohen's d bias corrected) and accompanied by 95% confidence intervals as per EEF specifications.

Report tables

All the tables will be structured according to the EEF trial report template[4].

[4] https://educationendowmentfoundation.org.uk/evaluation/resources-centre/writing-a-research-report/

References

Carpenter, J., Goldstein, H., & Kenward, M. (2011). REALCOM-IMPUTE Software for Multilevel Multiple Imputation with Mixed Response Types. *Journal of Statistical Software*, *5*, 1-14.

Ebesutani C., Reise S. P., Chorpita B. F., Ale C., Regan J., Young J., . . . Weisz J. R. (2012). The Revised Child Anxiety and Depression Scale-short version: Scale reduction via exploratory bifactor modeling of the broad anxiety factor. *Psychological Assessment*, *24*, 833-845.

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire (SDQ). *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 1337-1345.

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy, 28*, 487-495.

Pampaka, M., Hutcheson, G., & Williams, J. (2017). Handling missing data: Analysis of a challenging data set using multiple imputation. *International Journal of Research & Method in Education, 39*, 19-37.

Appendix 1

Average distribution of class entry across Kent

1 class entry	**51%**
2 class entry	**27%**
3 class entry	**17%**
4 class entry	**4%**

Timetable

	Trial Management	**Project Delivery**	**Process Evaluation**
Aug	Ethical Clearance Finalise trial protocol		
Sep			
Oct		Contact with schools	
Nov			
Dec			
2016			
Jan	Service agreements & Baseline collection		
Feb			
Mar	Follow up window (ensure against attrition) Scoring of baseline data Randomisation		
April		Intervention begins	Observations Interviews Case study data
May			
Jun			
Jul			
Aug			
Sep		Booster 1	
Oct			
Nov		Booster 2	
Dec	Post-test (non-academic)		
2017			
Jan			
Feb			
Mar			
April			
May			
Jun			
Jul		KS2 Testing	
Aug			
Sep			
Oct			
Nov			
Dec			