Education Endowment Foundation

FRIENDS for life

Evaluation report and executive summary

December 2018

Independent evaluators:

Michael Wigelsworth, Garry Squires, Liz Birchinall, Afroditi Kalambouka, Ann Lendrum, Louise Black, Patricio Troncoso, Joao Santos, Emma Ashworth & Philip Britteon

The University of Manchester

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

Manchester Institute of Education, University of Manchester

A team from the Manchester institute of Education, University of Manchester were responsible for this independent evaluation of 'FRIENDS for life' in English primary schools. The team comprised of: Michael Wigelsworth, Garry Squires, Liz Birchinall, Afroditi Kalambouka, Ann Lendrum, Louise Black, Patricio Troncoso, Joao Santos, Emma Ashworth & Philip Britteon.

The lead evaluator was Dr. Michael Wigelsworth

Contact details:

Dr. Michael Wigelsworth
Manchester Institute of Education
The University of Manchester
Oxford Road
Manchester
M13 9PL

P: 0161 306 1763
E: Michael.wigelsworth@manchester.ac.uk

Contents

Executive summary

The project

FRIENDS for life is a school-based, cognitive-behavioural therapy programme, designed to promote emotional resilience in order to prevent (or stabilise) the development of negative feelings of anxiety and depression. The programme consists of a series of 10 weekly sessions covering topics such as recognising symptoms of anxiety, emotional self-management and supporting peer relationships. Each session is enhanced by a series of additional activities that can be used to practice the newly learned skills as well as homework activities. The Salus Group (Salus), a training provider offering a range of programmes in schools, within the community and at home, delivered the intervention. Salus delivers FRIENDS in the UK and is not affiliated with the original FRIENDS developers.

The project was a cluster-randomised control trial. 122 Year 5 classes in 79 participating primary schools were randomised to either receive FRIENDS or to act as a business as usual control group. The evaluation tested the impact of FRIENDS on the attainment of pupils, measured by combined Key Stage 2 English and Maths scores, and also on health measures of self-rated worry, anxiety and depression, and teacher-rated difficulties. The process evaluation involved interviews with various stakeholders, class observations and focus groups with pupils. The trial took place in schools between March 2016 and July 2017. The programme was co-funded by the Department for Education as part of an EEF funding round on Character Education.

Key conclusions	
1.	The project found no evidence that FRIENDS had a positive impact on children's academic attainment, overall. This result has a high security rating.
2.	Among pupils eligible for Free School Meals, those in the FRIENDS classes made 1 additional months' progress on a combined maths and reading measure compared to children in other classes. These results may have lower security than the overall findings because of the smaller number of pupils.
3.	The project found no evidence that FRIENDS has a positive impact on children's health outcomes, overall. This result has a very high security rating. Pupils eligible for Free School Meals experienced a small increase in self-rated anxiety and depression, although these results may have lower security than the overall findings because of the smaller number of pupils.
4.	Overall, time was found to be the biggest pressure in ensuring consistency and quality of delivery. Schools often struggled to fit the FRIENDS sessions, which varied in length, within the school timetable. This lead to a variability in the amount of time dedicated to FRIENDS.

EEF Security rating

The findings for the KS2 Maths and Reading Combined score outcome have a high security rating, and the findings for the self-rated worry outcome have a very high security rating. This trial was classified as an efficacy trial by the Education Endowment Foundation. It tested whether the intervention worked in a number of schools, and was a well-designed, 2-armed randomised control trial.

The outcome for self-rated worry was well-powered. Although the outcome for KS2 was comparatively not as well powered (and therefore lost a padlock), this study is still (to the evaluators knowledge) the largest trial of FRIENDS ever conducted. Relatively few pupils (8%) who started the trial were not included in the final analysis. The pupils in FRIENDS classes were similar to those in the comparison classes in terms of prior attainment.

Additional findings

This evaluation found no evidence that FRIENDS has a positive impact on children's health based outcomes overall, and in fact showed a small increase in self-rated anxiety and depression for Free School Meal pupils. There was no evidence of a negative effect on those pupils who were identified 'at risk' of emotional symptoms at the beginning of the trial.

Staff and pupils had very positive feedback about the programme. Around half of the case-study schools reported reductions in aggressive or negative behaviour, and increased peer support. Schools also noted that students were tackling difficult subjects (such as learning about Anne Frank) more seriously, and using a more adaptive approach to learning.

The programme was delivered by external Project Officers and overall, this approach was effective. Different Project Officers had different approaches to delivery, but there was no evidence that this affected children's' outcomes.

High quality delivery of the programme was a predictor in reducing self-rated worry. Quality of delivery was measured across a number of areas, including preparedness of the Project Officer, their clarity in expressing concepts and activities, and how they related to the class as a whole. The trial also found that greater engagement from the Project Officers is related to higher levels of self-rated worry.

Cost

For an average[1] single form entry school, the cost of delivering FRIENDS is £1,728 a year, or £64 per pupil per year (£43 per pupil per year for double-form entry). This figure includes (i) delivery of the full 10 sessions of the FRIENDS curriculum by a trained and licenced external implementer, (ii) a group leaders' manual for delivering the session, and (iii) an accompanying activity workbook for each child.

Beyond an initial orientation meeting between the implementers and school staff, the external delivery model means there is no need to train school staff or maintain resources.

Table 1: Summary of impact on primary academic outcome (Key Stage 2 maths and reading combined score) and primary health outcome (self-rated worry)

Group	Effect size (95% confidence interval)	Estimated months' progress	EEF security rating	No. of pupils	P value	EEF cost rating
KS2 combined maths and reading	0.01 (-0.065, 0.078)	0	🔒🔒🔒🔒	1476	0.908	**£**£££
KS2 combined (FSM)	0.05 (-0.058; 0.159)	1	n/a	371	0.587	**£**£££
Self-rated worry	0.05 (-0.019; 0.124)	n/a	🔒🔒🔒🔒🔒	1476	0.269	**£**£££
Self-rated worry (FSM)	0.15 (0.039; 0.256)	n/a	n/a	371	0.109	**£**£££

** For the primary health outcome, a negative value would indicate a reduction in self-rated worry.*

[1] Assuming a class size of 27 (DfE, 2017d)

Introduction

Intervention

FRIENDS for life is a school-based, cognitive-behavioural preventative programme, designed to promote emotional resilience in order to prevent (or stabilise) the development of negative feelings of anxiety and depression. Anxiety and depression are recognised as amongst the most common childhood mental health disorders. It is suggested that by the age of 18, 1 in 10 children will have suffered from an anxiety disorder (Costello et al, 2003), with many more children experiencing serious symptoms that fall below clinical criteria for diagnosis. As FRIENDS contains universal, selective and indicated modes of delivery, it is intended to address all levels of prevention, early intervention, and treatment within a school setting, across a wide range of ages (see below).

FRIENDS has developed an impressive international evidence base since its official launch in 1998, with positive results for anxiety and/or depression reduction found in Australia (e.g. Lowry-Webster, Barrett, & Lock, 2003), Germany (Essau, Conradt, Sasagawa & Ollendick, 2012), Mexico (Gallegos, Linan-Thompson, Stark & Ruvalcaba, 2013), and recently, England (Stallard et al., 2013). FRIENDS is in use in at least 20 countries (including Ireland, where it forms part of the curriculum) and is currently endorsed by agencies such as the Substance Abuse and Mental Health Services Administration (who have included it on their national registry of evidence-based programmes and practices). It is the only childhood anxiety prevention program acknowledged by the World Health Organization (WHO, 2004). However, although the weight of evidence (see, for example Briesch, Hafermoser Sanetti & Briesch, 2010) implies FRIENDS can *potentially* be effective, more recent reviews have cast doubt on the consistency by which positive results may be achieved, and for whom, noting both implementation and evaluation limitations in research (Maggin & Johnson, 2014).

In order to provide a comprehensive and transparent description of the FRIENDS programme, we utilise an adapted version of the Template for Intervention Description and Replication (TIDieR) (Hoffmann et al., 2014), as per recommended reporting guidance (Humphrey et al.,2016):

1. Name of intervention

'FRIENDS for Life' (hereafter referred to as 'FRIENDS').

'FRIENDS for Life' is one of three versions of the FRIENDS curriculum (in contrast to FUN FRIENDS; ages 4-7, and My FRIENDS youth; ages 12-16), and is designed specifically for ages 8-11 years.

FRIENDS is an acronym, broadly reflecting the programme content (**F**eelings; Remember to relax; **I**nner helpful thoughts; **E**xplore solutions and coping plans; **N**ow reward yourself; **D**o it every day; **S**tay strong inside).

2. Why (rationale / theory)

Throughout various updates in content, the theoretical model underpinning FRIENDS can be consistently described as Cognitive Behaviour Therapy (CBT). Based on the work of Aaron Beck (1964), CBT links thoughts with feelings and behaviour. By developing an understanding of their thoughts, individuals can learn to recognise when their thoughts are unhelpful and this can lead to more helpful thinking, moderation of feelings or changes in behaviour. The programme therefore teaches children skills in the following domains; cognitive (identify and challenge anxiety increasing cognitions), emotional (identify and manage anxiety), and behavioural (problem solve and face feared situations). The theoretical model for the prevention and early intervention for anxiety in specific relation to FRIENDS is shown in appendix A. CBT has been described as the "gold standard" in the treatment of anxiety disorders (Otte, 2011, p.413); however, this recommendation has been largely based on clinical treatment in adults, with far less known about its utility as a preventative approach with children.

3. Who (recipients)

FRIENDS is a universal intervention, and the iteration of FRIENDS delivered as part of this trial was delivered in whole-class settings. It can be adapted for use as a selective and targeted intervention as either an alternative or as an augment to universal delivery. In the specific context of this trial, FRIENDS was delivered exclusively as a universal intervention to year 5 classrooms.

4. What (materials)

Individuals responsible for implementing FRIENDS are provided a single copy of the group leader's manual to accompany the pupil workbooks (one per pupil). The group leader's manual provides a highly prescriptive syllabus, including session goals, verbatim scripts and activities lasting between 5 – 10 minutes each. The pupil workbook provides template activities aligning to the exercises in the group leader's manual. For example, pupils have to look at a series of pictures and using different coloured pencils to write 'feeling' words; then think about and write other feelings. An overview of session content can be seen in appendix B.

5. What (procedures)

FRIENDS is a taught syllabus. FRIENDS consists of 10 weekly sessions, which aim to promote various protective factors, such as recognising physiological symptoms (e.g. session 2), emotional self-management (e.g. session 4), and supporting peer relationships (e.g. session 8). Indeed, the name of the programme is designed to help pupils recall and utilise the various skills taught in the sessions (as above). Each session is designed to last 60-90 minutes, but the manual recommends a delivery model of two 30-35 minute periods instead. Additionally, there are two one-hour whole-class booster sessions reinforcing the application of FRIENDS to real life situations that can be held approximately 1 to 3 months after completing the program, and 2 information sessions for parents of approximately 2 hours length each. Booster sessions were completed, but parental sessions were not. There have been revisions across the life span of the intervention in response to feedback received from recipients and stakeholders. Later editions (used from 2010 onwards) provide more activities to increase the flexibility of programme delivery. New content has been included to cover aspects of resilience (e.g. attention training, diet and exercise) and an increased focused on addressing externalising difficulties (Barrett, Cooper & Guajardo, 2014),

As the implementation of FRIENDS is encapsulated within an existing delivery team (see 6. who), there are minor adaptations and additions specific to the context. First, there is a pre-implementation phase in which the delivery agent (commonly known as 'coach', or in this specific context, 'project officer') forms a relationship with a member of the school. Second, parental lessons are not implemented in this specific trial. Although part of the original intervention, parental sessions have not been identified as core component of the intervention and have been omitted. All other aspects of the intervention remained unchanged.

6. Who (provided)

For this trial, FRIENDS was implemented by an external delivery team; 'Salus'. In the context of the specific trial, external delivery agents are called 'project officers'. Typically, one project officer is assigned to a number of different schools (in the case of Salus, one project officer was responsible for delivering FRIENDS in up to 10 different schools).

7. How

FRIENDS is explicitly timetabled during the normal school day.

8. Where

For the purpose of this trial, FRIENDS was implemented in school, within classrooms.

9. When and how much

As noted above, FRIENDS is delivered through 10 weekly sessions. The length of each weekly session varies (e.g. between 60 – 90 minutes) and therefore, there is not a precise mapping of week = activity, as implementers are required to deliver the intervention within the constraints of school timetable. This means that dosage of specific activities listed in the instructor's manual may vary (being truncated or omitted as time allows) and/or sessions carried over to later weeks.

10. Tailoring

FRIENDS is a highly prescriptive, manualised intervention, designed to optimise fidelity of implementation. However, given variable session lengths (in contrast to fixed school timetables) and the context of implementing any 'real world' intervention in complex social environments, adaptation is inevitable. This typically takes the form of truncation (or omission), substitution or expansion of activities by implementers.

11. How well (planned)

Implementers receive 1 day of training. This training is only available to staff who are already working with children and young people in a professional capacity (e.g. teachers, pastoral support workers, youth workers). In the context of the current trial, training was provided by 'Interactive Connections' a licensed training company (now retired). Salus is now licensed to both train and provide licensed facilitators.

Background evidence

Anxiety and depression are amongst the most common mental health disorders experienced by children. It is suggested that by the age of 18, between 10-20% of children will have suffered from an anxiety disorder (Costello, Mustillo, Erkanli, Keeler & Angold, 2003; Essau, Conradt & Petermann, 2000), with many more children experiencing serious symptoms that fall below clinical criteria for diagnosis. Longitudinal data demonstrates a strong link between early experiences with anxiety and depression and a range of challenging later outcomes, including conduct problems, diagnosable mental health disorders and substance abuse in early adulthood (Woodward & Fergusson, 2001; Bittner et al., 2007). Educational underachievement is also a recognised consequence of mental health difficulty (Moilanen, Shaw & Maxwell, 2010). For instance, the UK mental health survey shows that for children experiencing emotional disorder (e.g. anxiety and depression), there is more likely to be a lag in intellectual development in comparison to their peers (those otherwise not identified with a mental health disorder). Similarly, children with generalised anxiety and depression were seen to have the highest absenteeism, with a quarter reporting more than 15 days absence in one term. Truancy was also shown as prevalent (approximately 30% prevalence in those with anxiety and depression compared to 3% among those with no disorder (Green, McGinnity, Meltzer, Ford & Goodman, 2005)). The associated consequences of early childhood mental health difficulty are immense (Belfer, 2008), representing one the largest economic and disease burdens face by current and future cohorts (Centre for Mental Health, 2010). However, most children and young people experiencing these difficulties do not have access to comparable levels of support (Dvorsky, Girio-Herrera & Owens, 2014; Kelvin, 2014), owing, in at least part, to lack of availability of support services, and the predominant 'wait to fail' model in which services only become available at a point where problems have become severely entrenched.

This presents a compelling case that school is a primary setting in which many initial concerns such as mental health can be effectively remediated prior to adulthood (Masia-Warner, Nangle & Hansen, 2006; Greenberg, 2010; Caan et al., 2014). In the UK, schools have been seen as a means to early intervention in mental health since the 1990s and was recognised in Child and Adolescent Mental Health Services (CAMHS) strategy documents which referred to mental health as 'Everybody's Business'. More recently the role of schools has been strengthened in the National Pilot CAMHS and School Link (DfE & DoH, 2015). School is universal, meaning access to all, can arguably mitigate common barriers to accessing support (e.g. time, stigmatisation and location) offering a low-cost, low threat environment (Barrett & Pahl,

2006). On a practical note, schools represent a significant proportion of time through engagement with children and young people (totalling more than 15,000 hours; Rutter, 1982). Unsurprisingly, the responsibility of the school in the promotion of children's mental health and wellbeing as a core element of their educational experience continues to grow in importance (Domitrovich, Durlak, Staley, & Weissberg, 2017; Harden et al., 2001).

Action towards the positive promotion of mental health in schools is evidenced by several national strategies, including Every Child Matters (Department for Education and Skills, 2003) and the Department of Health's 'Future in Mind' report (Department of Health, 2015). Correspondingly, schools can choose to implement a number of different frameworks and strategies to address wellbeing through a number of different ways. This may typically take the form of a tiered system of support, dependent on intended nature of the outcome and the inclusion criteria for the selected participants (Durlak & Wells, 1997; Weare & Nind, 2010):

- Highly targeted interventions (also known as tertiary prevention) provide support for pupils experiencing significant or multiple difficulties. This may typically operate with very small groups and/or on a one-to-one basis, removed from whole-class activity.
- Selective approaches (also known as secondary prevention) are designed for those showing early signs of difficulties or are considered 'at risk' by virtue of the presence of particular factors (e.g. being children of an anxious parent (Spence & Dadds, 1996)) and may typically operate in small groups, again, likely removed from whole class activity.
- Universal approaches (also known as primary prevention), designed to promote positive outcomes for all children. Although this can take the form of explicit taught curricula, likely displacing other scheduled and regular activity, universal approaches are typically delivered as whole school or whole class approaches, regardless of symptoms or risk factors (Barrett & Turner, 2004).

There has been an expansive proliferation in the availability of interventions schools may elect to implement specifically in relation to anxiety and depression (see, for example: Feldner, Zvolensky, & Schmidt, 2004; Greenberg, Domitrovich, & Bumbarger, 2001; Neil & Christensen, 2009; Horowitz & Garber, 2006). However, efforts in demonstrating their effectiveness outside of academically sponsored research trials have not kept pace (Salovey, Brackett & Mayer, 2004). Accordingly, there is a great deal of variation in the level of empirical support for individual school-based programmes designed to address or prevent anxiety and depression, with some programmes offering little to no demonstrable research. Conversely, some of the more established programmes have a substantive trial history (incorporating randomised control methodology) across multiple countries and settings. Arguably one of the most established of these is the FRIENDS programme, having been the subject of approximately 50 studies since its development in 1998, and identified as the only school-based anxiety prevention program with sufficient empirical support to warrant classification as an evidence-based practice (WHO, 2004).

The FRIENDS programme

Overview of evidence base

FRIENDS is generally accepted as one of the more effective universal programmes for addressing specific elements in child mental health (specifically, the interrelated components of worry, anxiety and depression) on the basis of a successful (international) trial history and supporting evidence for its underlying approach based on cognitive behavioural therapy (Fisak, Richard & Mann, 2011). Although the weight of evidence (see, for example Briesch, Hafermoser Sanetti & Briesch, 2010) implies FRIENDS can *potentially* be effective, more recent reviews have cast doubt on the consistency by which positive results may be achieved, and for whom, noting both implementation and evaluation limitations in research (Maggin & Johnson, 2014). For instance, evidence for the universal application of FRIENDS (in reference to the above taxonomy of intervention tiers) is partially drawn from studies that have focused exclusively on 'high risk' (i.e. those pupils demonstrated elevated difficulties at baseline) (e.g. Bernstein, Layne, Egan & Tennison, 2005; Miller et al., 2011; Siu, 2007*).* Further issues have been raised in relation to clearly

distinguishing between prevention and treatment effects within different sub-groups of children, differences in effect as a result of implementer (i.e. class teacher vs. external delivery agent), and general methodological quality, including number of independent trials, and failure to account for clustered data. In addition to this, there has been little prior consideration of FRIENDS within the English context or consideration of its effects on attainment (with the notable exceptions of Stallard et al., (2013) and Skybraina, Taylor & Stallard (2016)). These issues will be briefly considered in turn.

Differential gains

A high degree of heterogeneity is typically evident in universal school-based programmes and that might have effects over the observed effect sizes on different sub-populations (Durlak, Weissberg, Dymicki, Taylor, & Schellinger, 2011; Sklad, Diekstra, Ritter, & Ben, 2012; Wigelsworth et al., 2016). In the case of FRIENDS, a wide variety of treatment and prevention modes are included in the programme's aims, which notes significant education and resilience outcomes for non-clinically anxious children and those experiencing symptoms (allowing a return to the 'normal range' of ability to cope with day-to-day life experiences). This indicates potential differential treatment effects across a wide range of needs within a cohort. Although several studies of FRIENDS have included risk status as an a-priori subgroup for analysis, there are differences in identifying such groups, on the basis of different criteria being used across trials[2] indicating that risk status is worthy of investigation, especially as meta-analytic evidence suggests a differential effect for high risk students (Hedges g = .37) in comparison to 'low risk (Hedges g = -.26). However, the meta-analysis itself has received criticism for inconsistent selection of participant groups (including trials which exclusively focus on anxious children) (see Barrett et al., 2017), highlighting difficulties in this area.

In relation to the primary remit of the EEF of improving outcomes among children and young people from socio-economically disadvantaged backgrounds, there is evidence demonstrating a relationship between prevalence of anxiety and depression and disadvantaged social circumstances (Fryers, Melzer & Jenkins, 2003), but we are not aware of any trial that has examined socio-economic disadvantage in relation to a potential moderator of sub-group effect as a result of implementing FRIENDS. The closest comparison appears to be Stallard et al., (2013) who note the socio-demographic characteristics of the trial sample. However, most schools were below national average for %FSM eligibility, and family affluence was not seen as a significant moderator of treatment effects.

Academic outcomes

Given the primary purpose of the FRIENDS programme as a cognitive-behavioural preventative and treatment programme, there has been almost no prior consideration regarding its impact on academic attainment. Such paucity of published research is perhaps unsurprising given that attainment is not cited as one of the intended benefits of the programme, and instead can be inferred as a distal outcome as a result of more immediate improvements in mental health.

There has been just cause to hypothesise a relationship between broad psycho-social outcomes and attainment (e.g. Jennings & Greenberg, 2009) and more empirical evidence has linked depression, worry and heightened anxiety with poorer academic attainment (Keogh, Bond & Flaxman, 2006; Owens, Stevenson, Hadwin & Norgate, 2012). However, it is important to distinguish between generalised, chronic anxiety and depression, and the more situational and acute test anxiety (e.g. Putwain, Connors & Symes, 2010). In the context of school-based interventions, a number of meta analyses have shown improvements in academic attainment attributed to the implementation of social and emotional learning programmes (Durlak et al., 2011; Sklad et al., 2012; Wigelsworth et al., 2016), though as Wigelsworth et al (2016) note, programmes containing academic components were not specifically excluded from these analyses, and

[2] e.g. Strength and Difficulties Questionnaire (Iizuka, Barrett, Gillies, Clayton & Miller, 2014), self-reported anxiety (Hunt et al., 2009; parent-teacher reported anxiety (Miller et al., 2011)).

therefore improvements in attainment cannot necessarily be directly attributable to improvements in pupils' psycho-social outcomes.

Although the literature base might infer just cause to consider the possibility of an academic impact as a result of FRIENDS, early evidence does not provide support for this hypothesis. A recent cluster-randomised study by Skryabina, Taylor & Stallard (2016) found no effect of the FRIENDS programme on primary-aged pupils', maths, reading and writing). Further work is required to more closely examine the inter-relationships between anxiety, worry and depression and later academic attainment.

Implementation and process

A key consideration in the implementation of FRIENDS is who is responsible for delivery. This can be either an 'insider' such as the schoolteacher responsible for the target class or an external delivery agent, unknown to the school and pupils. FRIENDS is typically delivered by an external delivery agent, although the qualification criteria is broad[3], for practical reasons teachers or trained researchers are often selected as the delivery agents. Trials have compared para/professional (e.g. school nurses) and teacher led delivery, though findings are mixed. For instance, Barrett and Turner (2001) found no difference between psychologist and teacher delivery, but conversely, Stallard and colleagues (2014) notes an increased effectiveness when FRIENDS was delivered by 'health facilitators'. Differences in effect between these two conditions have been shown to be notable. Briesch et al. (2010) shows an average difference of ES=0.34 (averages drawn across several trials show an external delivery condition with a mean of ES=0.56, whereas teacher-led conditions showed a mean of ES=0.22). Although Stallard et al. (2014) hypothesise that implementer characteristics may be responsible, no study of FRIENDS to date provides empirical evidence to indicate why such a difference might occur, and no study of FRIENDS has yet considered the characteristics of implementers as a potential explanatory mediator in understanding the differences in effect. This element is worthy of further consideration given the implications for additional scale up and cost of the intervention.

The UK Evidence base

CBT has been described as the "gold standard" in the treatment of anxiety disorders (Otte, 2011, p.413) however, this recommendation has been largely based on clinical treatment in adults. Far less is known about the use of CBT when utilised as a preventative, universal intervention, especially within the English education system (Stallard, et al., 2014). Findings from CBT use in children have tended to be community focused (Mychailyszyn et al., 2011) and/or been used with specialist populations, such as those identified with Autism Spectrum Disorder (e.g. McNally Keehn, Lincoln, Brown, & Chavira, 2013). The UK evidence base for FRIENDS is limited, currently restricted to a small number of publications resulting from a recent trial of the programme (e.g. Stallard et al., 2014). Main effects were found for self-reported anxiety and depression scores, though no effects were found for the parent or teacher measures. These results bear further investigation, as the trial was not adequately powered to examine the subgroup of highly anxious pupils (therefore failing to fully consider differential gains in treatment vs. prevention), and there was little to no consideration of process evaluation or cost effectiveness. This is particularly important, as although Stallard noted that health professionals were, on average, more likely to generate greater effects than trained teachers, there is little in the way of empirical evidence to indicate why this might be so, or the relative cost/reward for different implementers. There is also an opportunity to examine the theory of change more closely, moving beyond summative outcome effects, in an effort to understand more closely how anxiety and depression themselves impact upon later attainment.

As a final note, we believe that the current societal and education policy context is ideal for a trial of this type. There are serious concerns in UK society regarding the health-related burden of child mental health and the considerable associated economic and societal costs (Belfer, 2008). The mental health of children

[3] e.g. an undergraduate university degree in a relevant discipline and/or appropriate professional backgrounds such as, psychology or, in the case of Stallard et al. (2014), nursing

and young people is being increasingly recognised as a global priority. FRIENDS aligns closely with previous calls for early intervention both within the academy (e.g. Levitt, Saka, Romanelli & Hoagwood, 2007) and from a policy perspective, as the government has made early intervention for mental health and wellbeing in children a priority and focus (Department for Education, 2015d).

Evaluation objectives

The primary aims of this evaluation were to examine the impact of the FRIENDS for Life programme specifically in relation to its impact on primary school children's; i) academic attainment at Key Stage 2, and; ii) health related outcomes (specifically worry, anxiety, depression and behaviour). In addition, subgroup effects were examined for; i) children eligible for free school meals, and; ii) children with elevated internalising difficulties at baseline. The analysis also examined whether the way in which FRIENDS is implemented or the characteristics of the implementers is associated with variability in outcomes.

a) Universal effects

H1: 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.

Effects outlined in H1 and H2, will also be present for:

H1a. Maths scores only

H1b. Reading scores only

H2: 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

Effects outlined in H2, will also be present for:

H2a. Self-rated total anxiety and depression score

H2b. Teacher rated emotional symptoms

H2c.Teacher rated difficulties conduct problems

b) Subgroup effects

H3a: For the outcome variables listed in H1 (a-b) & H2 (a-c); There will be a significant interaction between allocation to condition and Free school meal status (EverFSM)

H3b: For the outcome variables listed in H1 (a-b) & H2 (a-c); There will be a significant interaction between allocation to condition and pupils identified with elevated internalising symptoms (scoring in the 'slightly raised' or above category for teacher-rated emotional symptoms at baseline ("at risk")).

c) Implementation

H4: Variation in implementation fidelity (H4a), quality (H4b), engagement (H4c), programme and reach (H4d) will moderate education-related outcomes (H1) in schools implementing FRIENDS.

H5: Variation in implementation fidelity (H5a-d) will moderate health related outcomes (H2) in schools implementing FRIENDS

H6: Variation in implementer characteristics, specifically emotional self-efficacy (H6a), teaching efficacy (H6b) and views of social and emotional learning (H6c) are related to pupil outcomes (attainment / self-rated worry)

Further details of the protocol can be found at: **https://educationendowmentfoundation.org.uk/our-work/projects/friends.** In comparison to the initial protocol, the order of hypotheses were rearranged for clarity, but all research questions are still covered.

Ethical review

All trial documentation and procedures were reviewed and approved by Manchester's University Research Ethics Committee (REF: 16012). The trial was also prospectively registered with the ISRCTN: reference number SRCTN13721202.

Consent/assent was managed through multiple stages of the research:

Opt-in by schools: Participating schools signed a Memorandum of Agreement (MoA) indicating their willingness to participate. The MoA contained detailed information about what participation entailed (e.g. data collection procedures and requirements). The MoA explained the nature of the RCT (e.g. that only half of participating year 5 classes in each school would receive FRIENDS, and that this would be determined by a random allocation procedure), in addition to what schools could expect in return for their participation (e.g. FRIENDS in one half of the year 5 classes). Single form entry schools, were randomly allocated to receive either FRIENDS or £1,000 payment at the end of the school year).

Parental consent: Participating schools distributed consent forms to the parents and carers of all eligible pupils (those in year 5 during 2015/16). This was achieved through multiple methods (e.g. letters sent home, texting services, email, school websites) as befitting the normal communication routes per school. Parents and carers who did not want their child to participate in the trial completed an 'opt-out' section on the consent form which was returned via a freepost address to the University of Manchester or alternatively correspondence was emailed direct to the principle investigator. In total 71 parents exercised their right to opt their children out of the trial.

Pupil assent: Children were provided with information about the study (including their guarantee of anonymity and right to withdraw) and were asked to give their assent to participate. No child declined assent or exercised their right to withdraw from the study.

Case study focus groups: An additional consent process was followed for any children attending case study schools and who were nominated to participate in the pupil focus groups. This followed an explicit (e.g. opt-in, as opposed to opt-out) parental consent procedure. Standard protocols were followed in respect to confidentiality and disclosure during the conduct of these focus groups. Children were assured that their responses would remain anonymous and confidential except in the event of the disclosure of information indicative of a child protection issue, at which point the school's designated safeguarding lead would have to be informed. No such disclosures took place.

Anonymity and confidentiality were ensured through data management procedures as follows: security for online surveys (e.g. teacher-rated SDQ) was ensured using hypertext-transfer-protocol-secure data encryption. Data matching (e.g. across time) was achieved through the use of a unique pupil number. All qualitative data were anonymised during the transcription process, with pseudonyms given to any personally identifying information. The University of Manchester and Microsoft Best Practice guidelines for data storage were followed, ensuring that data was held safely on secure drives behind internal and external firewalls, and physical transportation prohibited (e.g. flash drives).

Memorandum of Agreement, and parental information sheet and consent form can be found in appendix C.

Project team

Evaluation team:

Michael Wigelsworth: Principal investigator and lead author

Garry Squires: Specialist in cognitive behavioural therapy

Liz Birchinall: Specialist in primary education

Ann Lendrum: Specialist in implementation and process evaluation

Afroditi Kalambouka: Trial manager and qualitative analyst

Louise Black: Research assistant and qualitative analyst

Patricio Troncoso: Research assistant and quantitative analyst

Joao Santos: Research assistant and quantitative analyst

Emma Ashworth: Research assistant

Philip Britteon: Specialist in outcome modelling

Implementer:

Salus is a Community Interest Company who have over 20 years expertise in delivery of services for young people, families, schools and communities. In particular, Salus have experience of delivering the FRIENDS programme in Kent, and have worked with the original Australian developers of the programme. Accordingly, they have the license to deliver the intervention.

Trial registration

The trial was registered with ISRCTN (REF: 13721202). Details can be accessed here: www.isrctn.com/ISRCTN13721202

Methods

Trial design

A two-year cluster-randomised design was used, with class as the unit of randomisation. This design is advantageous in terms of the balance between scientific rigour, ethical considerations, and goodness-of-fit with the study aims and hypotheses. Class was selected as the unit of randomisation for a number of reasons; i) the intervention is typically delivered at class level, ii) the implementation by an external delivery agent and internalised context of the content (e.g. managing internal feelings) minimised risks of contamination to comparison classes, iii) allocation by class allowed most schools to receive the intervention, allowing for favourable pupil:school ratios for purposes of statistical power and continued participation by schools (e.g. drop out in similar trials is more typically associated with comparison schools).

Year 5 classes were randomly allocated to one of two trial arms: 1) To receive FRIENDS (intervention) or; 2) continue as normal (comparison). Classes in the intervention arm received the FRIENDS curriculum, commencing March 2016 for a period of 10 sessions. Classes allocated to the comparison condition continued their normal practice for the same period (see 'randomisation' for more details of the allocation process).

At the end of the school year (July 2016), schools were free to decide whether to continue (in the case of the intervention schools) or to start (in the case of usual provision schools) implementing FRIENDS. However, schools were not to implement or continue FRIENDS in the target cohort, who moved to year 6. This was to allow a year of normal practice for both arms of the trial, prior to collating Key Stage 2 data in 2017.

Participant selection

As 'FRIENDS for Life' is intended as a primary-aged school based intervention, the trial was comprised of a single cohort of year 5 pupils. Eligibility criteria for participation required schools to be state-funded primary schools (e.g. community, foundation or academy) who have not previously implemented FRIENDS. The implementation team (Salus) managed the initial identification of schools. Schools within the Kent Local Authority were approached using a range of techniques, including publicity events (e.g. a one-day event organised to inform schools of the project), email (of project flyers and associated materials) and telephone contact. Initial expressions of interest were formalized through the evaluation team (University of Manchester) by requiring schools to complete the aforementioned MoA. Participating schools that signed the MoA were then required to provide a minimum of 85% of baseline survey data prior to randomisation.

Outcomes measures

Primary Outcome Measures

The primary academic outcome measure was Key Stage 2 Maths and Reading combined (i.e. summed - equal weighting), as measured by the standardised curriculum tests conducted in July 2017. Data were obtained from Kent County Council (KCC). The specific variables used were "Reading Derived Scaled Score" and "Maths Derived Scaled Score" (variables KS2_READSCORE and KS2_MATSCORE respectively in KCC) and these were averaged for each pupil.

The primary heath outcome measure was self-rated worry as measured by the Penn State Worry Questionnaire for Children (PSWQ-C) (Meyer, Miller, Metzger, & Borkovec, 1990). This paper-based self-report measure was deployed at baseline (March 2016) and at post-test (December 2016). The PSWQ-C is a 14-item self-report questionnaire with responses scored on a 4-point Likert scale. An example item is "My worries really bother me", with responses: "never" = 0, "sometimes" = 1, "often" = 2, and "always" = 3. Additionally, 3 items were reversed coded. Cronbach's alpha for the baseline measures was 0.88 and for the post-test was 0.9.

Secondary Outcome Measures

Self-reported total anxiety and depression score (see Analysis section: H3a) was assessed by taking the total score from the Revised Child Anxiety and Depression Scale (RCADS 25) (Ebesutani, Korathu-Larson, Nakamura, Higa-McMillan & Chorpita, 2012). This paper-based measure was deployed at baseline (March 2016) and at post-test (December 2016). The RCADS25 is a 25-item self-report questionnaire with responses scored on a 4-point Likert scale. An example is: "I feel sad or empty", with responses: "never" = 0, "sometimes" = 1, "often" = 2, and "always" = 3. Cronbach's alpha for the baseline measures was 0.9 and for the post-test was 0.91.

Data pertaining to the PSWQ-C and RCADS-25 were collected simultaneously in a single, counter-balanced questionnaire pack, utilising whole class administration, and a strict, standardised administration protocol (see appendix D). Completed packs were stripped of identifying information (though retained an anonymous code for matching purposes) and couriered to the University of Manchester for scoring. A team of research assistants (blinded to condition) scored the completed questionnaire packs. Each pack was double-marked. There was 100% inter-rater reliability.

Both teacher rated internalising difficulties and teacher rated externalising difficulties (see Analysis section: H3b, H3C) were measured using the teacher version of the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 2001). As classes moved from Year 5 to Year 6 during the trial, the pre and post intervention teacher ratings were often completed by different teachers. This measure was deployed online at baseline (March 2016) and at post-test (December 2016). Data were collected through a secure online platform, by which schools were emailed individual password and log-on details. The teacher version SDQ is a 25-item questionnaire with responses scored on a 3-point Likert scale, with 5 items needing reverse coding. An example item is: "Often unhappy, downhearted…" with responses: "not true" = 0, "somewhat true" = 1, and "certainly true" = 2. Internalising difficulties and externalising difficulties are both constructed from a 10-item subdomain. Cronbach's alpha (total score) for the baseline measures and the post-test was 0.91.

Power and Sample size

The initial aim of the study was to recruit 110 classes from 77 schools (an estimate based on the number of single entry schools in the implementation area – see appendix E). All calculations assumed*: N= 28 per cluster (Department for Education, 2015a); ICC (class level) = 0.17[4], Power=0.8, Alpha=.05,[5] proportion of single form entry = 51%[6]*

Recruitment exceeded targets, resulting in a total of 128 classes (81 schools), resulting in a minimum detectable effect of 0.239. Nevertheless, 6 classes (2 schools) did not meet the inclusion criteria for randomisation (failure to return minimum level of baseline data), resulting in a total of 122 classes (79 schools), which is the total included in the main analysis. Post-hoc power calculations, including pre-test data give a 2-tailed minimum detectable effect size (MDES) of 0.151. In the interest of clarity, the MDES protocol, randomisation and analysis is presented in table 4 (see 'impact evaluation).

Randomisation

A cluster-randomised design was utilised, using matched pair or 'randomised block'. This is where each unit of allocation (i.e. class) was first paired on the basis of a blocking variable, in this case type of entry form – single; double; triple; quad. Within each pair, classes were randomly assigned to either receive FRIENDS or act as a comparison. During the randomisation, KS1 and baseline 'total anxiety and depression' scores (see: outcome measures) were used for minimisation.

[4] Estimate drawn from the largest academic ICC (KS2 results in writing) from the PATHS trial

[5] 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.

[6] Based on approximate data of Kent school sizes, provided by Salus.

For each participating school:

- Single form entry: Randomly allocated to receive either FRIENDS or £1000 (delivered at the end of the academic year).
- 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.
- Four form entry: Random allocation of two classes to receive FRIENDS, two classes to serve as usual practice.

Regarding split classes (e.g. a merged year 4/5 class); they were treated as classes of 15 years 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.

Analysis

Primary intention-to-treat (ITT) analysis (H1 - 2)

An ITT analysis was conducted for H1 and H2. This was done in accordance to intention-to-treat principles, e.g. ignoring noncompliance, protocol deviations and other events that take place after randomisation (Gupta, 2011). This analysis was carried out through fitting 3-level (schools, classes, pupils) hierarchical models to account for the nested nature of the data. Allocation to the FRIENDS intervention was at the class-level, which means that a multilevel model must be fitted to account for the different sources of variation. The school-level was included, as not doing so would have potentially biased the results, since two classes within the same school are more likely to be similar than two classes in different schools. From a statistical point of view, ignoring the school level would spuriously inflate the class variance and hence overestimate the effect of FRIENDS. All models were fitted using MLwiN Version 2.36.

First, we implemented an initial unconditional (empty) model to ascertain the amount of variance attributable to schools, classes and pupils, followed by models with treatment allocation (e.g. FRIENDS vs. comparison) included at class level. An intervention effect was noted if the coefficient associated with the trial group variable was statistically significant. This was subsequently converted to Hedge's *g* accounting for varying cluster sizes, as per EEF reporting guidelines (Hedges, 2007; Tymms, 2004).

Further details of the ITT, subgroup and implementation analysis models, including equations used for these estimations, are provided in appendices H and I.

Our principal analyses used fully observed data, as per EEF guidelines. Subsequently, we assessed the sensitivity of our findings by comparing both fully and partially observed data though the use of multiple imputation (MI). Commentary regarding any differences as a result of MI can be seen under each model within the impact evaluation chapter, which also contains details on how missing data was treated.

Subgroup analysis (H3)

Although Intent-to-Treat (ITT) is seen as the most unbiased form of analysis for randomised control trial methodology (Gupta, 2011), this approach fails to account for any heterogeneity in response between groups of participants (or 'recipients') of an intervention (Farrell, Henry, & Bettencourt, 2013). Being able to identify differential responses to 'treatment' can provide key information in how to maximise benefits of a programme (Rothwell, 2005). However, analyses must be pre-defined, carefully justified in relation to theory and evidence relating the intervention, and limited to a few important questions in order to avoid potential bias and over-interpretation of intervention effects (Petticrew et al., 2012).

Further models for each hypothesis were constructed including risk status (binary indicator) as a main effect and a cross-level interaction term (e.g. Risk Status*Allocation). An intervention effect at the subgroup

level will be noted if the coefficients associated with the interaction terms noted above are statistically significant. These will subsequently be converted to Hedge's g, as per EEF reporting standards. As per hypotheses 1 & 2, fully and partially observed data were compared, with a summary commentary provided.

This was also undertaken for pupils both currently and previously eligible for FSM (over the previous six years of schooling (EVERFSM).

Implementation Analysis (H4-6)

Implementation analysis was conducted on the FRIENDS trial arm only. Implementation variability scores from observations (fidelity/adherence, quality, reach and participant responsiveness) and self-reported implementer characteristics (emotional self-efficacy, teaching self-efficacy and views on social and emotional learning) were transformed to a z score prior to being analysed - see 'implementation and process evaluation' below for details on the measures themselves). The analysis was conducted for each hypothesis through the construction of 3-level hierarchical models[7] (pupil, class, Project Officer[8]) to account for nested nature of dataset using MLWin Version 2.36. Given that there are no agreed thresholds for implementation ratings (for example, it is not possible to definitively say what counts as an 'acceptable' level of fidelity/adherence), our approach was to identify and significant predictors of student outcomes and a relative comparison of effect size between the implementation variables. Following from the unconditional models created in the ITT analysis (see also appendix I), the models for each hypothesis were specified as follows:

H4 (implementation variability – attainment outcome): After checking for multicolinearity, the scale totals of each implementation variable were entered into the model simultaneously as scale-type predictors, with KS2 maths and reading score (combined) as the outcome.

H5 (implementation variability – health outcome): After checking for multicolinearity, the scale totals of each implementation variable were entered into the model simultaneously as scale-type predictors, with self-rated worry as the outcome.

H6 (implementer characteristics – attainment and health outcomes): After checking for multicolinearity, the scale totals for each of the implementer characteristics were entered into the model simultaneously as a scale-type predictors. Two models were run, one for attainment as the outcome variable and the one for self-rated worry as the primary outcome variable.

Following advice from Althouse (2016) there was no correction for multiple inferences primarily on the basis of the exploratory nature of the analyses and relatively small number of hypotheses tested, especially given the clear delineation between academic and health based outcomes.

Effect size calculation

In all cases, effect sizes are reported using Hedge's g (Cohen's d bias corrected) as per EEF specifications (Tymms, 2004) and 95% confidence intervals for the effect sizes were estimated as described by Fritz, Morris and Richler (2012), unless otherwise stated (for more detail see appendix I).

Implementation and process evaluation

Previous process evaluations of FRIENDS have typically been very limited, restricted to adherence to protocol, responsiveness, and limited study of acceptability (e.g. Kösters et al., 2012; Stallard et al., 2013). We would argue that this gap in knowledge is problematic for the continued scale up of the programme. Understanding implementation is vital for several reasons; not least that variability can affect outcomes

[7] *An alternative parameterisation was attempted in which schools were specified as fixed effects (instead of random effects, i.e. a second higher level above class) and robust standard errors were estimated. For the sake of parsimony, the 3-level specification was kept.*

[8] Project officer was included a level, not a fixed effect; because with PO as fixed effect there is perfect collinearity

(Humphrey, 2013. Understanding how and why schools vary their implementation, and what effect this might have on outcomes is vital.

Case studies of implementation

Richer, contextual detail regarding the implementation of FRIENDS was captured through the participation of 10 case study schools, selected in consultation with Salus on the basis of maximum variation sampling. Case study data included a number of inter-related data strands:

Independent observations of implementation: Addressing potential concerns of demand characteristics and impression management (Humphrey, 2013) through the use of self-report data, implementation of the intervention was recorded through direct observations of the class by trained researchers, as research demonstrates this to be much more closely related to intervention outcomes (Domitrovich, Gest, Jones, Gill & Sanford DeRousie, 2010). Observations were carried out once per project officer (10 observations, carried out in 10 separate classes) allowing inclusion of implementation variability in the analyses (e.g. to what extent implementation factors affect outcomes). We employed custom quantitative assessment tools, as in previous research (Humphrey, Barlow & Lendrum, 2018).

Construction of the observation schedule was based on theoretical and empirical implementation literature (Domitrovich et al., 2010; Berkel, Mauricio, Schoendelder & Sandler, 2011) and comprised of a total of 13 statements rated on a 10-point scale ('extremely poor to extremely good') with accompanying notes and observations, across 4 subdomains; fidelity/ adaptation (e.g. To what extent does the implementer follow the structure and sequence of activities outlined in the lesson guidance?) quality (e.g. How clearly does the implementer explain key concepts and activities in the lesson?); engagement (e.g. Rate the extent to which children in the class actively participate in the lesson activities (e.g. joining in role plays, answering questions) and reach (e.g. approximately what proportion of the class are present throughout the lesson?). A copy of the observation schedule is included in appendix G.

The schedule was explained and discussed in detail with the two research assistants responsible for observations, prior to the scheduled school visits. Video footage of an example FRIENDS session was captured by the implementers (Salus) from a local school implementing FRIENDS that was not part of the main trial. A process of refinement and agreement was conducted over several sessions, by which a shared understanding of the implementation indicators were developed through multiple reviews of the video footage. Shared understanding was agreed when both raters arrived at full agreement in their respective scoring.

Project Officer self-report: We collated brief report diaries from each PO, indicating what activities had been completed, and whether POs had deviated from the prescribed curriculum. This data was used as part of the qualitative process evaluation (see below) as well as constructing the descriptive 'fidelity' map of implementation across the 10 sessions (see case study data).

Interviews and focus groups with key stakeholders: Semi-structured interviews were conducted with a number of stakeholders including head teachers, members of the senior leadership team, pastoral leads and class teachers. Topics included reasons and criteria for why schools chose to participate in the research, why FRIENDS is seen as beneficial, as well as to explore the wider context influencing decisions of implementation (e.g. perceptions of social validity, acceptability and feasibility). Pupils were also included in the research through the use of small focus groups.

Usual practice survey: A critical consideration in the implementation of the FRIENDS curriculum is the extent to which it displaces, alters, or enables existing practice, both related, and independent to the programme objectives. FRIENDS may replace other class-based mental health initiatives, or be used to link and supplement other programmes (e.g. anti-bullying). This information was captured through a questionnaire completed by a member of school staff (either SLT and/or SENCO) from each participating school, detailing any provision in place by the time the intervention started, to assess the unique contribution of FRIENDS. This survey included a request for CAMHS referral data to assess possible pupil-

level changes in school provision. This survey was repeated post-intervention to establish whether there have been any changes (a copy of the usual practice survey is included in appendix F).

Qualitative Analysis

Qualitative data drawn from Project Officer self-report, observations, interviews and focus groups was subject to a thematic analysis using the principles and processes outlined by Braun and Clarke (2006) (e.g. familiarization, generating initial codes, searching for themes, reviewing themes, defining and naming themes, report production). A hybrid approach was taken, informed by the theoretical and empirical literature underpinning the observation schedule (Domitrovich et al., 2010; Berkel, Mauricio, Schoendelder & Sandler, 2011) in order to provide context and background the quantitative findings. Additional insight was probed through additional orientating concepts of implementation (specifically factors affecting implementation, and the social validity & perceived attractiveness of FRIENDS to stakeholders, and perceptions of impact (both congruent with programme theory and wider/unexpected benefits). Emergent coding allowed for the emergence of unanticipated themes specific to this project/context.

Costs

The cost of delivering FRIENDS in the context of the trial is exclusively based on estimates provided by Salus CIC as the implementer for the trial. Although schools assigned to the intervention received the intervention and materials free of charge, which included: (i) delivery of the full 10 sessions of the FRIENDS curriculum by a trained and licensed external implementer (in the case of the current trial; Salus), (ii) a group leaders' manual for delivering the session, and (iii) an accompanying activity workbook for each child, costs below are based on 'normal circumstances' outside of a trial infrastructure. Costs were calculated based on the direct costs of providing the intervention and per-pupil costs for associated resources (i.e. the activity workbook). We assume a three-year deployment of FRIENDS (though as there is little in the way of initial set-up costs (e.g. no staff training), this amount remains stable over a three-year period). We provide costs for single-form and double-form entry, as discount is provided when an implementer is able to deliver two sessions within the same school on the same day, assuming an average class size of 27 pupils per class (DfE, 2017).

Timeline

Contact with schools and recruitment took place between October and December 2015. In order to accelerate procedures and achieve the required sample, recruitment was carried out by the implementer (Salus) due to the organisation's knowledge of the locality and schools as well as its previously established links with schools owing to provision of other services.

Collection of baseline assessment data took place in February 2016. Once the baseline data was scored, classes were randomised to two arms in March 2016 and the schools were informed. Implementation of the programme started in April 2016 and was carried throughout the next two school half-terms till July 2016. There were two booster sessions at the start of Autumn 2016, and the post-intervention data was collected in December 2016.

Table 3: Timeline of project activities

Date	Activity
2015	
Aug-Sept 2015	Ethical Clearance Finalisation of trial protocol
Oct-Dec 2015	School recruitment
2016	
Jan	Collection of MoA forms
Feb	Baseline (T1) outcome data collection & scoring
Mar	Randomisation
April	Recruitment of case study schools
April-July 2016	FRIENDS Implementation (10 x sessions in each class)
May-June 2016	Case study fieldwork (10 schools)
Dec 2016	Post-intervention (T2) outcome data collection
2017	
Jan-May 2017	Qualitative data analysis
Sept 2017	Request for KS2 attainment data
Sept 2017-Feb 2018	Quant Data cleaning and analysis / report write up

Impact evaluation

Participants

An initial cohort of 129 classes (82 schools) were recruited, of whom 122 (3,284 pupils) were randomly allocated (62 usual provision; 60 FRIENDS). In terms of the primary outcome (Key Stage 2 combined Maths and English scores) data were available for 3,010 pupils (92%). All missing cases were due to pupils opting out or due to lack of match in the NPD. Taking into account the lower than expected ICC and the inclusion of a pre-test, the resultant complete case analysis of academic data was powered for an MDES of 0.151.

In comparison to the academic data, health-related response rates varied slightly due to the difference in data collection techniques (collected as part of in-class questionnaires). For self-rated worry, data were available for 3,087 pupils (94%). Missing cases were due to pupils opting out or being absent during data collection. The resultant complete case analysis of health-related data was powered for an MDES of 0.125

Table 4: Minimum detectable effect size at different stages of the trial: Academic outcomes data

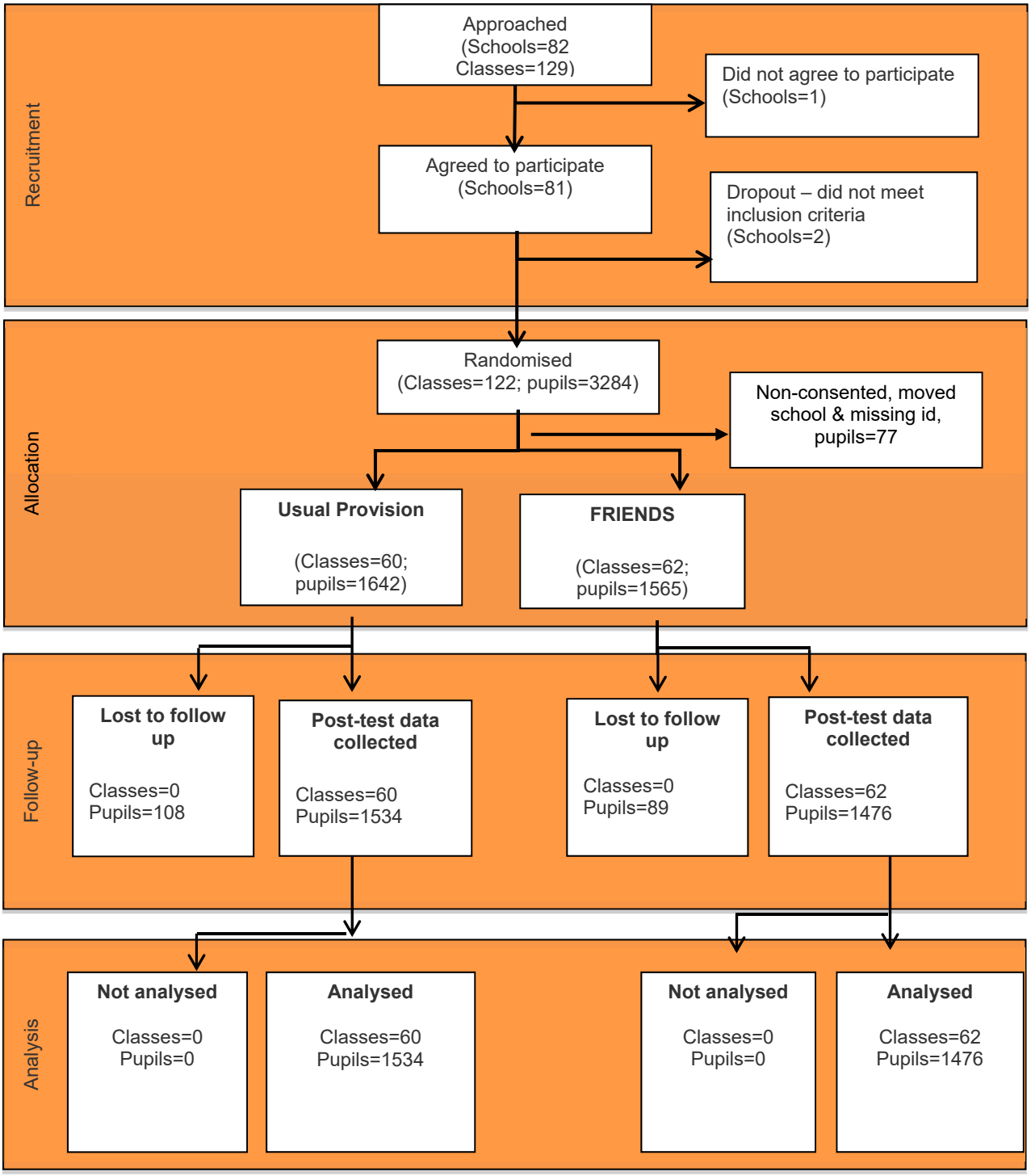
Stage	N [classes/pupils] (n=intervention; n=control)	Correlation between pre-test & post-test	ICC	Power	Alpha	MDES
Protocol	110/3,080 (55/1540; 55/1540)	-	0.17	80%	0.05	0.239
Randomisation	122/3284 (62/1565; 60/1642)	-	0.17	80%	0.05	0.227
Analysis	122/3010 (62/1476; 60/1534)	0.73[9]	0.11	80%	0.05	0.151

Table 5: Minimum detectable effect size at different stages of the trial: health outcomes (self-rated worry)

Stage	N [classes/pupils] (n=intervention; n=control)	Correlation between pre-test & post-test	ICC	Power	Alpha	MDES
Protocol	110/3,080 (55/1540; 55/1540)	--	0.02	80%	0.05	0.125
Randomisation	122/3284 (62/1565; 60/1642)	--	0.02	80%	0.05	0.124
Analysis	122/3010 (62/1476; 60/1534)	--	0.037	80%	0.05	0.125

[9] The timing of the study took place during the English Government's adoption of a new national curriculum and accompanying Key Stage 2 (KS2) SATS exams. At the time of the project set-up, no data were available regarding the estimated correlation between 'old style' Key Stage 1 tests (pre-test) and the new KS2 STATS. Estimates of the effect size at protocol and randomisation omitted KS1 as a pupil-level covariate (as there was no data for a reliable estimate), instead adding KS1 data in the post-hoc analysis, allowing a more refined MDES.

Figure 1: Participant flow diagram



School, class and pupil characteristics

Baseline demographic, attainment and wellbeing outcomes for the final sample are presented below (tables 6 & 7). The composition of trial schools (located in the South-East of England) mirrored national averages in terms of size, attendance and average attainment. Trial schools were seen to be higher in the proportion of pupils eligible for free school meals. Trial schools showed a much small number of pupils with special education and additional needs or disabilities and those speaking English as an additional language. Balance on key observables between trial arms was considered to be good as ES differences were observed to be to be negligible (table 7).

Table 6: Characteristics of the participating schools

Variable School-level (continuous)	National average	Randomised schools mean (SD, p)
Number of full-time equivalent (FTE) pupils on roll	279[10]	281.58 (162.95, p=0.888)
Sex – percentage of male students	50.97[8]	46.44 (13.68, p=0.004)
Overall absence (% half-days missed)	0.039[11]	0.04 (0.01, p=0.125)
Proportion of pupils eligible for free school meals	0.141[12]	0.26 (0.17, p<0.001)
Proportion of pupils speaking English as an additional language	0.206[8]	0.09 (0.09, p<0.001)
Proportion of pupils with special educational needs and disabilities	0.135[9]	0.01 (0.01, p<0.001)
Proportion of pupils achieving Level 4 (or above) in English and Maths at end of KS2[13]	0.853[14]	0.82 (0.11, p=0.008)

[10] Department for Education (2017a)
[11] Department for Education (2017b)
[12] Department for Education (2017c)
[13] Since 2016, this refers to the proportion of pupils reaching the expected standard in English and Maths at the end of KS2. National average = 0.696. Department for Education (2016b).
[14] Department for Education (2015c)

Table 7: Characteristics of the randomised classes and pupils

Class characteristic	National average	Intervention group (N=62) Mean (SD)	Usual provision (N=60) Mean (SD)	Balance at randomisation (Cohen's d, 95% CI)[15]
Sex – percentage of male students within class	50.97[8]	47.01 (14.32)	48.17 (10.72)	-0.091 (-0.45; 0.267) (negligible)
Class size	27.1[8]	24.09 (6.78)	25.64 (6.88)	-0.227 (-0.618; 0.102) (small)
FSM – proportion of pupils eligible for free school meals within class	0.141[8]	0.14 (0.12)	0.16 (0.14)	-0.129 (-0.478; 0.220) (negligible)
SEND– proportion of pupils with special educational needs and disabilities within class	0.135[10]	0.14 (0.11)	0.18 (0.14)	-0.31 (-0.660; 0.042) (small)
Pupil characteristic	**National average**	**Intervention group (N=62) Mean (SD)**	**Usual provision (N=60) Mean (SD)**	**Balance at randomisation (Cohen's d)13**
Baseline self-rated worry (PSWQ)	n/d	16.95 (8.46)	16.82 (8.35)	0.015 (-0.059; 0.123) (negligible)
Baseline self-rated total anxiety and depression (RCADS)	n/d	21.57 (12.21)	20.96 (12.03)	0.05 (-0.024; 0.123) (negligible)
Baseline teacher rated emotional symptoms (SDQ)	n/d	2.49 (2.78)	2.39 (2.71)	0.038 (-0.035; 0.112) (negligible)
Baseline teacher rated conduct problems (SDQ)	n/d	2.54 (3.11)	2.51 (3.09)	0.011 (-0.062; 0.085) (negligible)
KS1 average points score for English/Maths (2015)	16.1[16]	15.96 (3.3)	15.86 (3.37)	0.029 (-0.044; 0.102) (negligible)

Missing data

As educational data is drawn from the National Pupil Database (NPD), there was minimal missing data for the primary outcome. For the health outcome data, completely missing data was also minimal, given excellent responses at baseline and post-test. Differences between complete and missing cases were examined to establish patterns of missingness. Logistic regression was used to predict missingness, whereby each child is coded as providing complete (0) or incomplete (1) outcome data, with other study data as explanatory variables (as seen in Pampaka, Hutcheson & Williams, (2017)). These analyses revealed that missingness in the primary outcome (KS2 average derived scaled score in Maths and English) is significantly associated with raised externalising SDQ scores. Hence, a principled approach to deal with missingness was essential.

Analyses were conducted using complete cases and a sensitivity analysis using multilevel multiple imputation (via the REALCOM-Impute extension to MLWin) was later conducted. Multiple imputation procedures were carried out in REALCOM-Impute, using the missing at random (MAR) assumption (Carpenter, Goldstein, & Kenward, 2011). This enabled us to include both partially and completely observed cases of all schools and pupils in the analysis, thereby reducing the bias associated with attrition. Missing values were present for KS2, KS1, FSM eligibility, SDQ internalising and externalising scores, as

[15] Thresholds for quantification of effect size defined by Cohen (1992, p157). |d|<0.2 "negligible", |d|<0.5 "small", |d|<0.8 "medium", otherwise "large".
[16] Department for Education (2015b)

well as PSW and RCADS scores. For the case of the main analyses of this report, there is a total of 8% missing pupil-level data (listwise) (7.3% for the usual practice arm and 8.7% for pupils allocated to FRIENDS). The imputation models sought to impute values for all the above mentioned variables. We used all variables included in the model of interest that did not contain missing values as auxiliary variables, i.e. the constant or intercept, demographic variables (FSM eligibility) and class-level data (allocation and the minimisation variables). Following general guidelines about multilevel multiple imputation (Carpenter et al., 2011), REALCOM-Impute was set to run for 5,000 iterations, with a burn-in period of 500 iterations. We stored 10 imputed datasets, allowing for 500 iterations to run between them. Results presented in this report correspond to the analyses using complete cases, but pooled results from the multilevel analysis models of multiply-imputed datasets can be inspected in the appendices section (appendix J).

Outcomes and analysis

Table 8 below provides basic descriptive statistics for pupil-level outcome data at baseline and post-test (histograms and QQ plots can be seen in appendix L. Data does not appear to be indicative of an intervention effect for any outcome.

Table 8: Mean (SD) pupil outcomes at baseline and post-intervention follow-up.

	Intervention group		Usual provision group	
	Baseline Mean (SD)	Follow-up Mean (SD)	Baseline Mean (SD)	Follow-up Mean (SD)
KS1 average point score	15.96 (3.3)	-	15.86 (3.37)	-
KS2 maths and English combined score	-	104.06 (7.79)	-	104.07 (7.74)
Maths score	-	103.76 (7.89)	-	103.74 (7.91)
Reading score	-	104.37 (8.74)	-	104.4 (8.66)
Self-rated worry (PSWQ)	16.95 (8.46)	16.06 (8.53)	16.82 (8.35)	15.8 (9.03)
Self-rated total anxiety and depression (RCADS)	21.57 (12.21)	19.36 (11.87)	20.96 (12.03)	19.33 (12.43)

Table 9 summarises the universal effects (ITT) for H1 & H2.[17]

[17] A note on interpreting significance: Readers of this report may be familiar with assessing significance by examining whether confidence intervals cross zero (indicating significance). However, it is important to note that this method is only accurate for beta coefficients. Although effect sizes (as reported in the tables below) are derived from beta, they are not equal to this value, and therefore not comparable in the same way if attempting to assess significance through confidence intervals crossing zero. Instead, we direct readers of tables 9 & 10 to the right-hand column, which already provides values for statistical significance (as per standard conventions, any value less than .05 is considered significant).

Table 9: Universal effects - Intention to treat analysis

Outcome	Intervention group		Usual provision group		Hedges *g* (95% CI)	*p*
	n	Mean (95% CI)	n	Mean (95% CI)		
H1: Main effect of intervention - Academic (ITT)						
Maths and reading combined score	1476	104.06 (103.65, 104.47)	1534	104.07 (103.67, 104.46)	0.007 (-0.065, 0.078)	0.908
Maths scores only	1476	103.76 (103.34, 104.17)	1534	103.74 (103.33, 104.14)	0.003 (-0.068, 0.075)	0.953
Reading scores only	1476	104.37 (103.91, 104.83)	1534	104.4 (103.95, 104.84)	-0.013 (-0.084, 0.059)	0.83
H2: Main effect of intervention – health outcomes (ITT)						
Self-rated worry (PSWQ)	1476	16.06 (15.6, 16.53)	1534	15.8 (15.3, 16.3)	0.052 (-0.019, 0.124)	0.269
Self-rated total anxiety and depression (RCADS)	1476	19.36 (18.71, 19.99)	1534	19.33 (18.65, 20.01)	0.001 (-0.07, 0.073)	0.983
Outcome	**n**	**Percentage (95% CI)**	**n**	**Percentage (95% CI)**	**ES (95% CI)**	**P**
Teacher rated internalising symptoms (SDQ)[18]	1476	10.37 (8.81, 11.93)	1534	9.67 (8.19, 11.15)	-0.032 (-0.251, 0.187)	0.773
Teacher rated externalising difficulties (SDQ)	1476	6.94 (5.64, 8.24)	1534	7.32 (6.02, 8.62)	0.018 (-0.218, 0.254)	0.878

H1: Children in classes implementing FRIENDS demonstrate measureable improvements in Key Stage 2 Maths and Reading combined scores (equal weighting) when compared to those children attending comparison classes. Effects outlined in H1 are also present for; Maths scores only (H1a) and Reading scores only (H1b).

As can be seen from the above table, the overall interpretation is that FRIENDS did not impact upon pupils' attainment. This finding is consistent for both subject areas (maths (H1a) & reading (H1b)). Probability is firmly non-significant (p= .83 - .95) and effect sizes are far below what might be considered 'practically significant' (Hill, Bloom, Black & Lipsey, 2008).

Sensitivity analysis (appendix J): Statistical models using multiple imputation to account for missing data support the findings of the complete case analysis, i.e. findings are not sensitive to missingness. All differences in the coefficient between complete case and the MI is to the second decimal place, a difference of effect far below what might be considered practically significant.

H2: Children in classes implementing FRIENDS for the duration of the programme demonstrate measureable decreases in their self-rated worry when compared to those children attending comparison

[18] SDQ scores had a J-shaped distribution, which is impossible to transform into a normally-distributed variable. The only viable alternative to model this is to dichotomise the original scores. Therefore, the outcome for internalising and externalising scores is the binary indicator "raised" (1= raised, 0=not raised). These were analysed using a logistic multilevel model. Odds ratios are converted to effect size estimates (ES) following the procedure in Chinn (2000).

classes. Effects outlined in H2, are also present for Self-rated Total anxiety and depression score (H2a); Teacher rated emotional symptoms (H2b) and Teacher rated difficulties conduct problems (H2c)

Regarding health outcomes, there is weak evidence to suggest a small harmful effect of the FRIENDS intervention. For the health related outcomes (PSWQ (H2a) & RCADS (H2b)) a decrease in difficulties is recognised by a reduction in score (i.e. a negative value). Although the direction of effect is consistent with 'treatment' for teacher rated internalising difficulties, as the direction of effect for externalising difficulties indicate a positive relationship, there is a slight worsening of reported symptoms in those allocated to FRIENDS. However, for all health related outcomes, probability is firmly non-significant and effect sizes are far below what might be considered 'practically significant' (Hill et al., 2008), analogous to a little to no practical difference between trial arms, hence, no effects.

Sensitivity analysis (appendix J): As with Hypotheses H1 & H2, statistical models using multiple imputation to account for missing data support the findings of the complete case analysis, i.e. findings are not sensitive to missingness. All differences in the coefficient between complete case and MI is to the second decimal place, a difference of effect far below what might be considered practically significant.

The findings of the complete case analyses were borne out in all of these models; put another way, our findings were not sensitive to missing data.

Table 10: Summaries of sub-group effects for H3a & H3b.

Outcome	Intervention group		Usual provision group			
	n	Mean (95% CI)	n	Mean (95% CI)	Hedges *g* (95% CI)	*P*
H3a: Interaction effect of intervention – pupils eligible for free school meals						
Academic outcomes						
Maths and reading combined score	371	100.63 (99.79, 101.48)	430	101.03 (100.25, 101.81)	0.05 (-0.058, 0.159)	0.587
Maths scores only	371	100.64 (99.78, 101.51)	430	100.8 (99.99, 101.6)	0.057 (-0.052, 0.165)	0.544
Reading scores only	371	100.63 (99.67, 101.58)	430	101.26 (100.4, 102.11)	-0.018 (-0.126, 0.091)	0.851
Health outcomes						
Self-rated worry (PSWQ)[19]	371	0.11 (0.003, 0.23)	430	-0.027 (-0.135, 0.081)	0.148 (0.039, 0.256)	0.109
Self-rated total anxiety and depression (RCADS)[20]	371	0.18 (0.061, 0.301)	430	0.008 (-0.103, 0.12)	0.245 (0.136, 0.354)	0.009

[19] These variables are scored such that higher scores reflect larger iatrogenic (adverse) effects. Thus, a negative effect size suggests favorable effects.
[20] These variables are scored such that higher scores reflect larger iatrogenic (adverse) effects. Thus, a negative effect size suggests favorable effects.

Outcome	n	Percentage (95% CI)	n	Percentage (95% CI)	ES (95% CI)	*P*
Teacher rated internalising symptoms (SDQ)	371	17.39 (13.25, 21.53)	430	20.21 (16.18, 24.24)	-0.087 (-0.399, 0.226)	0.587
Teacher rated externalising difficulties (SDQ)	371	21.12 (16.66, 25.58)	430	19.95 (15.93, 23.96)	0.069 (-0.242, 0.379)	0.665
H3b: Interaction effect of intervention – 'at risk'						
Academic outcomes						
Maths and reading combined score	142	100.88 (99.46, 102.29)	144	101.11 (99.74, 102.48)	-0.004 (-0.172, 0.165)	0.979
Maths scores only	142	100.38 (98.92, 101.84)	144	100.36 (98.98, 101.75)	0.089 (-0.08, 0.258)	0.522
Reading scores only	142	101.37 (99.8, 102.95)	144	101.86 (100.26, 103.46)	-0.003 (-0.171, 0.166)	0.985
Health outcomes						
Self-rated worry (PSWQ)	142	0.45 (0.26, 0.64)	144	0.46 (0.23, 0.68)	-0.09 (-0.259, 0.078)	0.531
Self-rated total anxiety and depression (RCADS)	142	0.61 (0.41, 0.81)	144	0.42 (0.22, 0.63)	0.192 (0.056, 0.327)	0.181
Outcome	**n**	**Percentage (95% CI)**	**n**	**Percentage (95% CI)**	**ES (95% CI)**	***P***
Teacher rated internalising symptoms (SDQ)	142	33.9 (25.36, 43.44)	144	46.55 (37.47, 55.63)	-0.216 (-0.578, 0.146)	0.241
Teacher rated externalising difficulties (SDQ)	142	18.64 (11.62, 25.67)	144	24.14 (16.35, 31.93)	-0.229 (-0.633, 0.174)	0.266

H3a: For the outcome variables listed in H1 (a-b); There will be a significant interaction between allocation to condition and Free school meal status (EverFSM)

There was no discernible impact of FRIENDS for those pupils identified as eligible for free school meals for any of the academic outcomes, including combined score (ES = 0.05, CI = -0.058 to 0.159), maths scores (ES = 0.057, CI = -0.052 to 0.165) or reading scores (ES = -0.018, CI = 0.126 to 0.091).

For health-related outcomes, no discernible effect was evident for self-rated worry (ES = 0.148, CI = 0.039 to 0.256). However, an iatrogenic effect was seen for pupils' self-rated anxiety and depression scores, showing a small but significant increase at post-test as a result of being allocated to FRIENDS (ES = 0.245, CI = 0.36 to 0.354).

For the health related outcomes (PSWQ (H2a) & RCADS (H2b), allocation to FRIENDS did not appear to impact the likelihood of a teacher identifying a pupil with internalising symptoms (ES = -.087, CI = (-0.399 to 0.226) or externalising difficulties (ES = 0.069, CI = (-0.242 to 0.379). The pattern of findings shown for H2 indicating weak evidence to suggest a small harmful effect of the FRIENDS intervention - specifically a very small (and non-significant) decrease in reported internalising symptoms but a very small (and non-significant) increase in reported externalising difficulties are repeated for FSM eligible pupils.

H3b: For the outcome variables listed in H2 (a-c); There will be a significant interaction between allocation to condition and pupils identified with elevated internalising symptoms (scoring in the 'slightly raised' or above category for teacher-rated emotional symptoms at baseline ("at risk").

There was no discernible impact of FRIENDS for those pupils identified as 'at risk' for any of the academic outcomes, including combined scores (ES = -0.004, CI = -0.172 to 0.165), maths scores (ES = 0.897, CI = -0.08 to 0.268) or reading scores (ES = -0.003, CI = -0.171 to 0.166).

For health related outcomes, there was no discernible effect for any of the outcomes, including the primary outcome of self-rated worry (ES = -0.09, CI = -0.259 to 0.078) and self-rated anxiety and depression (ES = 0.192, CI = 0.056 to 0.327). Allocation to FRIENDS did also do appear to impact the likelihood of a teacher identifying a pupil with internalising symptoms (ES = -0.216, CI = -0.578 to 0.146) or externalising difficulties (ES = -0.229, CI = -0.633 to 0.174), though confound effects are possible given that is also the same instrument used to identify 'at risk' groups.

Costs

The costs of implementing FRIENDS are relatively stable over a three-year period given the external delivery model and set number of sessions (10 sessions), but discount is offered for double (or quad) entry schools, on the basis that two sessions (e.g. morning and afternoon) can be delivered within the same day.

For a single class, the cost of delivering FRIENDS is **£1,727.73** a year, which equates to approximately **£63.99 per pupil per year**[21].

For two classes within the same school (e.g. double of quad form entry), the cost of delivery FRIENDS is **£2,305.26**, which equates to approximately **42.69 per pupil per year**.

(i) delivery of the full 10 sessions of the FRIENDS curriculum by a trained and licensed external implementer (in the case of the current trial; Salus), (ii) a group leaders' manual for delivering the session, and (iii) an accompanying activity workbook for each child (£8.43 per child).

Table 18. Costs of delivering FRIENDS

	Cost for single class	Cost for two classes
Cost per school, per annum	£1,727.73	£2,305.26
Cost per pupil, per annum	£63.99	££42.69
Cost per school over 3 years	£5,183.19	£6,915.78
Cost per pupil over 3 years	£191.97	£128.07

These costs include (i) delivery of the full 10 sessions of the FRIENDS curriculum by a trained and licensed external implementer (in the case of the current trial; Salus), (ii) a group leaders' manual for delivering the session, and (iii) an accompanying activity workbook for each child (£8.43 per child). As there is little in the way of initial set up or training costs to be borne by the school this presents a stable cost over a three-year period. For instance, beyond an initial orientation meeting between the implementers and school staff, the external delivery model means there is no need to train school staff or maintain resources (beyond the price-per-pupil workbooks, already covered in the above estimate).

[21] Assuming an average class size of 27 pupils per class (DfE, 2017).

Implementation and Process evaluation

This section examines wider context surrounding the implementation of FRIENDS. First, this chapter begins with an examination of any changes of practice within trial schools, with specific focuses on examining a) evidence for displacement (in wider school practice) and, b) either compensatory rivalry and/or resentful demoralisation effects (in usual practice classes). Second, observation data taken from case study schools is used to examine implementation factors in relation to pupil outcomes, specifically: fidelity and adherence to protocol, quality of delivery, participant engagement and responsiveness and reach. This data is compared and contrasted with interview data taken from the same schools. Third, implementer characteristics are also considered in relation to their association with pupil outcomes. Finally, wider aspects of implementation and process are considered in relation to wider factors affecting implementation, including a consideration of perceived needs and usefulness (Social validity).

Usual practice activity

This section reports on findings generated from the 'usual practice survey', a single self-reported form that recorded the 'state of play' in relation to the current implementation status of similar initiatives (see appendix F). The survey was administered to all schools at pre-test (94.9% response rate) and post-test (96.2% response rate). The data is supported with interview data from case study schools.

As the Usual Practice Survey contains data at school, rather than class-level, data is organised by school type: i) 'practice as usual' - (1-form entry schools allocated to the usual practice arm), ii) 'FRIENDS' (1-form entry schools allocated to the FRIENDS arm), and iii) 'Mixed' (multi-form entry schools with both usual practice and FRIENDS classes). Numbers and percentages for each group are shown below in Table 11.

Table 11: FRIENDS group of schools – Usual Practice Survey groups of schools

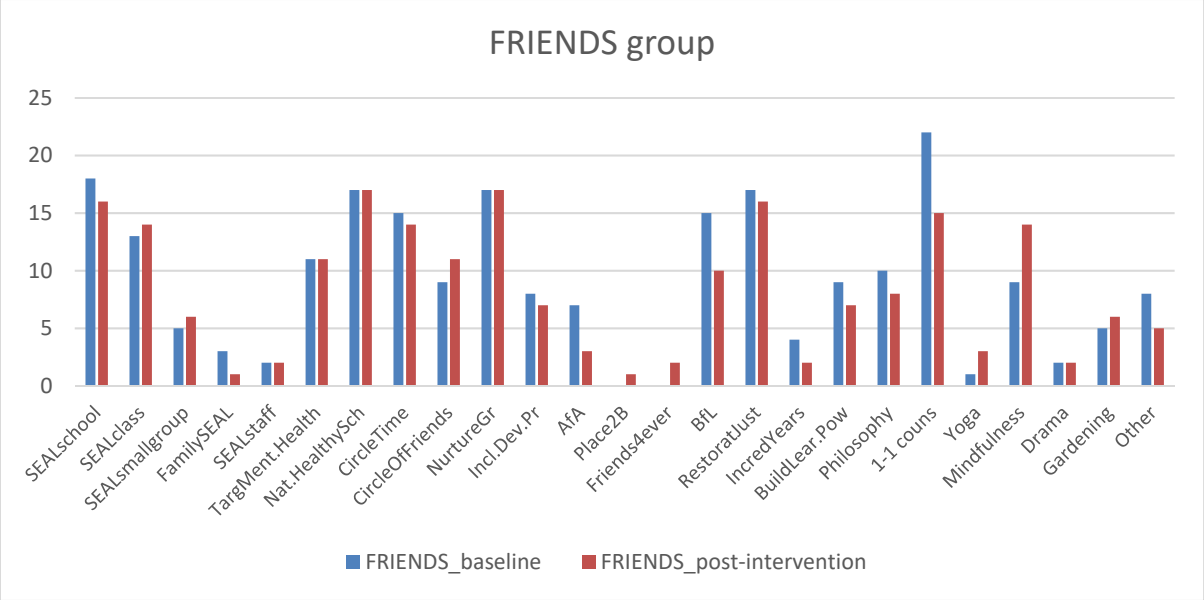
Usual Practice Survey schools	Number of schools	Percentage
Control Schools (1-form entry allocated to usual practice)	24	30.4
FRIENDS (1-form entry allocated to intervention)	25	31.6
Mixed group (multi-form entry with both 'control' and FRIENDS classes)	30	38.0
Total	79	100.0

Displacement

We explored possible displacement effects associated with implementing FRIENDS by comparing changes in the use of named initiatives between pre- and post-intervention within the schools implementing FRIENDS.

A range of universal, indicated and targeted interventions were seen to be present in all schools at pre-test. Figure 2 shows the number of schools reporting using named initiatives at both pre-test and post-test.

Figure 2: Number of FRIENDS schools using each named initiative – Baseline and Post-intervention



In terms of numbers of schools using each initiative, there was little observable change in the reported use of named initiatives between time periods, indicating a lack of evidence for displacement. Interview data indicated FRIENDS was seen to reinforce (rather than replace) existing approaches:

> *We've got a group of boys that come out and do gardening and we do some social skills, self-esteem work in that... [FRIENDS] Just reinforces what we're doing, absolutely… (ID30, SLT)*

> *We had Challenger Troup in who are… they're an army based, sort of team building group really. There's a lot on positive behaviour and positive team things and positive role models really for the children and working collaboratively and, that that really had a had a big impact ... (ID19, CT)*

Pupils also provided occasional associations between FRIENDS and other well-being/mental health *initiatives:*

> *I'm having Lego therapy as well to try and help me with the group and that's given me help extra help (ID61, Pupil3)*

In one school, despite describing FRIENDS as part of a complimentary package of work the school implemented around wellbeing, some staff recognised that the programme was offering different elements compared to previously used approaches:

> *You know, even the breathing I don't think some of them had ever done that before, some of them had never thought about planning steps to reach a goal or to solve a problem or there's different ways of dealing with problems with friends like that* (SLT interview, ID30)

Reported increases in Mindfulness as an approach (from 9 in pre-intervention to 14 in post-intervention) is also consistent with the idea of 'complimentary strategy use', as there are parallels in the theoretical underpinnings of CBT and mindfulness.

However, there was reported reduction in the number of schools implementing the universal framework 'Behaviour for Learning' (from 15 pre-intervention to 10 post-intervention) and 1-to-1 counselling (from 22 pre-intervention to 15 post-intervention). Data from case study schools indicates that discontinuation may be for a number of reasons, including perceived overlap with existing strategies:

> *So compared to the FRIENDS, to be honest a lot of the work that's in FRIENDS is very similar to the kind of work I would do for the interventions anyway ... so a lot of the relaxation techniques, building self-confidence, what happens to our bodies, the feelings, that kind of thing is all very similar really ... (PO7)*

And also potential overlap with other previous initiatives offered by the implementer:

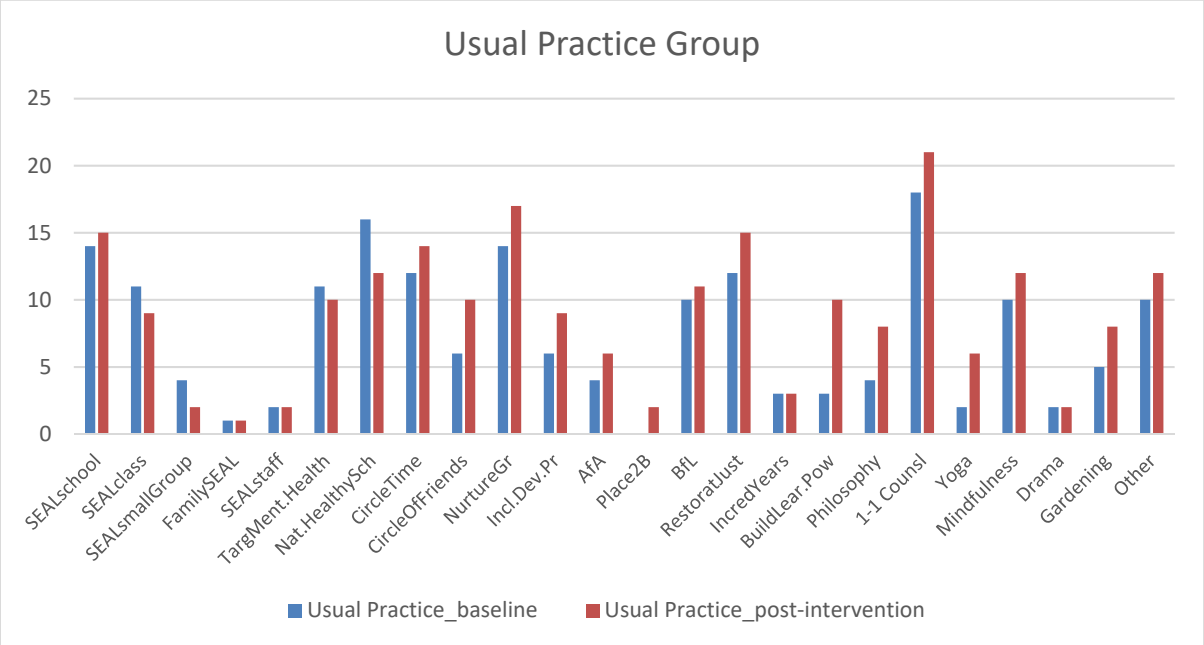
> *one of the reasons we've decided to implement FRIENDS is because … we've worked with project Salus in the past and it was a very successful, so when we got offered the chance and sort of adapt our PSHE programme and we have had some kind of friendship issues linked to resilience in our previous Year Six and started to bubble in Year Five ... (ID14, CT)*

In summary, intervention approaches were seen to remain relatively stable during the trial period (indicating a lack of displacement) with interview data supporting a complimentary approach in relation to existing strategies. Use of 'Behaviour for Learning' and '1-to-1 counselling' was seen to reduce, but data does not indicate that this was directly attributable to the implementation of FRIENDS.

Compensatory rivalry

Change in the use of named mental health and emotional learning intervention practices between pre- and post-intervention in the usual practice group was explored in order to investigate any attempts of a compensation activity. We did not find compelling evidence of compensatory rivalry.

Figure 3: Number of Control schools using each named initiative – Baseline and Post-intervention



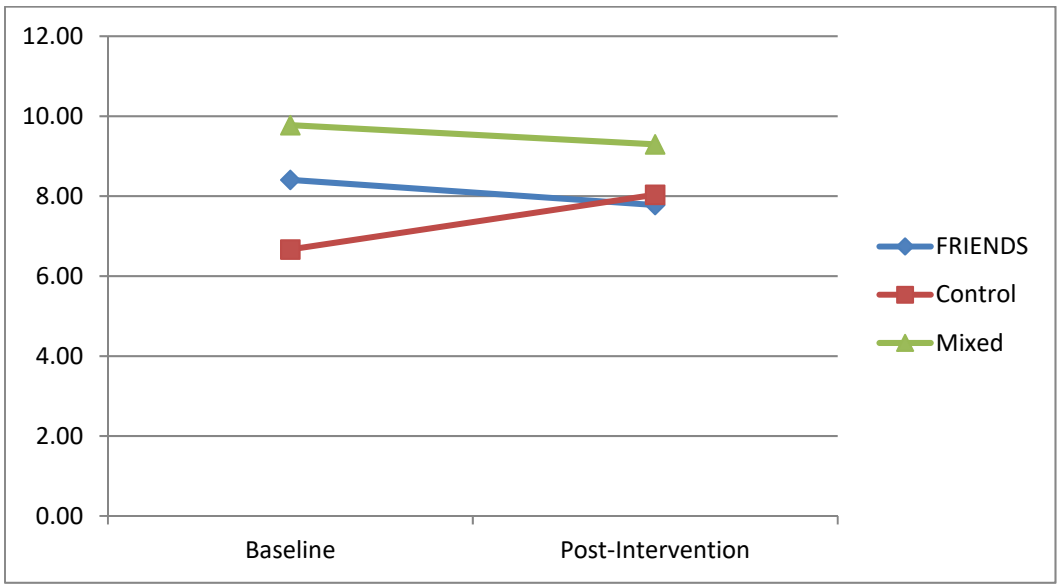
One-to-one counselling was the most commonly used initiative in the total sample of schools reported in the post-intervention survey followed by 'SEAL whole school resources' and the National Healthy Schools Programme. There was an increase of the number of schools reporting use of Nurture Groups and Mindfulness.

In order to assess compensation rivalry specifically, changes in the number of emotional health activities was compared across school type (FRIENDS only, mixed classes, usual practice only). Figure 4 below shows the mean number of named emotional health activities used in schools at pre- and post-intervention (T1 & T2) for each type of school. The drop of the mean number of activities used in FRIENDS schools between T1 and T2 was not seen to be statistically significant (t-paired test of differences, p>.05). In

interpreting the values in context, control schools gained approximately 1.3 points in emotional health activity adoption at post-test – this is the approximately the equivalent of schools 'just getting started' with a new initiative. This is comparatively very little progress, spread across different initiatives (as in conjunction with figure 3, there is no clear 'competitor' programme identified in control schools) in relation the FRIENDS group, who each received a full delivery of the programme. This evidence suggests a minor move towards compensation rivalry, but not to the extent to which there is alternative threat to the internal validity of the study

Alternatively, the minor uptake in new initiatives within control schools reflects the current practices of schools, and that those allocated to FRIENDs reduced their uptake of new initiatives as a result of this allocation. This is not compensatory rivalry, but instead reflects a potential line of enquiry as to the unintended consequences of the adoption of trial architecture within a school's ecosystem. For either scenario, the effect is small.

Figure 4: Mean number of emotional health activities used in all schools across groups of schools in Baseline and Post-Intervention



Case study data

In this trial, FRIENDS was implemented as a universal manualised cognitive-behavioural programme that aims to improve social and emotional skills, build resilience and prevent childhood anxiety. As such, the programme consists of 10 weekly sessions followed by two booster sessions, delivered in the following term. Although FRIENDS is a manualised intervention, variability in implementation can be affected by a variety of factors (e.g. characteristics of deliverer of the programme, background characteristics of the recipient, time availability, etc.) and can appear as in different dimensions (quality, dosage, etc.). Examining those aspects of implementation and process evaluation is key in understanding possibly variability in outcomes (Humphrey, 2013; Lendrum & Humphrey, 2012) This section therefore looks closely at the various aspects of implementation in order to provide an understanding of where variability in implementation was found, why variability existed, how this could affect outcomes and finally provide some recommendations based on this analysis (see relevant section *Evaluation Objectives* – hypotheses H4-H6).

Case study schools

Data regarding implementation variability was taken from 10 case study schools. Selection of the case studies took place in agreement between the University of Manchester (evaluation team) and Salus (implementation team). Maximum variation sampling was used to approach schools within the wider trail cohort on the basis of: Project officer (the title given to those employed by Salus to implement FRIENDS), size of school, % of pupils receiving FSM, % of pupils with EAL, % of pupils enrolled on the SEN register and absence.

The characteristics of the final sample are shown in Table 12. Most of the schools were one-form entry and had below average pupils with EAL and with SEND. There was a nearly even split between schools with percentage of pupils eligible for FSM above and below the national average.

Table 12: Case study school sample characteristics

School	Number of Y5 classes in school	FSM[1]	EAL[2]	Absence[3,5]	SEND[4]	L4+ Eng/maths[5, 6]
A	single form	above average	below average	+	below average	-
B	single form	below average	below average	-	below average	+++
C	single form	above average	below average	+	below average	--
D	single form	above average	above average	-	below average	++
E	single form	above average	below average	-	above average	+++
F	single form	below average	below average	+	below average	++
G	double form	average	below average	-	below average	+
H	single form	above average	below average	++	below average	--
I	triple form	below average	below average	-	below average	++
J	single form	above average	below average	-	below average	++

[1] *As compared to national average in 2016. National average in 2016 = 14.5% in primary schools*
[2] *As compared to national average in 2016. National average in 2016 = 20.1% in primary schools*
[3] *As compared to national average in 2015/16. National average in 2015/16 = 4.0% in primary schools*
[4] *Department for Education (2014), (*As in 2014 – State funded primary schools average 16.6)
[5] key: '+' up to 5% above national average; '++' 5%-10% above; '+++' > 10% above; '-' up to 5% below; etc.
[6] *As compared to national average in 2015. National average in 2015 = 80%*

Table 13: Summary of qualitative data collected and analysed from the Case Study schools

School	Observation of a session	Focus group with pupils	Interview with PO	Interview with school staff
A	√	√	√	Teaching Assistant in Y5
B	√		√	Class teacher Y5 Member from Senior Leadership Team
C	√	√	√	Class teacher Y5 SENCO
D	√	√	√	
E	√	√	√	Member from Senior Leadership Team
F	√	√	√	Class teacher Y5
G	√	√	√	
H	√	√	√	Class teacher Y5 Learning Mentor
I	√	√	√	Member from Senior Leadership Team Class teacher Y5 - 1 Class teacher Y5 - 2
J	√	√	√	Class teacher Y5 Member from Senior Leadership Team

Implementation variability and its association with pupil outcomes

Observation data from trained researchers (see methods) was used to create summaries of implementation variability, which was then modelled against pupil outcomes. Given that there are no agreed thresholds for implementation ratings (for example, it is not possible to definitively say what counts as an 'acceptable' level of fidelity/adherence), our approach was to identify and significant predictors of student outcomes and a relative comparison of effect size between the implementation variables. Table 14 shows the mean and associated 95% CIs for each implementation variable as well as the standardised coefficient and associated significance and effect size in predicting academic outcomes (KS2 maths and reading scores). The definition of these dimensions can be checked in Appendix G.

Table 14: Implementation variability associated with KS2 maths and reading combined scores

Implementation dimension	N (pupils)	Mean (95% CI)	Coefficient (Std. error)	Hedges g (95% CI)	P
Fidelity	1460	8.03 (6.68, 9.38)	0.007 (0.035)	0.022 (-0.029, 0.074)	0.850
Quality	1460	9.52 (9.11, 9.31)	-0.042 (0.135)	-0.138 (-0.190, -0.087)	0.759
Engagement	1460	8.65 (7.92, 9.38)	-0.007 (0.064)	-0.022 (-0.073, 0.0229)	0.917
Reach	1460	9.50 (8.81, 10.20)	0.007 (0.048)	0.023 (-0.029, 0.074)	0.886

Hypothesis 4: Variation in implementation fidelity (H4a), quality (H4b), engagement (Hbc), programme and reach (H4d) will moderate education-related outcomes (H1) in schools implementing FRIENDS.

No dimension of implementation variability was seen to be associated with pupil academic outcomes. However, this result is consistent with academic attainment operating as a potential later outcome of the intervention effect. As proximal changes appear non-significant (i.e. self-rated worry), there is little opportunity for then distal effects to show significance.

Table 15: Implementation variability associated with self-rated worry scores

Implementation dimension	N (pupils)	Mean (95% CI)	Coefficient (Std. error)	Hedges g (95% CI)	P
Fidelity	1460	8.03 (6.68, 9.38)	-0.024 (0.023)	-0.049 (-0.101, 0.002)	0.353
Quality	1460	9.52 (9.11, 9.31)	-0.226 (0.096)	-0.469 (-0.523, -0.415)	0.019
Engagement	1460	8.65 (7.92, 9.38)	0.114 (0.047)	0.237 (0.185, 0.289)	0.015
Reach	1460	9.50 (8.81, 10.20)	0.008 (0.034)	0.017 (-0.034, 0.069)	0.805

Hypothesis 5: Variation in implementation fidelity (H5a-d) will moderate health related outcomes (H2) in schools implementing FRIENDS

Quality of delivery by the Project Officers was seen to be associated with a statistically significant reduction in pupil's self-rated worry (ES = -0.469, CI = -0.523 to -0.415) A smaller but also significant effect was seen for level of engagement by Project Officer, however this effect is associated with a small increase in self rated worry score (ES = 0.237, CI = 0.185 to 0.289).

We now examine a summary of the qualitative commentary in respect to implementation variability.

Programme Delivery

Self-report information from Project Officers facilitated additional insight into one particular element of implementation variability: fidelity / adaptation (i.e. the extent to which FRIENDs was delivered as prescribed in the programme manual). Delivery of the intervention in the treatment groups as specified was seen to be high, although inevitably, there was evidence of variation and adaptation, as POs reported it difficult to implement FRIENDS with absolute fidelity. Evidence indicated that variation was attributable to both PO and nature of the particular session, as shown in figure 5 below.

Figure 5: Delivery map of session content by PO

Fidelity map: Each cell represents the % compliance with the prescribed activities in the FRIENDS group leader's handbook, for each of the 10 weekly sessions. Where POs indicated they had adapted or changed a portion of the session (i.e. an activity), this is represented by a corresponding reduction in % compliance.

Figure 5 shows a pattern of adaptation by PO (for instance, POs 7, 8 & 9 adapted a higher proportion of content for sessions 3, 3 & 5 in comparison to other POs) and by session (sessions 8,9 & 10 were adapted by almost all POs). Although not directly associated to an overall effect upon pupil level outcomes (see

table 15), a more granular interpretation suggests the degrees of variation (i.e. adaptation) may be worthy of further investigation. We now interrogate some of the reported adaptations below.

Adaptations to FRIENDS prescriptive programme refers to the degree and nature of changes that POs made to the sessions in the day-to-day delivery of the programme. We have categorised into the following themes: deviation and omission, failure to engage and additional provision.

Deviation and omission: Forms of deviation and omission in which POs did not follow the prescribed activities and/ truncated the overall session plan were observed by just under half of POs. One PO described choosing different activities for different classes:

> "*Activities in one session you might do one, three and five for [School A] because you know that it will have a great impact on those children and then when you go to [School B] you might go 'oh one, two and four are going to work better with them'*" (PO6 interview)

There was also evidence that this adaptation at the planning stage was used to consider different children's needs:

> *"Because the book's so big I'm I will write out a much shorter plan which is an A4 sheet, which I will then take into the school that I'm in, and adapt as necessary, which can range from we haven't got enough space in a classroom to do an activity to there's a child that can't read properly so I have to do it pictorially, but largely speaking it's planning an adaptation of the programme, obviously from the book, and delivering it in schools and then afterwards evaluating how it's gone."* (PO2 interview)

This approach, in combination with evidence from around half of POs that activities were omitted as they were felt to be "very repetitive" (PO1 and 2) suggests that designed reinforcement of ideas and techniques may have been missed. One tactic used by POs to address time issues, (and possibly also an issue around the programme being heavily prescribed), was to discuss activities rather than complete them as directed to in the group leader's manual. However, this potentially limits understanding and application to children's unique circumstances. There was also evidence of the compression of separate concepts from distinct activities in the manual into a single activity or "*kill[ing] two birds with one stone"* (PO2 interview) as well as evidence (PO4, PO5, PO7, PO8) that occasionally POs *'sent home a copy [of a specific activity] for homework to be completed'* (PO4). POs sometimes also decided not to move on in response to the group:

> *"This is a chatty group and they were doing extremely well with the challenging red thoughts and behaviour which I felt was more beneficial to them and would make the activity easier to understand and do. Talked about the stories and picked out some statements they could underline, to do in their own time."* (PO10 POSR5)

Reactive omissions in response to time were also reported frequently by all POs (*see 'factors affecting implementation'*).

Failure to engage: Failure to provide opportunities for children to apply FRIENDS to their own particular circumstances was seen from evidence that around half of POs omitted homework activities. It is not always clear why homework was omitted though one PO reported that this omission was due to books needing to stay at school, whilst another PO mentioned the workload that pupils have for their main curriculum and other after school activities:

> "*Only downside to session five I would say is there was a lot of home activities in that session and as much as the children wanted to do home activities, especially with homework and other clubs out of school often meant that sometimes they were worrying about getting their home activities done …"* (PO7 interview).

Data also demonstrated that POs would often provide scenarios for children to work through rather than taking the time for children to come up with a relevant scenario from their own experience. For instance:

> *"The children found it very difficult to think of a coping step plan to work on for themselves so I asked them to work on being able to hold a spider, I took a plastic spider with me, if they weren't scared of spiders then what they would do to help someone who is. This worked very well, children worked in their pairs."* (PO5)

It is possible that POs found it particularly difficult to challenge children to apply techniques to their own lives, since they did not know these children as well as a teacher.

Additional provision: Examples whereby POs augmented (rather than deviated or omitted) activities were seen to fall broadly into two categories: the addition of resources or props, and the provision of additional context or examples to new concepts. Additional resources included posters, props, PowerPoint visuals and video clips, while additional context was seen in situations such as the following observed by a researcher:

> *"Activity begins with students providing ideas of things that worry children their age and putting these in a "stress bucket". It seems that as the bucket fills up it starts to overflow. All students' ideas are recognised. PO then leads a discussion (as per activity 5) of identifying ways to cope with difficult situations. Students then complete the activity in their workbook."* (PO6 observation)

The researcher felt this additional activity helped children to identify relevant situations. Such an approach may be particularly valuable since it is clear from the negative adaptations above, that other POs often struggled with connecting activities to the wider context and to pupils' lives.

Similarly, in contrast to omitting activities that were perceived to be repetitive, one PO used such a situation to provide extension:

> *"A [activity] 6 was very easy and obvious having just done lots of work on who is in their support team and why, as such I adapted this by getting each child to stand up and say an adjective that describes a good friend and then they wrote them quickly in their books."* (PO2)

In some cases, POs made adaptations based on the acceptability of resources and cultural context suitability but still upheld the core components intended in the manual, for instance:

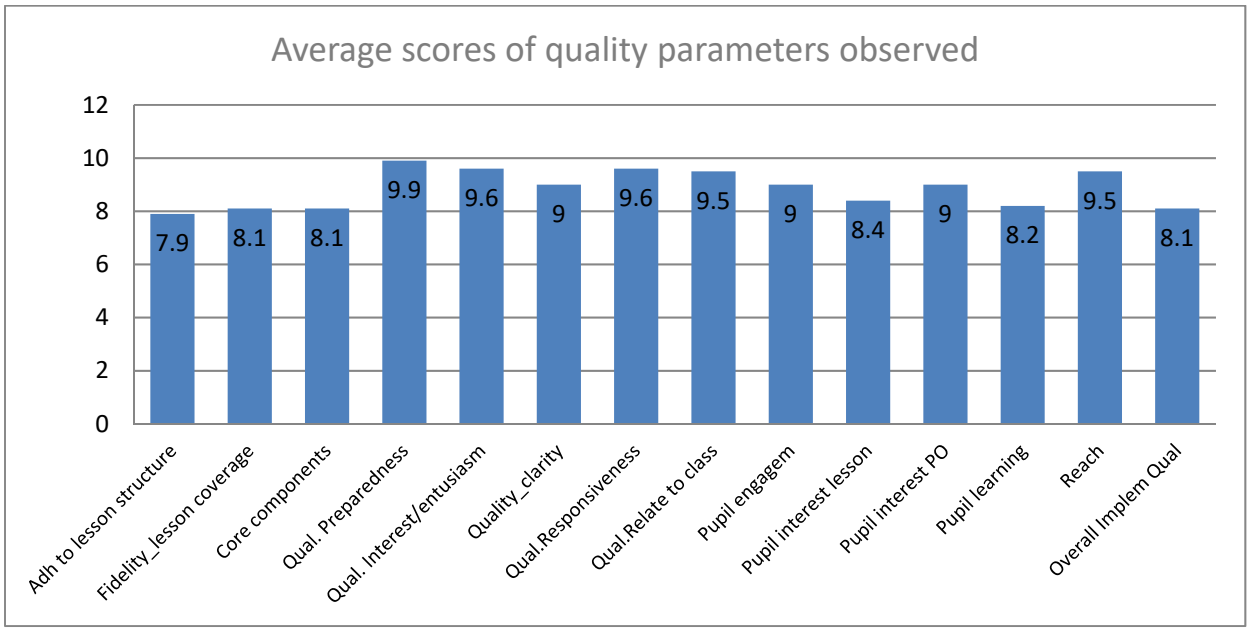
> *"I couldn't relate to this poem and used my story I'd made up of a Year 5 boy named Daniel This story covered the objective of 'coping with worries'. This story was told before Activity 1 to give purpose to Activity 1 and Activity 2."* (PO9 POSR1)

Quality of Delivery

Information on the level to which FRIENDS was delivered by facilitators was mainly gathered through observations during case studies but also through interviews with the other key stakeholders.

Observation data from the case studies showed that researchers generally rated highly the overall quality of delivery, with a total average score of 8.1 (Likert scale in Observation Forms where 1= "extremely poor" to 10= "extremely good") (see figure 6). As the graph below shows, the average scores for the 10 POs, for all aspects of implementation on the days of the observation, were high. The quality of delivery - in terms of preparedness of POs, their interest and enthusiasm in delivering the lessons, the clarity of expressing concepts and activities, the way they responded to each pupil requests and, finally, how they related to the whole class overall, received the higher ratings from the observers, with a minimum of a '9' out of 10.

Figure 6: Mean scores of lessons' observations as rated by the observers



Researchers indicated that overall, POs delivered the intervention with confidence, were well prepared, enthusiastic in delivering the programme and engaged their audience.

The data providing evidence for the quality of implementation is closely linked to adaptations, since good quality was often reported by researchers where high levels of fidelity were also reported and where concepts were made clear through additional questioning, resources or context. Inversely, poor quality was reported where it was felt concepts were not adequately conveyed or presented in enough depth or even omitted. Most POs were described as enthusiastic, fully involved, providing constant encouragement, addressing questions confidently, knowledgeable and appearing to understand the concepts underpinning the activities, able to manage classroom incidents and overall well prepared. Superior quality was noted when POs for instance used activities to further embed programme concepts:

> *"Students form a large circle (as guide), students all have the opportunity to share their Happy Thought. PO reflects on the importance of HTs and how they impact on our feelings"* (PO5 observation)

This kind of additional context was not only presented through planned activities, but also as a reaction to children's disengagement or lack of understanding. However, in several cases, behaviour management was also noted as a barrier to quality in a couple of instances and class teachers made occasional references to effective behaviour management being linked to quality. Researchers and POs, on the other hand, commented on the effective intervention of class teachers and teaching assistants in assisting with the right behaviour management interventions in classroom and therefore, impacting the overall quality of the session. Finally, we extensively describe (*see also Factors Affecting Implementation* further below) the possible effects of the lack of time on the quality of delivery. Several comments also referred to the long duration as after 90 minutes' children may get tired and distracted, losing focus especially if the intervention is the last session of the week. Similarly, the space (occasional change of classroom, very small cramped classrooms) might have prevented some elements to be carried as recommended.

Participant reach

As a universal intervention, there was an expectation that all pupils accessed the programme throughout the 10 weeks by being present in the sessions. There was a general consensus that the intervention

overall reached participants, although in varied ways and levels. For example, there were several comments referring to the positive ways that the intervention particularly reached vulnerable children, the 'quiet ones', those 'needy in terms of attention' or those with additional needs:

> *"The students I've seen the most change in I think are the ones that always get left behind, the ones that are not bad enough to warrant extra help or get those extra lessons ...[…] ... or things like that or the ones that are bright enough to be able to keep up and be the one that everybody the teachers look at and teachers talk to perhaps a little bit more but it's those ones that are just underneath, you know, that almost get forgotten, what I call the forgotten children underneath […] They do their work but nobody ever listens to them particularly well" (PO10)*

> *"I think some of our children who perhaps haven't got an ASD diagnoses, erm, but, erm, display behaviours that are are kind of on the spectrum or that those types of behaviours have I've I I think some of the ways mean that it's another way of explaining about some of the concepts that we're trying to use, erm, in other ways. So when we're talking about, erm, behaviour and what's appropriate and what's not it's just another way of showing.... and discussing that behaviour so it sort of trickles in more."* (ID19 SLT interview).

On the other hand, there is the very occasional comment of pupils being withdrawn from the sessions. As mentioned under Classroom Climate, a few individual children were identified as being 'noisy'

One class teacher questioned whether these children achieved something by just listening in on the background (*see perceptions of impact*) though there was also evidence of such children being excluded altogether:

> *"it just got to the point where I couldn't teach certain children so I had to [exclude them] because the teacher was very non-communicative. Just send the children out for the benefit of the children that were left remaining ... which is a great big shame because those children who were misbehaving were probably the ones who need it the most but if you can't get it across to the majority then, you know ... you've got to get it across as best as you could."* (PO2 interview)

In some instances, these children were considered to need separate intervention by POs or school staff, though this mostly appeared to be unstructured and ad hoc, taking the form of additional conversations with children after sessions. There was one comment by a respondent on whether some children should be completely withdrawn from the programme as it may not reach them and also for the benefits of the rest of the class, or, whether they should 'left' there listening, have small group work or even one-to-one. In one instance, there was clear evidence that teachers and a PO (PO8) worked together to timetable a more structured one-to-one session (PO8).

Participant Responsiveness

Participant responsiveness refers to the extent children in the class engaged with the FRIENDS lessons. The overall response from the pupils is described in very positive terms. For instance, one pupil described the FRIENDS' impact as:

> *"It's helped me from where when something I'm worried about or like something's happening like it's a big change it helps me so I can speak and have confidence to tell my friends because usually I just keep it in myself and I don't say anything."* (ID10 FG)

Teachers, POs (in interviews as well as in self-reports) and researchers (in observations) commented that pupils loved the programme, looked forward to it, felt enthusiastic, excited, attentive and responsive. Pupils in focus groups interviews overall confirmed this general response to the programme. POs thought that in the majority pupils 'embraced it' and they usually 'compete' to participate. Finally, the FRIENDS acronym was seen as engaging for children by two POs though this name was felt to be confusing by some as it did not highlight the programme's focus on resilience.

There were comments that related to variation in the responsiveness between different groups of children, individual children and even between different sessions/topics and elements within each session. Children's different needs and abilities were often seen by POs, teachers and researchers to affect participant responsiveness and therefore implementation in a varied way. For example, the higher ability ones grasped the concepts fully and more quickly than those on the lower end of academic ability who might have found it more difficult to participate at an equal level in all the components of the programme. In some instances, POs felt they needed to "simplify" (PO2 interview) activities or that more differentiation was needed:

> *"they almost understood it to a point, but it was almost like I kind of wanted to kind of split the class in half and be like 'right this half do this and this half do that'"* (PO1 interview)

There was evidence of this opinion and perceived effect of class size on responsiveness from the majority of case study schools and smaller classes were seen as a way of differential implementation leading to programme success and dealing with different needs and abilities.

Implementer characteristics and their association with pupil outcomes.

Implementer (PO) characteristics

The intervention was delivered by 10 project officers (POs), all female, aged between 27-55 years old (mean=40 years). Two POs reported that their highest qualification was an Undergraduate Degree or professional equivalent (e.g. BEd) whilst six POs had a Postgraduate Certificate (eg. PGCE). The remaining two POs reported 'other' highest qualification (one of them a Diploma, the other did not specify). All POs had previous professional educational experience in schools (three out of ten had between 2-5 years of past experience, whilst the remaining seven had 5+ years of experience). Past roles, described as either teaching or pastoral or both, referred to examples such as: teaching SEAL/PSHE and supporting children who had been excluded and/or school refusers, teaching assistant(s) (behavioural TA/HLTA), early intervention project officer, targeted mental/emotional health project officer, pastoral educational support, supporting self-esteem in small/large groups, other peer mentoring and anti-bullying strategies, secondary school teacher and lead in numeracy.

Self-efficacy, Knowledge, Understanding and Approaches to Delivery

Prior to the implementation, POs were asked to self-rate three areas: their emotional self-efficacy, their teaching self-efficacy and their views on social and emotional learning. There were overall high average and total scores reported for both emotional and teaching self-efficacy, however social desirability may have played a role in the high reported levels of self-efficacy. Self-rated scores were then subsequently modelled against pupil outcome data at post-test (see methods).

Tables 16 and 17 shows the results of this analysis for the main academic and health outcomes.

Table 16: Implementer characteristics associated with maths and reading combined score

Implementer characteristic	N (pupils)	Mean (95% CI)	Coefficient (Std. error)	Hedges g (95% CI)	P
Emotional self-efficacy	1460	145.9 (135.5,156.3)	0.007 (0.007)	0.023 (-0.028, 0.075)	0.339
Teaching self-efficacy	1460	99.6 (94.0,105.2)	-0.011 (0.0015)	-0.036 (-0.087, 0.015)	0.459
Views on SEL	1460	52.2 (50.2, 54.2)	0.013 (0.018)	0.043 (0.043, -0.008)	0.478

Table 17: Implementer characteristics associated with self-rated worry score

Implementer characteristic	n(pupils)	Mean (95% CI)	Coefficient (Std. error)	Hedges g (95% CI)	P
Emotional self-efficacy	1460	145.9 (135.5,156.3)	-0.007 (0.007)	-0.015 (-0.066, 0.037)	0.289
Teaching self-efficacy	1460	99.6 (94.0,105.2)	0.010 (0.013)	0.021 (-0.030, 0.072)	0.428
Views on SEL	1460	52.2 (50.2, 54.2)	-0.017 (0.017)	-0.035 (-0.086, 0.017)	0.330

Hypothesis 6: Variation in implementer characteristics, specifically emotional self-efficacy (H6a), teaching efficacy (H6b) and views of social and emotional learning (H6c) are related to pupil outcomes

There are no discernible effects of implementer characteristic on pupil outcomes (attainment / self-rated worry).

Insights into PO's knowledge and understanding of the FRIENDS programme as well as their self-efficacy in its delivery were also drawn from data on POs' own accounts (brief self-report and interviews) as well as from researchers' observations. Analysis of this data provided further insights into how different POs interpreted and defined their role and consequently, into differences in delivery of the intervention. For example, whilst some POs saw themselves as there to deliver the programme, others saw themselves as providing a broader function. For instance, in line with the manual's design, PO7 described that despite a wide range of experience and roles in the county, that "*as far as doing FRIENDS… FRIENDS was an hour and a half session in two schools per day... and that was that was going in, delivering the session from the book ...*" On the other hand, the following description arguably reveals less understanding of the programme's intended delivery method *(see also Fidelity):*

> *"I feel that a project officer is a person that has an understanding of children, …, so they extend on their understanding by being able to come in and deliver a programme that works with children but in a very open manner"* (PO6 interview)

Further to this, whilst some POs described the content and their understanding as "common sense" or "not rocket science" (PO3), others suggested the programme was unique:

> *"I feel this programme is really important for the children to be able to get on in life and I honestly believe that those children that haven't had this programme will find it hard because even getting up on a stage for a child of this age is really hard and difficult and I believe now after doing the programme the children I've taught can do that now and I only wish when I was at school they had a programme like this."* (PO4 Interview)

There was also evidence that the majority of POs felt their previous experience and work beyond FRIENDS influenced their approach and delivery, suggesting that differentiation in terms of the POs' portfolio of work should also be considered in this external delivery model. Evidence of POs sharing experience and practice at weekly meetings and of adapting the programme through supplementation from other interventions can also be seen:

> *"to be honest a lot of the work that's in the FRIENDS is very similar to the kind of work I would do for the interventions anyway .... you know, so a lot of the relaxation techniques, building self-confidence, what happens to our bodies, the feelings, that kind of thing is all very similar really ... to what I normally do."* (PO7 interview)

All these variations in knowledge, understanding and attitudes to social-emotional learning as well as POs' background characteristics must have influenced their actual approaches to delivery of the intervention. Preparation, use of materials and adherence to the programme's manual, amongst others, varied amongst

POs. for example, whilst some POs followed the manual as their 'bible' others found it restrictive (see more below in *Is the Intervention Attractive to Stakeholders?*).

Perceived need and benefit

There was particularly rich data from POs on their perceptions of the programme's perceived purpose (since this was probed more in PO interviews than in other instruments). The programme was viewed mainly as designed to promote resilience (reported by 9 out of 10 POs). POs (9 out of 10) also discussed how the programme worked by equipping children with skills, or described CBT processes such as identifying feelings and behaviours or focusing on the positive. Around half of POs mentioned that the programme should equip pupils to deal with difficult situations or problems as they might arise in the future, but only one PO directly referred to anxiety:

> *"it's a resilience programme and its kind of teaching children how to cope and how to manage their feelings and basically how to kind of get better on in life, … those skills that we're teaching them to hopefully deal with the anxiety and other things they might be faced with."* (PO1 interview)

There were different opinions in regards to which children were more in need of the programme, with two POs claiming that children in deprived areas were most in need, while two others stating the opposite, that children from more affluent areas had less experience of difficulty and were therefore less prepared for it. Only one PO explicitly stated that FRIENDS was inappropriate for children in need of treatment (PO9), although there were a few references to it as a programme serving treatment purposes in a general sense as well as for friendship difficulties and angry or disruptive behaviour.

Summary of factors affecting implementation

Several factors were identified as affecting the implementation of FRIENDS. These are organised by 'level' of intervention: intervention characteristics, school characteristics and classroom characteristics. In addition to those, time was unanimously considered as a barrier to the delivery of FRIENDS and as it cuts across all these levels, it is examined firstly below.

Time

Time was considered a significant barrier to implementation by all POs. POs frequently omitted activities because of lack of time within sessions which therefore affected adherence to the programme on day to day delivery. '

Difficulty in fitting in all session content was consistently reported by POs and children, and some teachers also noted that sessions could feel rushed. Even where POs, as external deliverers, described having some flexibility in session timings from schools (e.g. extended session time), POs still ran out of time and subsequently omitted activities. This difficulty stems from the 'mismatch' between variances in prescribed content (some sessions take longer than an hour) and the fixed school timetable, described explicitly as "*challenges for timetabling [in] schools*" (PO10 interview). However, longer sessions (i.e. beyond 1 hour) were seen to be detrimental in terms of concentration and focus by pupils:

> *"they're not very good at listening for [90 minutes] ..... and I've found that that their attention's started to, I've been in and out a few times... and I know Courtney, our TA's, said that, actually after about an hour usually.... they're starting to they're starting to lose focus. "(*ID19 CT interview).

A member of the SLT in the same school suggested that if the programme were to be continued, sessions would likely be shortened, further curtailing the activities prescribed in the manual. In the main, where POs experienced particular difficulty with sessions and time, these were the sessions that were the longest in length: session 7 and session 9 (both of which take approximately 2 hours to complete). The variety of session lengths in itself was therefore problematic since primary school timetabling did not allow for such

variation, suggesting an expectation that all sessions should be the same length and should fit into an 'standard' school timetable.

Time was also seen to affect quality of delivery (*see also quality*). Time was needed to prepare resources, adapt the programme to the needs of different classes and respond to children. The pattern of POs was divided amongst those who took time to share information and catch up with teachers, compared to those who ran "closed sessions" (*see also External Delivery model and SLT involvement*), and also had an impact on how much time was needed by POs to set up, prepare and debrief once at school. The requirement to hear from each child (e.g. an activity requires ach pup to share their happy thought for the week) was increasingly difficult with class size, especially given the time constraints (*see also participant reach*). One PO described this compromise in some detail:

> *"The only thing that I would say in terms of a challenge was that sometimes you would go to the sessions for whatever reason there may have been something that you needed to either pick back up on from the previous session or recap over for a little bit longer and on a on a couple of times I thought 'ah there's so much to fit in', that you just sort of ran out of time before covering everything which had that sort of domino effect because I wanted to cover everything but obviously you only have so much time in a in a day to be able to do that, or also sometimes the children would want to discuss it a little bit more, not everything but there would be something really poignant that they wanted to talk about and I felt it was important that they had that opportunity so in terms of the difficulty I would only say, sometimes I'd prefer the sessions to have been a little bit longer ..."* (PO3 interview).

Intervention Level characteristics

There were a number of programme level characteristics seen to potentially affect implementation including the external delivery model and, relating to this, partnership between deliverers and recipients. Other characteristics such as content and qualities of the programme (see *Perceived Attractiveness to Stakeholders* section) and programme differentiation are examined in other sections.

External Delivery Model

Broadly, various stakeholders positively described the external delivery model of FRIENDS, with POs receiving explicit praise from school staff or pupils. Positive interactions between the PO and class were noted though the independent observations:

> *"The interactions between the PO and the class are excellent…The PO has obviously formed strong relationships with the students. They seemed excited to see her at the start of the session."* (PO6 observation)

The external delivery of the intervention required high levels of collaboration between schools and POs, working in positive partnership with each other. Overall, communication between schools and POs was mostly seen as effective, with each feeling supported by the other. Several comments highlighted that stakeholders saw such effective working relationships as key to the success of the programme:

> *"I think [clear communication is] really essential, [the PO's] aware and she's been very good as well at making sure that prior to starting that she gleaned all the essential information for the children and she has a a quite a challenging group of children anyway with quite complex needs..."* (ID25 SLT interview)

Stakeholder interviews suggested that there was effective information sharing between POs and school staff. For approximately half of the cases, information was deliberately shared in discussions before and after sessions, while the rest held "closed" sessions at alternative times (ID25 SLT, PO4, PO6). Similarly, some school staff reported being given an overview of sessions (e.g. ID14 CT), while others indicated they

would have liked one. A number of further positive characteristics of external delivery were reported. First, it was noted that by 'booking' an external implementer, it was easier to ensure consistent delivery, as the programme could not be *"push[ed] to one side when you need to get a bit more numeracy in or you need to do a bit more literacy" (*PO5 interview). Second, POs' increasing expertise were noted:

> *"Because we're delivering it a lot, I you get better and better at delivering it. I think if you were just delivering it once a year in your school it would be quite limited. I think you do gain huge amounts by delivering it across the board .."* (PO8 interview)

Third, POs, teachers and pupils often felt that the programme being delivered by external facilitators was a positive aspect as it helped them in being more receptive and open to POs who were viewed as being separate from the rest of children's school experience. Several children, school staff and POs noted that children might have felt able to self-disclose due their relative unfamiliarity with POs, in line with the *stranger on a train* effect (Ignatius and Kokkonen, 2007):

> *"I think because it's someone different I think if we were doing it with a teacher from the school you might kind of think 'cos they're kind 'cos they've known you throughout school they know more things about you whereas if it's someone different then they know nothing about you so you can say really, not really anything, but you can say a lot more than you might actually want to in front of someone who's known you for a lot of your life."* (ID24FG)

> *"... because also I think, especially with children, an outsider there's not that fear of 'oh I can't say that can't say this' because they don't know them... ..., so it's almost a little bit safer ..."* (ID30, CT interview)

> *"I've found a lot of a lot of children just did want someone to talk to but they could felt that they couldn't open up to their teachers, or people in school because they're too busy and they seem to feel like they get overlooked"* (PO2 interview)

Partnership between Deliverers and Recipients of the Intervention

Although the external delivery model was positively viewed by all stakeholders in the few instances where positive working relationships and effective communication between stakeholders were not achieved, delivery appeared to suffer:

> *"the school that didn't know how many children they had in their class ....., I think was very put out that I couldn't begin that session ... so they were quite on the back foot I think for the best part of six sessions"* (PO6 interview)

Similarly, since the school FRIENDS manual is intended to be carried out by school staff (mainly teachers) and promote links beyond sessions, the degree to which the programme was reinforced in other than the dedicated FRIENDS sessions, could also have suffered due to the external delivery model (*see more in external delivery model*). One possible drawback of the external delivery model could be the lack of reinforcement of the programme throughout the school life. There are only a couple of references to school staff prompting children in the playground or displaying strategies, and one teacher recognised that a lack of change in pupils may be due to "*the fact that [he or she hadn't] really followed-up and kept reinforcing it"* (ID86 CT).

School characteristics

School Involvement

Programme acceptability by schools was considered important for FRIENDS implementation, especially as this was delivered externally. However there were less clear indications of optimal levels of schools' involvement. In terms of SLT involvement, in around half of case-study schools, there was clear positive

involvement from head teachers or other senior leaders, while in the remaining cases, their involvement was unclear or limited to early conversations. One staff member linked the involvement of the headteacher to "*very easy*" implementation (ID30 SLT interview), while in another case the headteacher highlighted the importance of the school being involved more and their disappointment of that not being the case:

> *"I'd have liked to see a skeleton over-run of of how what their focuses were going to be over a period of time ... so to have got a more detailed idea and then I can then see the impact of that ..... and possibly then engage the children in very subtle conversations about things and bring things up and I think that would have enabled the school community to then get ... an idea of what these children are experiencing…I think that would have been really really helpful because there is this bit about the, as I said, they go into FRIENDS... and they come out…And actually, you know, the door's closed ..... and it's a bespoke group and actually sometimes it would be nice to have a bit of that filtering through."* (ID25 SLT)

However, in around half of case-study schools there seemed to be a lack of shared decision-making within schools, with senior leaders deciding to implement the programme, while day-to-day decisions relating to running it were made by CTs and TAs.

Competing priorities

In around half of case-study schools there was evidence of competing priorities ranging from problems with finding an appropriate space for the activities that was away from disruption (ID25, PO2, ID28, PO5) to timetabling (*see also time*) and children being withdrawn (*see also participant reach*). A couple of POs reported a "mixed bag" (PO6 and PO8 interviews) while in some cases, the programme was clearly prioritised by the school who flexibly fitted around the timetables of POs. The sometimes less than effective school-deliverer relationships and levels of communication were also perceived as barriers to implementation (*see school involvement*).

Classroom characteristics

Teacher Involvement

Similar to SLT involvement, classroom staff's involvement also varied, and the case-study data suggests a relatively even split between teachers and TAs who chose to attend FRIENDS sessions, and those who did not. In another case, a teacher felt that would like '*for us maybe to be more involved… ... you know, the staff to be more involved*' (CT ID19) but it is worth noting that POs also had views on whether sessions were open or not (*see external delivery model above*). Some teachers who chose to participate reported doing so for reasons such as:

> "*as a classroom teacher, obviously, I feel that I know my children well and I know they've got some needs and I felt for it to be a real benefit and for us to be able to then follow-up with strategies… Otherwise I felt it was a bit 'stand alone'"* (ID86 CT)

However, others felt it was beneficial to let their class participate without a teacher present and some chose instead to catch up with POs at lunch time:

> *"we actually quite like [the teachers not being present for sessions] ... I think the kids quite like that as well it's their special time…They feel mature I think… I sort of say 'right off you go'... 'to to, yeah, your projects and things' ... and it's great."* (ID14 CT interview)

For teachers and TAs who attended sessions, there was a variety of approaches. Some sessions were seen to be disrupted by the inclusion of school staff, whereas in other cases teachers reporting feeling they should observe and keep a low profile. In some teachers were reported to play a key role, for instance by contributing to *"a very nurturing environment"* (PO9 interview) or help with the classroom management. One senior leader also stated:

> *"PO1 is so easy to get on with and, you know, I love what she does so it's been a pleasure for me to be in there because then I can see and support the children and and carry it on"* (ID30 SLT interview)

This varied involvement is likely to have contributed to school staff's varied familiarity with the programme. Around half of school staff from case-study schools did not feel they had good knowledge and understanding of the programme, though a similar number reported familiarity with the resources or demonstrated knowledge of concepts and activities (see also programme characteristics, RQ3). In some cases, this lack of confidence seemed to come from a desire for *"more detailed understanding of the programme itself"* (ID25 SLT) rather than unfamiliarity. This is also suggested by statements such as:

> *"I've been given the week by week overview ..... so I know I do know what they're covering ... yeah but I just don't know how amazingly they're covering."* (ID14 CT interview)

Classroom Climate & Behaviour

Behaviour was broadly reported to be mixed with some good days and others where children were more distracted or, for instance *"A couple of the children were having a difficult morning, (No behavioural assistant today so loss of stabilising influence)"* (PO10 observation). However, certain classes were identified as more consistently "*challenging"* (ID19 SLT) by school staff, researchers and POs and in these cases the groups as a whole were described as *"lively"* (PO4 and PO9 observations) or *"silly"* (PO1 POSR 9). POs and researchers often commented on the important role of CTs and TAs in assisting with the classroom behaviour management. Finally, the levels of literacy of some pupils, or occasionally of the class as a whole, were reported as factors that potentially hinder accessibility to specific activities and therefore to the programme (see also: Participant Reach)

Perceived Attractiveness to stakeholders

As a whole, the intervention appeared to be attractive to the various stakeholders, each of whom have offered their own views in relation to the perceived need of the intervention (see also sections: *Participant Responsiveness, Perceptions of Need and Benefit* and *Perceived Impact*).

In terms of content of the intervention, POs and pupils were mostly very positive about symptom-focused and warm-down (relaxation) activities throughout the 10 weeks. On the other hand, schema-focussed activities (identifying and moderating thoughts and feelings) were seen to be valuable but sometimes challenging aspects of the programme by pupils, CTs and POs. This is not surprising as previous research has reported various levels of satisfaction and acceptability of the various components of the programme (Barrett, Shortt et al., 2001).

In terms of resources, the manual received mixed responses, ranging from one PO describing it as "*my bible*" (PO4 interview) and another stating that it was "*a lot more restrictive than I was hoping for"* (PO2 interview) which consequently affected the ways and degree in which it was used (several described using it as an aid rather than reading directly from it, some preferred to adapt their own prompts from the original, and three POs reported feeling able to run sessions without direct reference to the manual once they had taught a session several times). There were several comments that the tone of resources was not suitable, suggesting resources were "*geared towards… Australian lifestyle*" (PO1 interview), or as needing "*context*" (PO9 interview). However, the majority felt it was useful even if it was a bit *"wordy"* (PO7 and PO2 interviews).

The activity books were positively viewed by pupils, POs, CTs and members of senior leadership and there were several positive comments in relation to games and interaction. Some participants felt, however, that there was too much written work and that more "*active… activities…. would be more useful*" (CT ID19 interview). There was an agreement that children each having their own workbook seemed to contribute to their engagement in the programme.

> "*We're really going to use them' and so many of them said about taking their books home and and just going through the stuff and maybe colouring in the pictures or finishing off anything that wasn't quite finished in the sessions ...*" (PO3 interview)

Examining the attractiveness versus sustainability issue, there was evidence that CTs or SLTs in just under half of case study schools felt it would be worth continuing with the programme. For a couple of school staff, the investment of time and money was a concern in terms of the continuation of FRIENDS, though around half of case-study schools indicated that they might wish to continue with the programme but in an adapted, unsupported format to make savings. For instance:

> "*Invariably as a small school what we tend to do is we take some of the philosophies that underpin it ... and then we make a bespoke programme that's very similar or we can look at costing it in depending on ... finance, erm, but yeah it's been really positive to see*" (ID25SLT interview).

Additional costs were seen in the addition of resources, and the issue of class size (*see participant reach*) demonstrates that more expensive cases, in terms of adult-child ratios, were perceived to be more effective.

Outcomes

Expected outcomes

Though one PO explicitly stated that FRIENDS was inappropriate for children in need of treatment (PO9), a number of POs, pupils and school staff described the programme serving treatment purposes in a general sense as well as for friendship difficulties and angry or disruptive behaviour:

> *"[pupils are] not very good at collaborative group work or team games or anything so we thought it might be quite a nice thing for them to, you know, to have this to see if we can build solid friendships and those relationships."* (ID19 CT)

School staff identified pupils' needs such as poor self-regulation, lack of cohesion, academic stress and low confidence. Their perceptions of the programme's purpose were less around resilience, although there were frequent comments around commitment to wellbeing in more general sense and some also commented on the programme's contribution in preparing children for puberty and school transition. School staff also made more comments, compared to POs, on the possible contribution of the programme to general academic and learning purposes through an understanding that general wellbeing and confidence are necessary conditions for successful learning and "*might help… attitude to learning*" (ID19 SLT) and as it "*just puts them in the right frame of mind for learning*" (ID30 SLT). There were very scarce references to such potential outcomes by all groups of respondents and no benefit for academic stress was discussed:

> *"I'm of the mind that resilience, once you've got that kind of confidence and self-belief, if you don't have it it affects all your lessons and if you do have it it improves everything so really really important for for an hour a week I think that a lot more than that is gained in in the long term"* (ID14 CT interview)

Pupils described the programme's purpose as to aid in dealing with and providing strategies for problems now or in the future. There were single references to each of depression, building friendships and resilience. However, though children sometimes connected using skills to overcoming future difficulty, strategies were expressed in very simple terms:

> *"when we're in a bad time we can always reuse some of the some of the strategies that she has learnt us to do."* (ID25 FG)

> *"how to deal with situations more easily, like being sad or depressed or something like that, you can learn to get happier"* (ID28 FG)

Perceptions of Impact

The main perceived impact relates to emotional or mental health difficulties and was mostly positive with general increases in positive feelings reported by pupils and teachers. POs, CTs and children reported a change in children's self-reported mood, with less "*worry books*" (ID86 CT interview) being used and reduced overall worry. Sometimes this is linked to improved support networks and increased use of these which were commented in over half of case-study schools by all both school staff and POs.

Pupils, SLTs and POs from around half of case-study schools, reported reductions in aggressive, angry or negative behaviour, with particular effects on individuals prone to such incidents, or those who staff were particularly concerned about:

> *"We've seen a little lad who, prior to FRIENDS, we've been monitoring for the last, infinite amount of years, he has significant anger issues, he has low self-esteem, anxiety, it manifests himself in being aggressive and confrontational and staff and I have noticed that there are still elements of those things there ... but actually he seems a lot calmer this term ... and engaging and more positive about himself […] it's quite stark because we were like 'well what's changed?' and the regularity of FRIENDS has been one of the contributing factors we feel ... to the way that he is, you know."* (ID25 SLT interview)

> *"Without [FRIENDS], I would be a completely different person…not completely different, but I would still be throwing things and getting cross…I sometimes do it and now that it happens, I do it a lot less"* (ID19 FG)

Some SLTs felt they had been less involved with high-profile individuals and it is possible that this is linked to a consistently reported change in perspective taking. CTs, SLTs, POs and children in over half of case-study schools observed changes in pupils' ability to see each other's point of view. This related both to learning to recognise how children's own negative behaviour might come across, as well as to be more tolerant of others' challenging conduct:

> *"I think it's been more useful for maybe our children who ... don't necessarily recognise that their behaviour can be seen as being, ..., sort of aggressive or splitting into groups, they think that they're being good friends to somebody else but actually looks like they're ganging up on somebody else ... and I think it's those sorts of children that it's helping them to realise actually what does a good friend look like ... and what how could I behave in this situation?"* (ID19 SLT)

One teacher felt that *"the FRIENDS programme is probably helping them work it out themselves before they come to us."* (ID86 CT), while another reported that children were now *"starting to… put themselves in the position of others ...which is really good."* (ID30 CT). A child also confirmed this:

> *"P2: Before that, children when they got sad at playtime and they couldn't explain what's happening they used to run off saying 'go away' and now they now after it's happened they just sit down and talk about it."* (ID19 FG)

In around of half of case-study schools, POs, CTs and pupils reported improvements in confidence, sometimes for the whole cohort, but mostly for shy individuals who were noted gradually to speak out and share more as the programme went on, as well as for those identified with SEND:

> *"last week we had a staff meeting and we were talking about girls with autism and two of the children in the year group have got autism, two girls particularly, and they've noticed a difference with them being calmer and they're kind of, the way they're describing their emotions, it's helping them." (ID86 SLT* interview*)*

> "*And now, he's actually able to project his voice, read things aloud, present his own work. It's made a huge difference to him*" (ID86 CT interview)

There was evidence of pupils generalising elements of the programme both at home and at school in just under half of case-study schools. At home, most examples related to using symptom-focussed skills in difficult circumstances. For instance:

> *"our next door neighbour we've kind of been having trouble with her and problems so like the police have been round and it's been kind of nerve racking what's been happening so had to do some deep breathing exercises."* (ID61 FG)

There were however, also suggestions in just under half of case-study schools from CTs, POs and pupils that skills were not being generalised, possibly, as one teacher observed, due to the lack of reinforcement of the programme outside the dedicated FRIENDS space.

In around half of case-study schools, academic and learning skills were seen to be affected in the general sense, through children taking difficult scenarios more seriously (such as learning about Anne Frank); increased peer support when struggling; and through a more adaptive approach to learning. This last dimension was most consistently cited, and included children and teachers observing the direct application of programme strategies:

> *"I used something from the FRIENDS programme today because we were doing test and I was using red angry thoughts 'cos I got really stuck on the question I really couldn't get through and they always tell us to leave it but I couldn't leave it I was sort of stuck to it so I was doing the deep breathing and I changed the red thought into a green thought like I can do this ... and I got the question right I think. …Hopefully (laughter)."* (ID86 FG)

> *"I've seen one or two of them 'oh nothing good's happened' 'what did you do with ...?' 'oh yeah, oh yeah, oh yeah' straight away their mind-set has changed…Yes but it just puts them in the right frame of mind for learning because all of a sudden they're positive so yeah….And I, as I said before, the things that's worked out really well that I think is when they have worked together and worked out the scenario themselves..."* (ID30 SLT interview)

Unintended or wider impacts

There were some perceptions relating to wider or unintended impacts noted by a small number of pupils, POs and CTs. The most common of these was that teachers felt their practice was influenced by ideas they want to emulate or continue and a couple of adults on the project felt they had benefited psychologically from the programme. There were also a couple of instances of children teaching skills to family members. For instance:

> *"the eldest sister was starting to panic [about GCSEs] and the little girl was showing, the younger one, Year Five, showing the older one how to do her breathing. …and helping her with that stress ... and she said that they had she could both the parents both parents have noticed a difference at home ... how they were with their siblings."* (PO10 interview)

The only negative impacts reported relate to sadness and worry from pupils, CTs and POs, particularly for those with pre-existing difficulties in symptom-focussed activities in line with other research (Feldman, Gleeson & Senville, 2010). For instance:

> *"P3: I don't like doing the relax the relaxation at the end because we have like when she says find a place where's a happy place and I always find place where is happy but it's worse and it shifts into ... a place where it's not ... and it really I don't really want to be reminded of those places….*
>
> *P2: That's the same with me….Because I had a friend who died….I don't really want to be reminded of it.*

I: Oh I see so even though it's intended to take you to a happy place sometimes it can maybe take you to a not happy place. Yeah?

P5: Ah well it's sometimes been a bit bad but sometimes it's been good because, it's brought up some like quite some painful topics ..." (ID25 FG)"

There were also examples where children with existing difficulty were seen as struggling in sessions *"because… problems were too big"* (PO9 interview). For instance, one teacher described a child *"getting very anxious in the lesson. He was anxious about coping with his anxiety…He was crying"* (ID86 CT interview).

Conclusion

Key conclusions	
1.	The project found no evidence that FRIENDS had a positive impact on children's academic attainment, overall. This result has a high security rating.
2.	Among pupils eligible for Free School Meals, those in the FRIENDS classes made 1 additional months' progress on a combined maths and reading measure compared to children in other classes. These results may have lower security than the overall findings because of the smaller number of pupils.
3.	The project found no evidence that FRIENDS has a positive impact on children's health outcomes, overall. This result has a very high security rating. Pupils eligible for Free School Meals experienced a small increase in self-rated anxiety and depression, although these results may have lower security than the overall findings because of the smaller number of pupils.
4.	Overall, time was found to be the biggest pressure in ensuring consistency and quality of delivery. Schools often struggled to fit the FRIENDS sessions, which varied in length, within the school timetable. This lead to a variability in the amount of time dedicated to FRIENDS.

Interpretation

The current study is, to the best of our knowledge, the largest randomised control trial of FRIENDS conducted worldwide, and is only the second to consider standardised academic attainment as an outcome. It is also the second randomised trial of FRIENDS to take place in England. The study failed to show any intervention effects for any of the study outcomes (both academic and health related) at the ITT level, with a small iatrogenic result shown at the subgroup level for pupils eligible for free school meals. At first glance, these results appear incongruent with the wider literature base supporting FRIENDS as an evidenced intervention, however a number of important distinctions are worthy of consideration in interpreting this outcome.

In relation to the ITT analysis of the academic outcomes (H1), findings are consistent with the emergent literature. For example, Ahlen, Hursti, Tanner, Tokay and Ghaderi (2018) reported no effects on child attainment from a recent cluster randomised control conducted in Sweden; however, it is worth noting that academic performance was assessed through teacher ratings. A more compelling comparison is with that of Skryabina, Taylor and Stallard (2016), whose results show no effect on academic performance 12 months' post-intervention FRIENDS, specifically Key Stage 2 (maths, reading and writing) in English Primary school children. A lack of meaningful findings in this regard is broadly consistent with an emerging literature base, which demonstrates a lack of conclusive evidence for internalising symptoms predicting later academic attainment (Deighton et al., in press; Panayiotou & Humphrey, 2017). Utilising developmental cascade work, conclusions drawn from longitudinal work has implications for predicting the likely magnitude of any academic impact of health-based interventions (especially for universal populations), indicating that if results are to be seen, they are likely to be small. However, it is worth noting that this line of enquiry in its infancy (early health outcomes explaining later educational outcomes).

For H2, which examined ITT effects for health based outcomes, the study did not find any discernible effects for self-rated worry, self-rated anxiety and depression or teacher rated difficulties (emotional symptoms or behavioural difficulties). For studies that have shown positive effects as a result of implementing FRIENDS, these have not always been shown across the full range of outcome measures. Although self-report has been described as potentially contentious as a reliable outcome measure for assessing FRIENDS (Maggin & Johnson, 2014), a number of studies have shown an effect for self-report measures whilst reporting null findings for teacher-based assessments. Therefore, our subsequent considerations focus on self-rated worry as the primary health related outcome, and that most closely aligned to similar measures used in prior studies. In interpreting the lack of significant effect for self-rated

worry, it is crucial to reconsider the intended purpose and approach of the intervention, with reference to intended target(s) and outcomes. As a programme that is delivered universally with a specific focus on negative thoughts and behaviours, it is to be expected that a large proportion of the recipients (i.e. pupils) are without symptoms or immediate need of the intervention itself. This is in contrast to more selective programmes or approaches (which typically might deploy alternative delivery approaches such as small group work), which typically feature selection criteria for individuals at comparatively greater risk of developing difficulties and/or displaying early-onset symptomology. It is a common phenomenon in intervention and prevention literature that targeted interventions typically produce larger effects in comparison to universal approaches (Reivich, Gillham, Chaplin, & Seligman, 2005). This is particularly important as, although FRIENDS is delivered as a whole class approach, evidence for its effectiveness has sometimes been drawn from exclusively high-risk populations (e.g. Pereira, Marques, Russo & Barros, 2014; Liddle & Macmillan, 2010; Bernstein et al., 2005). Comparators for effectiveness drawn from these types of studies (indicating effects between 0.44 - 0.84) are overestimates for universal effects. In calculating a suitably more contextual benchmark, effects drawn from participants with low or absent initial symptomatology are shown to be close to 0.24 (Briesch et al, 2010) or lower, as with Maggin and Johnson (2014) report effects as low as 0.12 at immediate post-test (dissipating at 12 month follow up). Even smaller effects are suggested by Barrett et al., (2017), who consider a range between 0.04 – 0.19 as potentially "successful", given that results are drawn from reduction in symptomology measured against an expectation of prevalent floor effects. For the current study, an intervention effect of g = 0.06 was shown for self-rated worry, which is included within Barrett's range for considering success, but is far lower than average estimates drawn from similar studies. Possible explanations include the presence of a sleeper effect and/or the possibility and an emphasis on treatment vs. prevention. In regards to the former, both Essau and colleagues (2012) (trialling FRIENDS) and Stallard et al., (2013) (trialling a similar CBT based intervention) found effects at 12 month follow up, having initially reported null findings. This is consistent with the idea that, as a prevention programme, skills and behaviours are deployed when required later in the life course, rather than displayed as an immediate benefit at post-test. As the current trial does not allow for any follow-up effects, it is not possible to more conclusively address this possibility. Closely related to this is the idea that FRIENDS is comparatively more efficacious in providing treatment (rather the prevention) allowing children to cope with current difficult and/or anxiety provoking situations in the present study. Accordingly, we now examine sub-group effects in relation to evidence towards treatment effects.

H3 examined the possibility of subgroup effects with specific reference to eligibility for free school meals (H3a) and 'at risk' status (defined by elevated teacher-rated emotional symptoms at baseline) (H3b). There was no evidence of sub-group effects for pupils identified 'at risk' though pupils eligible for free school meals showed a small increase in their self-rated anxiety and depression. In terms of the latter, although socio-economic disadvantage is an established predictor of both poor mental health and impaired academic attainment (Bradley & Corwyn, 2002), the factors behind this relationship are complex. For instance, SES is a proxy of several factors, including access to community resources and exposure to stressors (Hetzner, Johnson & Brooks-Gunn, 2010). There is little direct literature exploring differential treatment effects for this subgroup, especially regarding FRIENDS specifically. Indicated treatment programmes would likely recognise eligibility for free school meals as a potential risk factor, however wider literature indicates the need to consider a number of different variables or scenarios in determining 'risk', for instance gender (Deighton et al, 2017). Lack of discernible effects for 'at risk' pupils appears incongruent with the wider literature, however, in interpreting this outcome, two principle elements require consideration; identification of 'at risk' subgroup and implementation of the programme. Regarding the identification of 'at risk' pupils within the current study, self-report was ultimately rejected in favour of the teacher rated SDQ on the grounds of rigour. Utilising self-report, especially where there is no objectively derived cut-off and the selection criteria for risk is equivalent to the main treatment effect (self-rated worry) runs the risk of over-interpreting data. As several previous studies are not immune to this critique (e.g. (Hunt, Andrews, Sakashita et al., 2009), estimates for sub-groups may be elevated within the FRIENDS literature base. However, even the most conservative of effect size estimates (as above) show the current

study to lack discernible impact. We now turn to a wider discussion around the implementation of FRIENDS.

H4 & H5 showed that quality and engagement were significant predictors of health based outcomes. This indicates that, even for heavily prescribed interventions, the mode and method of delivery by the implementer is important. The wider IPE data suggests the deployment of an external implementer was an important factor in delivery. It was acknowledged that an external delivery model promoted a commitment to the delivery of FRIENDS, ensuring that it was not relegated due to other competing priorities (see below). Even so, time was still recognised as a constraining factor with interview data highlighting how time restrictions intersected with several of the programmes implementation dimensions, such as coverage, adaptations and omission of content/activities, adherence to the programme on day-to-day delivery and, finally, quality of overall implementation. Regarding the external delivery role, there were mixed attitudes from both POS and teachers regarding whether school staff should be present within the sessions. Some participants commented on the fact that continuity between FRIENDS sessions and the rest of the school day was not overt and possibly useful techniques for reducing anxiety could be missed if it was done '*in isolation it's not really going to kind of embed into general practice as well as it might*'. (PO2). In maintaining an external delivery model, there are opportunities to consider how best to integrate school staff into the intended objectives of the programme, however this is not without difficulty. Although there is a general paucity of research discussing the implementation of FRIENDS, it has been noted that one proposed benefit of external delivery is an independent opportunity to discuss worries and concerns (Stallard, 2010), which may be voided should teaching staff be in attendance. However, it was also noted within the current study that some of the adaption to the programme was as a result of POs being unfamiliar with pupil's backgrounds, limiting the opportunities for strategies to be personalised – a core component of the programme.

The above commentary highlights a current tension in the English policy context – how best to develop health and education partnerships within school based settings. In the specific example of the current FRIENDS trial, this is encapsulated by the training and background of the implementers (project officers). Although H6 did not identify any specific implementer characteristics as being directly related to pupil outcomes, self-report questionnaires showed a background drawn exclusively from previous professional educational experience in schools, including roles as teaching assistant(s), pastoral educational support and teaching. Formalised training in relation to FRIENDS was reported as one day's training. This is in contrast with the background of implementers in previous trials of FRIENDS, which have deployed trained therapists (though typically when implementing in 'at risk' populations e.g. Barrett, Sonderegger & Xenos (2003)) and trained psychologists (e.g. master's or above in the clinical sciences) (Lock & Barrett, 2003). Expansive formal training in the areas of mental health are hypothesised to allow a focus on the programme theory, facilitating a greater awareness of mental health issues raised and discussing during implementation of the programme. However, trained health facilitators may not have comparable levels of experience in relation to classroom management and related pedagogical practices. It is broadly accepted that as programmes move from universal to indicated and treatment-based modes the importance of a health background may supersede that of pedagogical expertise (Barrett et al, 2017). This element may explain a lack of subgroup effects, however there is a need for further work in the broader literature to explore this further.

Regarding the wider ecology of the school environment, it is worth noting the possibility of a lack of unique gain of FRIENDS over and above existing provision, as encapsulated in the quote below.

So compared to the FRIENDS yeah, to be honest a lot of the work that's in the FRIENDS is very similar to the kind of work I would do for the interventions anyway ... you know, so a lot of the relaxation techniques, building self-confidence, what happens to our bodies, the feelings, that kind of thing is all very similar really ... (ID28)

The usual practice surveys demonstrated a range of interventions present within each school, though interviews did not elicit any substantial commentary regarding an integrated or co-ordinated strategy for

their use. Integrated or 'blended' strategies by which a range of complementary programmes are selected (e.g. across universal, selected and treatment-based approaches) may provide additive or multiplicative effects through a synergistic approach (Domitrovich et al., 2010). There is little evidence regarding the suitability of FRIENDS as an integrated package, but this may be an approach in addressing concerns around a) the integration of teachers in promoting broader objectives outside of FRIENDS sessions and b) pressures of time in relation to programme delivery.

In light of the various findings discussed above, although the wider evidence base indicates FRIENDS can be effective, both in regards to prevention and treatment, however there needs to be a closer and more specific consideration regarding factors for its success. Key elements for consideration include an increasing need to delineate the evidence base in regards to approaches in assessing the ambiguity between prevention and treatment (both in terms of delivery and evaluation), a more careful examination of the implementer and the training provided, and the role of the teacher regarding the intersection between health and education.

Strengths and Limitations

The current study demonstrates a number of strengths, securing confidence that the principal conclusions are accurate and representative. The trial was large and well powered, with an MDES smaller than demonstrated by most SEL interventions (Wigelsworth et al., 2016) meaning that the smallest of improvements in pupil outcomes are likely to be reliably captured, even for the sub-group of 'at risk' pupils (a limitation of some earlier trials of FRIENDS). Randomisation was conducted independently of the evaluation team, and the use of minimisation variables ensured a good balance between trial arms at baseline. Retention of pupils, classes and schools was extremely good with 0% loss-to-follow up after randomisation at the school level and 12% at the pupil level (for academic measures). Return of health measures (e.g. worry) was even better, at 6%, likely aided by the external delivery model of the intervention. Selection of measures was robust, all having an established psychometric history, drawn from both pupils and teachers, and having been deployed in a number of previous evaluations of FRIENDS. In addition, all analysis and reporting is based on CONSORT standards and EEF guidelines as reported in the statistical analysis plan, although the identification of the 'at risk' subgroup was subject to revision (see below). Our principal findings relating to the impact of FRIENDS on pupil-level outcomes at the ITT level were not sensitive to any changes in our modeling parameters. In terms of generalisability, the composition of trial schools mirrored that of primary schools in England in terms of size, attendance and average attainment (table 7). However, trial schools were seen to be higher in the proportion of pupils eligible for free school meals and showed a much small number of pupils with special education and additional needs or disabilities and those speaking English as an additional language. The study authors are not aware of empirical evidence to suggest that this variation is necessarily directly related to the intervention mechanism (though any effect of this variation may be worthy of future analysis).

A slight revision to the original protocol regarding the identification of the 'at risk' subgroup was made shortly before the main analysis began. Despite an understanding that differential response to intervention is likely on the basis of 'need' (e.g. those presenting with elevated symptoms at baseline), wider literature shows a lack of consistency in the identification of this sub-group. A number of different measures and participant groups have been used in previous trials, including parent, teacher and/or self-reported difficulties, behaviour and/or worry (Hunt et al., 2009; Miller et al., 2011; Iizuka et al., 2014). Although self-reported worry was initially chosen on the basis of its alignment with the primary health outcome of the current trial, the use of self-report raised some questions regarding its suitability in providing a dichotomous cut-off. The PSWQ does not have clinically identifiable categories, and literature indicated difficulties in reliability distinguishing an appropriate cut-point (Viana, Rabian & Beidel, 2008). As pre-test teacher SDQ data was provided at the at the end of the school year and as the instrument has UK-normed clinical cut-offs, this data source was considered a better choice for identifying 'at risk' status (specifically those rated as 'elevated' or above for emotional symptoms) with the additional benefit that the risk status was no longer

selected from the outcome itself ('inherent to treatment'). Although this is an arguable refinement to the final analysis, this is a change in protocol.

In regards to study design, the use of class as the unit of randomisation (rather than school) means that the possibility of contamination between trial arms cannot be completely ruled out. The likelihood of contamination is low given that the delivery of FRIENDS is by external implementer and the process date reports that several teachers were unaware of the content or lessons learnt by pupils. Peer-to-peer contamination remains a possibility, though the nature of the intervention is in teaching internalised skills (as per CBT theory) reducing the opportunities for overt demonstration of skills and replication by the comparison group (e.g. on the playground). Such risk is offset by the substantial advantages in reduced trial and implementation costs by having FRIENDS and usual practice classes on the same site, corresponding security in improved retention and minimisation of potential compensation rivalry (as double, triple and four-form entry schools were in receipt of FRIENDS for at least some classes). In the case of an observed uptake of similar programmes in single-form entry schools that were allocated to control only, differences in change in practice were shown to be non-significant. This indicates that if there is evidence of compensatory practices as a result of being allocated to control, it's effects are minor.

A final consideration is the lack of longitudinal observer data from research staff. Hypotheses 4 and 5 are based on a single 'snap shot' of observation. This presents the possibility that observation data may be subject to researcher/ observer effects (although all project officers came from education backgrounds in which practice was routinely observed). Multiple independent observations would allow mitigation of this potential limitation as well have allowed any temporal variation in implementation to be taken into account in the analyses.

Future research and publications

As the evidence base continues to build in relation to FRIENDS, especially in the English context, there are opportunities to continue to pursue several key lines of enquiry. We anticipate further publications to be generated in regard to the following:

First, there remains an ambiguity as to the relative cost / benefit in regards to alternative delivery modes, specifically in relation to teacher involvement. Interview data revealed a desire for schools to more directly engage with the intervention, therefore there are opportunities to consider the effect of scenarios such as teacher delivery with external technical support (e.g. a coaching model), such as those deployed in several social and emotional learning interventions (e.g. Promoting Alternative Thinking Strategies – PATHS). However, it should be noted that technical assistance was provided in the recent trial conducted by Stallard and colleagues who reported a low up take-up of support by teachers.

Second, the current report serves as a general call to consider alternative means for identifying and analysing sub-groups. There is no established consensus for establishing categorical cut-off's for differential treatment effects, and each method brings with it it's own benefits and disadvantages. For instance, in the case of the current trial, although the SDQ is a relatively well-validated instrument, internalising symptoms may not always be apparent to the external observer. Worry benefits from self-report, but is problematic when also used as a primary outcome. Future work may benefit by examining the utility of alternative person-focused' modeling (e.g. through identification of latent profiles amongst pupils).

Finally, as modes for examining implementation continue to develop in education trials, there is further work in examining more closely the 'critical components' required for effective pupil outcomes. A range of options can be developed (and are scheduled for future publication) to examine this aspect, including a consideration of the underlying logic model and closer examination of any differential effect of different theoretical components (e.g. a focus on expressing emotions vs. self talk).

References

Ahlen, J., Hursti, T., Tanner, L., Tokay, Z. & Ghaderi, A. (2018). Prevention of Anxiety and Depression in Swedish School Children: a Cluster-Randomized Effectiveness Study. *Prevention Science*, *19*(2), 147–158. Doi:10.1007/s11121-017-0821-1

Althouse, A. (2016). Adjust for Multiple Comparisons? It's Not That Simple. *The Annuals of Thoracic Surgery, 101, 1644-1645. Doi: https://doi.org/10.1016/j.athoracsur.2015.11.024*

Barrett, P. (2005). *Friends For Life: Group Leaders' Manual for Children*, Australia: Barrett Research Resources Pty Ltd.

Barrett, P.M., Cooper, M. & Guajardo, J.G. (2014). Using the FRIENDS programs to promote resilience in cross-cultural populations. In Prince-Embury, S., & Saklofske, H. (Eds). *Resilience Interventions for Youth in Diverse Populations*. (85–108). New York: Springer.

Barrett, P. M. & Pahl, K.M. (2006). School-based intervention: Examining a universal approach to anxiety management. *Australian Journal of Guidance and Counselling*, 16, 55−75. http://doi.org/10.1375/ajgc.16.1.55

Barrett, P. M., Cooper, M., Stallard, P., Zeggio, L., Gallegos-Guajardo, J., Barrett, P. M. & Stallard, P. (2017). Effective Evaluation of the FRIENDS Anxiety Prevention Program in School Settings: A Response to Maggin and Johnson. *Education & Treatment of Children*, *40*(1), 97–110. https://doi.org/10.1353/etc.2017.0006

Barrett, P. M., Sonderegger, R., & Xenos, S. (2003). Using FRIENDS to combat anxiety and adjustment problems among young migrants to Australia: A national trial*. Clinical Child Psychology and Psychiatry, 8*, 241–260

Barrett, P. M. & Turner, C. M. (2004). Prevention strategies. In T. L. Morris, & J. S. March (Eds.), Anxiety disorders in children and adolescents (pp. 371−386). 2nd ed. New York: The Guilford Press

Beck, A. (1964). Thinking and depression: Theory and therapy. *Archives of General Psychiatry, 10*, 561-571.

Berkel, C., Mauricio, A. M., Schoenfelder, E., & Sandler, I. N. (2011). Putting the pieces together: An integrated model of program implementation. *Prevention Science, 12*, 23–33. doi: 10.1007/s11121-010-0186-1.

Belfer, M. L. (2008). Child and adolescent mental disorders: the magnitude of the problem across the globe. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 49*, 226–36. Doi:10.1111/j.1469-7610.2007.01855.x

Bernstein, G. A., Layne, A. E., Egan, E. A. & Tennison, D. M. (2005). School-Based Interventions for Anxious Children. *Journal of American Academy of Child and Adolescent Psychiatry*, *29*(6), 1118–1127. http://doi.org/10.1016/j.biotechadv.2011.08.021

Bittner, A, Egger H.L., Erkanli, A, Costello, J.E., Foley, D.L. & Angold, A. (2007). What do childhood anxiety disorders predict? *Journal of Child Psychology and Psychiatry, 48(12),* 1174–83. Doi: 10.1111/j.1469-7610.2007.01812.x

Bradley, R. H., & Corwyn, R. F. (2002). Socio-economic status and child development. *Annual Review of Psychology*, 53, 371–399. https://doi.org/10.1146/annurev.psych.53

Briesch, A. M., Hagermoser Sanetti, L. M., & Briesch, J. M. (2010). Reducing the Prevalence of Anxiety in Children and Adolescents: An Evaluation of the Evidence Base for the FRIENDS for Life Program. *School Mental Health*, *2*(4), 155–165. http://doi.org/10.1007/s12310-010-9042-5

Caan, W., Cassidy, J., Coverdale, G., Ha, M.A., Nicholson, W. & Rao, M. (2014). The value of using schools as community assets for health. *Public Health*, *129*(1), 3–16. Doi: 10.1016/j.puhe.2014.10.006

Carpenter, J.R., Goldstein, H. & Kenward, M.G. M. (2011). REALCOM-IMPUTE Software for Multilevel Multiple Imputation with Mixed Response Types. *Journal of Statistical Software*, *45*(5), 1–14. http://doi.org/http://dx.doi.org/10.18637/jss.v045.i05

Centre for Mental Health. (2010). The Economic and Social Costs of Mental Health Problems in 2009/10. London: Centre for Mental Health.

Cohen, J. (1992). A power primer. Psychological Bulletin, 112, 155-159.

Costello, E.J., Mustillo, S., Erkanli, A., Keeler, G., & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry*, *60*(8), 837–844. Doi: 10.1001/archpsyc.60.8.837

Daviss, WB., Birmaher,, B., Melhem, NA., Axelson, DA., Michaels, SM., & Brent, DA. (2006). Criterion validity of the Mood and Feelings Questionnaire for depressive episodes in clinic and non-clinic subjects. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*. 47*,* 927–34.

Deighton, J., Humphrey, N., Patalay, P., Belsky, J., Wolpert, M., & Vostanis, P. (in press). Longitudinal pathways between mental health difficulties and academi cperformacen during middle childhood and early adolescence. *British Journal of Developmental Psychology*.

Department for Education (2014). 'Children with special educational needs: an analysis'. London: DFE. https://www.gov.uk/government/statistics/children-with-special-educational-needs-an-analysis-2014 [accessed 5/2/2018]

Department for Education (2015). 'National curriculum assessments at Key Stage 2: 2012 to 2013'. London: DFE. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/483897/SFR47_2015_text.pdf

Department for Education (2015b). Phonics screening check and key stage 1 assessments: England 2015. Statistical first release SFR32/2015. Retrieved from: https://www.gov.uk/government/statistics/phonics-screening-check-and-key-stage-1-assessments-england-2015

Department for Education (2015c). National curriculum assessments: key stage 2, 2015 (revised). Statistical first release SFR47/2015. Retrieved from: https://www.gov.uk/government/statistics/national-curriculum-assessments-at-key-stage-2-2015-revised

Department for Education (2015d). Nicky Morgan speaks at Early Intervention Foundation conference. 12 February, 2015. Retrieved from: https://www.gov.uk/government/speeches/nicky-morgan-speaks-at-early-intervention-foundation-conference

Department for Education (2016a). 2016 key stage 2 scaled score conversion tables. Retrieved from: https://www.gov.uk/government/publications/2016-key-stage-2-scaled-score-conversion-tables

Department for Education (2016b). National curriculum assessments: key stage 2, 2016 (revised). Statistical first release SFR62/2016. Retrieved from: https://www.gov.uk/government/statistics/national-curriculum-assessments-key-stage-2-2016-revised

Department for Education (2017a). Schools, pupils and their characteristics: January 2017. Statistical first release SFR 28/2017. Retrieved from: https://www.gov.uk/government/statistics/schools-pupils-and-their-characteristics-january-2017

Department for Education (2017b). Pupil absence in schools in England: autumn term 2016. Statistical first release SFR20/2017. Retrieved from: https://www.gov.uk/government/statistics/pupil-absence-in-schools-in-england-autumn-term-2016

Department for Education (2017c). Special educational needs in England: January 2017. Statistical first release SFR37/2017. Retrieved from: https://www.gov.uk/government/statistics/special-educational-needs-in-england-january-2017

Department for Education (2017d). Class size and education in England evidence report. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/183364/DFE-RR169.pdf

Department for Education & Department of Health. (2015). *Child and Adolescent Mental Health Service and Schools Link Pilot Scheme*. London: Dfe & DoH

Department for Education and Skills. (2003). *Every Child Matters*. London: DfES Publications.

Department of Health (2015). Future in mind: Promoting, protecting and improving our children and young people's mental health and wellbeing. London: NHS England.

Domitrovich, C., Bradshaw, C., Greenberg, M., Embrey, D., Poduska, J., & Ialongo, N. (2010). Integrated models of school-based prevention: Logic and theory. *Psychology in the Schools,* 47, 71-88. https://doi.org/10.1002/pits.20452

Domitrovich, C., Durlak, J.A., Staley, K.C. & Weissberg, R.P. (2017). Social-Emotional competence: An essential factor for promoting positive adjustment and reducing risk in school children. *Child Development*, *88*(2), 408–416.http://doi.org/10.1111/cdev.12739

Domitrovich, C. E., Gest, S. D., Jones, D., Gill, S. & Sanford DeRousie, R.M. (2010). Implementation quality: Lessons learned in the context of the Head Start REDI trial. *Early Childhood Research Quarterly*, *25*(3), 284–298. http://doi.org/10.1016/j.ecresq.2010.04.001

Durlak, J.A., & Wells, A.M. (1997). Primary prevention mental health programmes for children and adolescents: A meta-analytic review. *American Journal of Community Psychology*, *25(2)*, 115–52. http://doi.org/10.1023/A:1024654026646

Durlak, J., Weissberg, R., Dymicki, A., Taylor, R. & Schellinger, K. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432. http://doi.org/10.1111/j.1467 10.01564.x

Dvorsky, M., Girio-Herrera, E. & Owens, J. S. (2014). School-based screening for mental health in early childhood. In M. Weist (Ed.) Handbook of School Mental Health (pp.297–310). New York: Springer.

Ebesutani, C., Korathu-Larson, P., Nakamura, B., Higa-McMillan, C., & Chorpita, B. (2017). The Revised Child Anxiety and Depression Scale 25-Parent Version: Scale Development and Validation in a School-Based and Clinical Sample. *Assessment, 34*, 712-728. doi: 10.1177/1073191115627012

Essau, C., Conradt, J., Sasagawa, S. & Ollendick, T. (2012). Prevention of anxiety symptoms in children: results from a universal school-based trial. *Behavior Therapy, 43*(2), 450–64. Doi:10.1016/j.beth.2011.08.003

Essau, C.A., Conradt, J. & Petermann, F. (2000). Frequency, comorbidity, and psychological impairment of anxiety disorders in German adolescents. *Journal of Anxiety Disorders, 14*(3), 263–27. http://doi.org/10.1016/S0887-6185(99)00039-0

Farrell, A. D., Henry, D. B. & Bettencourt, A. (2013). Methodological Challenges Examining Subgroup Differences: Examples from Universal School-based Youth Violence Prevention Trials. *Prevention Science*, *14*(2), 121–133. http://doi.org/10.1007/s11121-011-0200-2.

Feldman, G., Greeson, J. & Senville, J. (2010). Differential effects of mindful breathing, progressive muscle relaxation, and loving-kindness meditation on decentering and negative reactions to repetitive thoughts. *Behaviour Research and Therapy, 48*(10), 1002-1011. **https://doi.org/10.1016/j.brat.2010.06.006**

Feldner, M. T., Zvolensky, M. J. & Schmidt, N. B. (2004). Prevention of anxiety psychopathology: A critical review of the empirical literature. *Clinical Psychology: Science and Practice, 11(4)*, 405−424. http://doi.org/10.1093/clipsy.bph098

Fisak, B.J., Richard, D. & Mann, A. (2011). The Prevention of Child and Adolescent Anxiety: A Meta-analytic Review. *Prevention Science*, *12*, 255–268. http://doi.org/10.1007/s11121-011-0210-0

Fritz, C., Morris, P., & Richler, J. (2012). Effect Size Estimates: Current Use, Calculations, and Interpretation. *Journal of Experimental Psychology: General, 141*, 2-18. Doi: 10.1037/a0024338

Fryers, T., Melzer, D. & Jenkins, R. (2003). Social inequalities and the common mental disorders. *Social Psychiatry and Psychiatric Epidemiology*, *38*(5), 229–237. http://doi.org/10.1007/s00127-003-0627-2**.**

Gallegos-Guajardo, J., Ruvalcaba-Romo, N. A., Garza-Tamez, M. & Villegas-Guinea, D. (2013). Social Validity evaluation of the FRIENDS for Life Program with Mexican Children. *Journal of Education and Training Studies*, *1*(1), 158-169. **http://doi.org/10.11114/jets.v1i1.90**

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*, 1337-1345. DOI: https://doi.org/10.1097/00004583-200111000-00015

Green, HA., McGinnity, A., Meltzer, H., Ford, T & Goodman, R. (2005). Mental Health of Children and Young People in Great Britain. Newport: Office for National Statistics

Greenberg, M.T. (2010). School-based prevention: current status and future challenges. *Effective Education*, *2(1)*, 27–52. DOI: 10.1080/19415531003616862

Greenberg, M. T., Domitrovich, C. & Bumbarger, B. (2001). The prevention of mental disorders in school-aged children: Current state of the field. Prevention and Treatment, 4, 1-59. http://doi.org/10.1037/1522-3736.4.1.41a

Gupta, S. K. (2011). Intention-to-treat concept: A review. *Perspectives in Clinical Research*, *2*(3), 109–12. http://doi.org/10.4103/2229-3485.83221

Harden, A., Rees, R., Shepherd, J., Brunton, G., Oliver, S. & Oakley, A. (2001). *Young People and Mental Health: A Systematic Review of Research on Barriers and Facilitators.* London, UK: EPPI-Centre

Hetzner, N., Johnson, A., & Brooks-Gunn, J. (2010). Poverty: Effects on social and emotional development. In S. Jarvela (Ed.), *Social and emotional aspects of learning* (pp. 76–86). Oxford, UK: Elsevier.

Hill, C., Bloom, H., Black, A. R. & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, *2*(3), 172–177. http://doi.org/10.1111/j.1750-8606.2008.00061.x

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., …, Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ (Clinical Research Ed.)*, *348*(MARCH 2014), g1687. http://doi.org/10.1136/bmj.g1687

Horowitz, J. L. & Garber, J. (2006). The prevention of depressive symptoms in children and adolescents: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 74*(3), 401-415. http://doi.org/10.1037/0022-006X.74.3.401

Humphrey, N. (2013) *Social and emotional learning: a critical appraisal*. London: Sage Publications.

Humphrey, N., Barlow, A., & Lendrum, A. (2018). Quality matters: Implmentation moderates student outcomes in the PATHS curriculum. Prevention Science, 19, 197-208. Doi: 10.1007%2Fs11121-017-0802-4

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R. & Kerr, K. (2016). *Implementation and Process Evaluation (IPE) for Interventions in Education Settings: An introductory handbook*. London: Education Endowment Foundation.

Hunt, C., Andrews, G., Crino, R., Erskine, A. & Sakashita, C. (2009). Randomized controlled trial of an early intervention programme for adolescent anxiety disorders. *Australian & New Zealand Journal of Psychiatry, 43*, 300–304. doi:10.1080/00048670902721152

Ignatius, E. & Kokkonen, M. (2007). Factors contributing to verbal self-disclosure. *Nordic Psychology 59*(4), 362–391. DOI: 10.1027/1901-2276.59.4.362

Iizuka, C., Barrett, P., Gillies, R., Clayton, C. and Miller, D. (2014). The FRIENDS emotional health program for minority groups at risk. *Journal of School Health*, *84*(2), 124–132. doi:10.1111/josh.12127

Jennings, P.A. & Greenberg, M.T. (2009). The Prosocial Classroom: Teacher Social and Emotional Competence in Relation to Student and Classroom Outcomes. *Review of Educational Research*, *79*(1), 491–525. http://doi.org/10.3102/0034654308325693

Kelvin, R. 2014. “Mental Health Help “Needed in Schools.” BBC News Health, March 24. http:// www.bbc.co.uk/news/health-26685550

Keogh, E., Bond, F.W. & Flaxman, P.E. (2006). Improving academic performance and mental health through a stress management intervention: Outcomes and mediators of change. *Behaviour Research and Therapy, 44*(3), 339−357. **http://doi.org/10.1016/j.brat.2005.03.002**

Kösters, M. P., Chinapaw, M. J., Zwaanswijk, M., van der Wal, M. F., Utens, E. M., & Koot, H. M. (2012). Study design of “FRIENDS for Life”: process and effect evaluation of an indicated school-based prevention programme for childhood anxiety and depression. *BMC Public Health*, *12*, 86. http://doi.org/10.1186/1471-2458-12-86

Lendrum, A., & Humphrey, N (2012). The importance of studying the implementation of interventions in school settings, Oxford Review of Education, 38:5, 635-652, DOI: 10.1080/03054985.2012.734800

Levitt, J.M., Saka, N., Romanelli, L.H. & Hoagwood, K. (2007). Early Identification of Mental Health Problems in Schools: The Status of Instrumentation. *Journal of School Psychology, 45*(2), 163-191. Doi: https://doi.org/10.1016/j.jsp.2006.11.005

Liddle, I., & Macmillan, S. (2010). Evaluating the FRIENDS programme in a Scottish setting. *Educational Psychology in Practice, 26*, 53–67.

Lock, S., & Barrett, P. M. (2003). A longitudinal study of developmental differences in universal preventive intervention for child anxiety. *Behaviour Change, 20*, 183–199.

Lowry-Webster, H. M., Barrett, P. M. & Lock, S. (2003). A Universal Prevention Trial of Anxiety Symptomology during Childhood: Results at 1-Year Follow-up. *Behaviour Change*, *20*(01), 25–43. http://doi.org/10.1375/bech.20.1.25.24843

Maggin, D. M. & Johnson, A. H. (2014). A Meta-Analytic Evaluation of the FRIENDS Program for Preventing Anxiety in Student Populations. *Education and Treatment of Children*, *37*(2), 277–306. http://doi.org/10.1353/etc.2014.0018

Masia-Warner, C., Nangle, D.W. & Hansen, D. J. (2006). Bringing evidence-based child mental health services to the schools: General issues and specific populations. *Education and Treatment of Children, 29*(2),165−172.

McNally Keehn, R.H., Lincoln, A.J., Brown, M.Z., & Chavira, D.A. (2013). The Coping Cat Program for Children with Anxiety and Autism Spectrum Disorder: A Pilot Randomized Controlled Trial. *Journal of Autism and Developmental Disorders, 43*(1), 57-67. Doi: 10.1007/s10803-012-1541-9

Meyer, T., Miller, M., Metzger, R., & Borkovec, T. (1990). Development and validation of the penn state worry questionnaire. *Behaviour Research and Therapy, 28*, 487-495. https://doi.org/10.1016/0005-7967(90)90135-6

Miller, L. D., Laye-Gindhu, A., Liu, Y., March, J. S., Thordarson, D. S. & Garland, E. J. (2011). Evaluation of a preventive intervention for child anxiety in two randomized attention-control school trials. *Behaviour Research and Therapy*, *49*(5), 315–323. http://doi.org/10.1016/j.brat.2011.02.006

Moilanen, K. L., D. S. Shaw, & K. L. Maxwell. (2010). Developmental Cascades: Externalizing, Internalizing, and Academic Competence from Middle Childhood to Early Adolescence. *Development and Psychopathology, 22*(3): 635–653. doi:10.1017/S0954579410000337

Mychailyszyn, M.P., Beidas, R.S., Benjamin, C.L., Edmunds, J.M., Podell, J.L., Cohen, J.S. & Kendall, P.C. (2011). Assessing and Treating Child Anxiety in Schools. *Psychology in the Schools, 48*(3), 223-232. Doi: 10.1002/pits.20548

Neil, A. L. & Christensen, H. (2009). Efficacy and effectiveness of school-based prevention and early intervention programs for anxiety. *Clinical Psychology Review*, *29*(3), 208–215. http://doi.org/10.1016/j.cpr.2009.01.002

Otte, C. (2011). Cognitive behavioral therapy in anxiety disorders: Current state of the evidence. *Dialogues in Clinical Neuroscience*, *13*(4), 413–421.

Owens, M., Stevenson, J., Hadwin, J. A. & Norgate, R. (2012). Anxiety and depression in academic performance: An exploration of the mediating factors of worry and working memory. *School Psychology International*, *33*(4), 433–449. http://doi.org/10.1177/0143034311427433

Pampaka, M., Hutcheson, G. & Williams, J. (2016). Handling missing data: analysis of a challenging data set using multiple imputation. *International Journal of Research & Method in Education*, *39*(1), 19–37. http://doi.org/10.1080/1743727X.2014.979146

Panayiotou, M. & Humphrey, N. (2017). Mental health difficulties and academic attainment: Evidence for gender-specific developmental cascades in middle childhood. *Development and Psychopathology*, 1-16. Doi: 10.1017/S095457941700102X

Pereira, A. I., Marques, T., Russo, V., Barros, L., & Barrett, P. (2014). Effectiveness of the friends for life program in portuguese schools: Study with a sample of highly anxious children. *Psychology in the Schools*. http://doi.org/10.1002/pits

Petticrew, M., Tugwell, P., Kristjansson, E., Oliver, S., Ueffing, E. & Welch, V. (2012). Damned if you do, damned if you don't: subgroup analysis and equity. *Journal of Epidemiology and Community Health, 66*(1), 95–98. http://doi.org/10.1136/jech.2010.121095

Putwain, D. W., Connors, L. & Symes, W. (2010). Do cognitive distortions mediate the test anxiety-examination performance relationship? *Educational Psychology, 30*, 11–26. doi:10.1080/01443410903328866

Reivich, K.J., Gillham, J.E., Chaplin, T. M., & Seligman, M.E.P. (2005). From helplessness to optimism: The role of resilience in treating and preventing depression in youth. In S. Goldstein & R.B. Brooks (Eds.) *Handbook of Resilience in Children.* (pp. 223-237). New York: Kluwer Academic/Plenum Publishers

Rothwell, P. (2005). External validity of randomised controlled trials: "to whom do the results of this trial apply? *Lancet, 365*, 82-93. doi: 10.1016/S0140-6736(04)17670-8

Rutter, M. (1982). *Fifteen Thousand Hours Secondary Schools and Their Effects on Children*. Cambridge, MA: Havard University Press.

Salovey, P., Brackett, M. & Mayer, J. (2004). *Emotional Intelligence: Key readings on the Mayer and Salovey model*. New York: Dude.

Siu, A.F.Y. (2007). Using FRIENDS to combat internalizing problems among primary school children in Hong Kong. *Journal of Cognitive and Behavioral Psychotherapies*, *7*, 11–26.

Sklad, M., Diekstra, R., Ritter, M. & Ben, J. (2012). Effectiveness of school-based universal social, emotional, and behavioral programs: Do they enhance students' development in the area of skill, behavior, and adjustment?. *Psychology in the Schools, 49*(9), 892–909. http://doi.org/10.1002/pits

Skryabina, E., Taylor, G. & Stallard, P. (2016). Effect of a universal anxiety prevention programme (FRIENDS) on children's academic performance: Results from a randomised controlled trial. *Journal of Child Psychology and Psychiatry*, *57*(11), 1297–1307. http://doi.org/http://dx.doi.org/10.1111/jcpp.12593

Spence, S. H. & Dadds, M. R. (1996). Preventing childhood anxiety disorders. *Behaviour Change, 13,* 241−249.

Stallard, P. (2010). Mental health prevention in UK classrooms: the FRIENDS anxiety prevention programme. *Emotional and Behavioural Difficulties*, *15*(1), 23–35. http://doi.org/10.1080/13632750903512381

Stallard, P., Phillips, R., Montgomery, A.A., Spears, M., Anderson, R., Taylor, J., … Sayal, K. (2013). A cluster randomised controlled trial to determine the clinical effectiveness and cost-effectiveness of classroom-based cognitive-behavioural therapy (CBT) in reducing symptoms of depression in high-risk adolescents. *Health Technology Assessment*, *17*(47). http://doi.org/10.3310/hta17470

Stallard, P., Skryabina, E., Taylor, G., Phillips, R., Daniels, H., Anderson, R. & Simpson, N. (2014). Classroom-based cognitive behaviour therapy (FRIENDS): a cluster randomised controlled trial to Prevent Anxiety in Children through Education in Schools (PACES). *The Lancet Psychiatry*, *1*(3), 185–192. http://doi.org/10.1016/S2215-0366(14)70244-5

Tymms, P. (2004). Effect sizes in multilevel models. In I. Schegen & K. Elliot (Eds.), *But what does it mean? The use of effect sizes in educational research* (pp. 55–65). London: NFER.

Viana, A. G., Rabian, B. & Beidel, D.C. (2008). Self-report measures in the study of comorbidity in children and adolescents with social phobia: Research and clinical utility. *Journal of Anxiety Disorders, 22*(5), 781-792. https://doi.org/10.1016/j.janxdis.2007.08.005

Weare, K. & Nind, M. (2010). Identifying evidence-based work on mental health promotion in schools in Europe: An interim report on the DataPrev project. *Advances in School Mental Health Promotion, 3*(2), 37–45. https://doi.org/10.1080/1754730X.2010.9715679

WHO (World Health Organization) (2004). Prevention of mental disorders: Effective interventions and policy options. Geneva: WHO.

Wigelsworth, M., Lendrum, A., Oldfield, J., Scott, A., Ten-Bokkel, I., Tate, K. & Emery, C. (2016). The impact of trial stage, developer involvement and international transferability on universal social and emotional learning programme outcomes: A meta-analysis. *Cambridge Journal of Education*, *46*(3) http://doi.org/10.1080/0305764X.2016.1195791

Woodward, L.J. & Fergusson, D.M. (2001). Life course outcomes of young people with anxiety disorders in adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry, 40(*9), 1086–93. https://doi.org/10.1097/00004583-200109000-00018

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
£ £ £ £ £	*Very low:* less than £80 per pupil per year.
£ £ £ £ £	*Low:* up to about £200 per pupil per year.
£ £ £ £ £	*Moderate:* up to about £700 per pupil per year.
£ £ £ £ £	*High:* up to £1,200 per pupil per year.
£ £ £ £ £	*Very high:* over £1,200 per pupil per year.

Appendix B: Security classification of trial findings

Appendix 2: Security rating template – KS2 Maths and Reading Combined Score

Rating	Criteria for rating			Initial score		Adjust		Final score
	Design	**Power**	**Attrition[22]**					
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 **[0]**		4
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 **[0]**		
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 = Well conducted experimental design, MDES at randomisation was 0.227 for attainment outcomes and 8% attrition
- **Reason for adjustment for balance** (if made): None made as the imbalance was 0.029 for KS1
- **Reason for adjustment for threats to validity** (if made): There was some evidence of compensatory rivalry in the control group, but as the uptake of other programmes was incipient we do not consider this to be a relevant threat to the validity of the results
- **Final padlock score:** initial score adjusted for balance and internal validity = 4 padlocks

[22] Attrition should be measured at the pupil level (even for clustered trials) and from the point of randomisation to the point of analysis.

Appendix 2: Security rating template – Self-Rated Worry

Rating	Criteria for rating			Initial score		Adjust		Final score
	Design	Power	Attrition[23]					
5 🔒	Well conducted experimental design with appropriate analysis	MDES < 0.2	0-10%	**5**		Adjustment for Balance **[0]** Adjustment for threats to internal validity **[0]**		5
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%					
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%					
0 🔒	No comparator	MDES > 0.6	over 50%					

- **Initial padlock score:** lowest of the three ratings for design, power and attrition = Well conducted experimental design, MDES at randomisation was 0.124 for self-worry outcome and 8% attrition.
- **Reason for adjustment for balance** (if made): None made as the imbalance was 0.015 for the self-rated worry
- **Reason for adjustment for threats to validity** (if made): There was some evidence of compensatory rivalry in the control group, but as the uptake of other programmes was incipient we do not consider this to be a relevant threat to the validity of the results
- **Final padlock score:** initial score adjusted for balance and internal validity = 5 padlocks

[23] Attrition should be measured at the pupil level (even for clustered trials) and from the point of randomisation to the point of analysis.

Appendix C: Theoretical model of FRIENDS

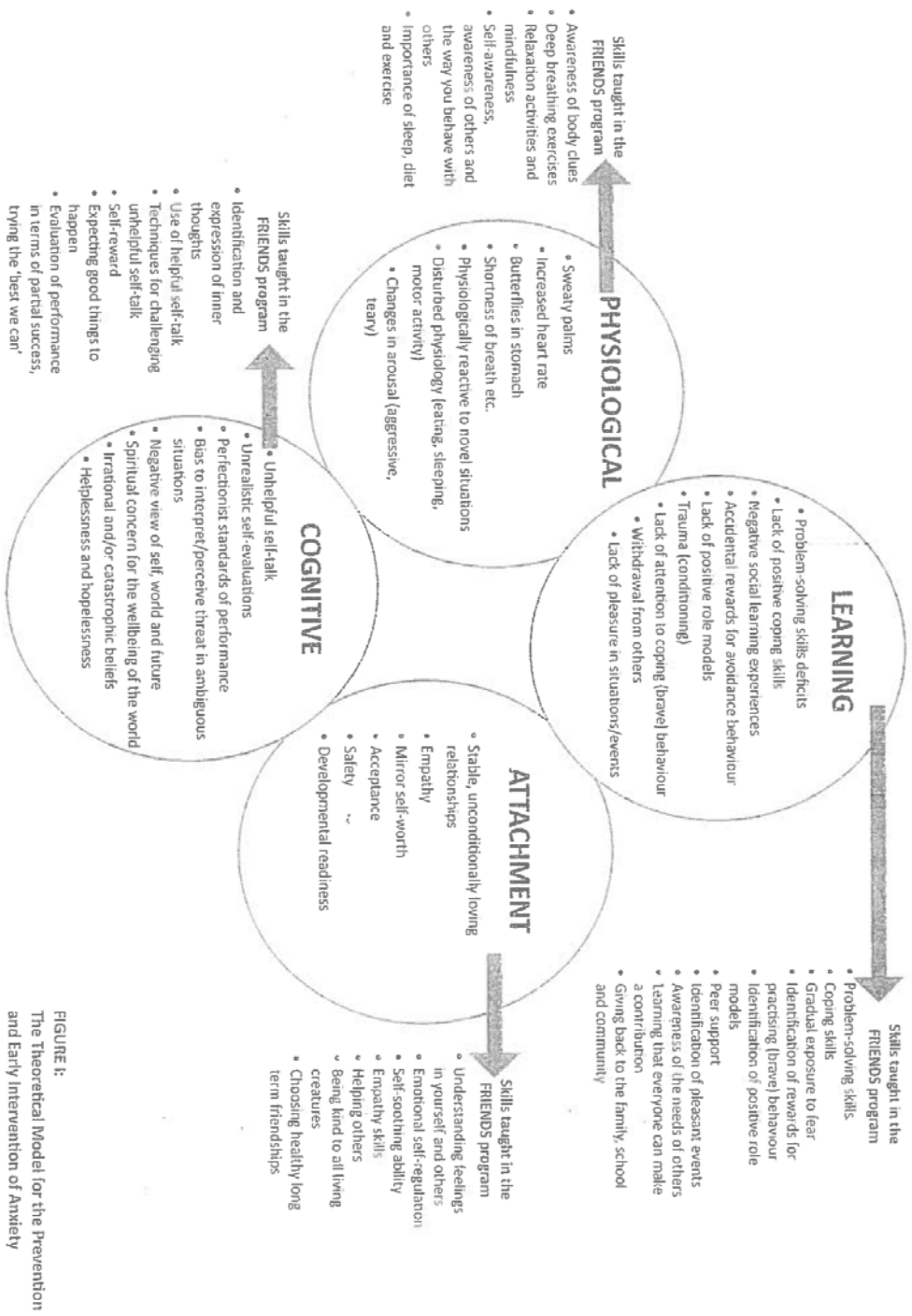


FIGURE 1:
The Theoretical Model for the Prevention and Early Intervention of Anxiety

Appendix D: Overview of the FRIENDS' session content

Session 1: Feelings – Understanding Feelings in ourselves and others. Learning to Help Others.

Objectives – The aim of the first session is to introduce group members to each other and to explain what the *FRIENDS for Life* program is about. Session 1's focus is on the recognition and sharing of feelings, the concept of being brave and accepting that there are similarities and differences between people and places.

Session 2: Introduction to Feelings

Objectives – The major aim for this session is to introduce participants to the concept of feelings, and to introduce Step 1 of the *FRIENDS for Life* plan.

Session 3: Introduction to Body Clues and Relaxation

Objectives – The goal of this session is to introduce participants to Step 2 of the FRIENDS plan.

Session 4: Helpful (Green) and Unhelpful (Red) Self-Talk

Objectives - The goal of this session is to introduce participants to Step 3 (I can do it! I can try my best!) of the FRIENDS plan.

Session 5: Changing Unhelpful Thoughts into Helpful Thoughts

Objectives - The objective of this session is to continue with Step 3 of the FRIENDS plan.

Session 6: Introduction to Coping Step Plans

Objectives - The objective of this session is to introduce participants to Step 4 of the FRIENDS plan.

Session 7: Learning From Our Role Models and Building Support Teams

Objectives - The goal of this session is to continue with Step 4 of the FRIENDS plan for feeling confident and brave, and to explore further solutions to difficult situations.

Session 8: Using a Problem Solving Plan

Objectives - The goal of this session is to continue with Step 4, Exploring Solutions.

Session 9: Using the FRIENDS Skills to Help Ourselves and Others

Objectives - The goal of this session is to introduce group participants to Steps 5, 6 and 7 of the FRIENDS plan: N = Now reward yourself! You've done your best! D = Don't forget to practice! S = Smile! Stay Calm, and talk to your Support Networks!

Session 10: Review and Party

Objectives – The first aim of this session is to establish strategies to maintain participants' coping skills for use in the future. The second aim is to congratulate group members for their participation and hard work.

2 Booster Sessions 1 & 2: Review and Practice

Appendix E: Memorandum of Agreement, Parental information Sheet & Consent Form for Parents/Guardians

FRIENDS Programme

Memorandum of Agreement

This memorandum of agreement outlines the key conditions for schools entering into partnership with Salus and the University of Manchester in the FRIENDS Programme and the associated research project. It outlines what schools that participate in the project will receive, and what they will be required to do in return. The aim is to have a completely transparent process so that all parties have a clear understanding of the project and shared expectations.

Schools that participate will either receive the full programme of support for select class(es), or will act as a comparison school, with an agreed financial incentive, for the purpose of researching the impact and implementation of the programme. Random allocation is an important part of this trial design, as it helps us produce rigorous findings. This means that for each participating school:

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

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.

Four form entry: Random allocation of 2 classes to receive FRIENDS, two classes to serve as usual practice.

Each school receiving the full programme will have support from an accredited practitioner, highly experienced in managing both interventions to improve wellbeing and whole class support.

The practitioner will deliver the initial programme, one session per week, for ten weeks to whole year 5 classes. They will then undertake two follow up sessions to reinforce learning and ensure skills are becoming embedded.

During this period, the practitioner will be available to offer more focused intervention to a small number of year 5 pupils who may require additional support.

The initial programme will need to be delivered between April and July 2016, with two short booster sessions delivered in September and November 2016, once the pupils are in year 6.

The exact timings of the programme within the school week will be as flexible as possible to accommodate the needs of the school and their pupils.

Section 1 – ABOUT YOUR SCHOOL

We need some key details about your school – please complete the form below:

Name of school	
LAESTAB code	
Address of school	
Postcode of school	
Telephone number of school	
Name of head teacher	
Email address of head teacher	

It is useful in projects like this to have a nominated 'link' person, who can co-ordinate the project within the school and act as our first point of contact. This FRIENDS Programme co-ordinator could be the head teacher, deputy/assistant head, SENCO, Key Stage 2 co-ordinator, PSHE or SEAL co-ordinator, or a class teacher from Year 5. Please provide details of the nominated link person below:

Name of FRIENDS co-ordinator	
Email address of FRIENDS co-ordinator	
Primary role within school	

Section 2 – KEY CONDITIONS OF PROJECT PARTICIPATION

In this section we outline the key conditions of project participation. Please read through them carefully.

(1) Randomisation – all schools signing this document agree that they can be randomly allocated to either (a) implement the FRIENDS programme with either one or two Year 5 classes (dependent on the size of school and randomisation) commencing in March 2016, or (b) continue their usual practice during this period.

(2) Focus – this project focuses on pupils who will be on roll in Year 5 at the start of the 2015/16 year only.

(3) Compliance with data collection requirements – all schools signing this document understand that they are committing to participation in a research project with certain data collection requirements. These are:

i) Pupil surveys – assessing social and emotional skills and health-related quality of life (once in Feb 2016, and again in December 2016)

a. Brief staff surveys – assessing emotional health and behaviour of the pupil's in their classes
b. School survey – a single survey, likely completed by the nominated contact, detailing current interventions and climate of the school

(4) For schools <u>randomly allocated to the FRIENDS group only</u>:

a. A commitment to implement the programme for one or two year 5 classes (dependent on the size of school and randomisation).
b. Observations of FRIENDS classes by University researchers
c. Requests for brief interviews with staff (e.g. head teacher, SENCO, class teacher) to explore why they have chosen to participate.
d. Requests for brief focus groups with pupils attending FRIENDS classes to explore their likes and dislikes of the programme.

Section 3 – WHAT PARTICIPATING SCHOOLS WILL RECEIVE

This section outlines what each participating school will receive as part of the FRIENDS programme. There is no financial cost whatsoever associated with participation – all of the items outlined below are covered by our project funding.

All participating schools – regardless of whether they are randomly assigned to the FRIENDS or comparison group, will receive:

(1) Bespoke aggregated feedback from our pupil and teacher surveys following the completion of the project

In addition, schools randomly allocated to the comparison group only will receive:

(1) £1,000 at the end of the school year

In addition, schools randomly allocated to the FRIENDS group (either one or two classes) will receive:

(1) FRIENDS for children in Years 5, comprising:
   a. 10 weekly sessions of approximately one hour and small group follow support
   b. 2 parent information sessions
   c. 2 follow up (booster) sessions

(2) FRIENDS curriculum materials, comprising:
   a. Lesson packs including 1 copy of 'FRIENDS for life' work book for each child
   b. All supplementary materials

The above items are subject to compliance with the data collection requirements outlined earlier. The details of these conditions are outlined in the next section.

Section 4 – WHAT WE NEED IN RETURN

In this section we outline the conditions that need to be fulfilled by each school in order for them to receive the items outlined in Section 3. These are as follows:

(1) In order to receive either payment as a comparison, or the FRIENDS programme as an intervention school, we require a minimum 85% response rate for the survey data. Specifically, 85% pupils to complete their self-report questionnaires, and 85% of pupils covered by the teacher brief reports. This is to ensure a quality trial by using this data to balance out the comparison and FRIENDS schools and to ensure we have enough responses to make the trial worthwhile.
(2) For the pupil surveys by the end of a pre-specified survey period.
(3) For schools randomly allocated to the FRIENDS group only, receipt of the training, curriculum materials and ongoing technical support and assistance is conditional upon compliance with section 2.4 above, namely assistance and co-operation in speaking with staff and pupils.

We recognise that involvement in this project requires a time commitment on behalf of participating schools. In recognition of this we will do everything we can to support all schools to meet our data collection requirements. This includes (but is not limited to):

- Members of Salus providing on-site support to facilitate survey completion if required
- Detailed, clear instructions for survey completion processes, given well in advance in order to allow for adequate planning at the school level

Section 5 – COMMITMENT TO PARTICIPATION

Having read all of the above, if you are happy to participate in the FRIENDS programme, please sign below on behalf of your school.

If you have any queries, about the FRIENDS programme or its delivery, please email Sally Williamson **sally.williamson@salusgroup.org.uk**. If you have a question relating to the research project contact the principal investigator Michael Wigelsworth at **michael.wigelsworth@manchester.ac.uk**

If you would like more information about Salus, please visit the website, **www.salusgroup.org.uk**, where you can find detailed information about the programme.

School commitment to participation

I confirm that I have read and understood all of the above and am happy to commit to participation in the FRIENDS Programme on behalf of my school:

(print school name)

(print name)

(print role)

(signature)

Salus and University of Manchester commitment

On behalf of the Salus and the University of Manchester, we confirm our commitment to provide the items specified in this document to participating schools in the FRIENDS Programme.

Principal Investigator

0161 306 1763

Michael.wigelsworth@manchester.ac.uk

Director

01303 817470

sally.williamson@salusgroup.org.uk

The FRIENDS Programme: An evaluation of academic and emotional health outcomes

Information Sheet for Parents/ Guardians

Your child's school has arranged to participate in research being conducted by the University of Manchester. We have been independently commissioned by the Educational Endowment Foundation to see if a programme called FRIENDS works for children in Year 5. We will be looking to see if FRIENDS works just as well in different schools, whether it is useful for different groups of pupils (e.g. boys and girls) and whether it helps improve pupil grades.

FRIENDS consists of 10 weekly class-based sessions for all pupils in a class. The classes aim to promote various skills for the children's emotional heath, such as helping them to recognise symptoms of anxiety & supporting them with their relationships with classmates. The classes are being delivered by 'Project Salus CIC'. More information about the FRIENDS programme can be seen on this website: **http://www.friendsforlife.org.nz/**

We are writing to you because your child's class may be taking part in these lessons as part of our research. Please take time to read the following information carefully and decide whether or not you would like your child to take part in the research.

If you would like any more information or have any questions about the research project, please contact me, Dr. Michael Wigelsworth on 0161 306 1763 or by email: **michael.wigelsworth@manchester.ac.uk**

Who will conduct the research?

The research will be conducted by myself and my research team at the Manchester Institute of Education, University of Manchester, Oxford Road, Manchester M13 9PL. My research assistants may visit your school to collect information about the FRIENDS programme.

Title of the research

"The FRIENDS Programme: An evaluation of academic and emotional health outcomes"

What is the aim of the research?

Our main aim is to examine whether FRIENDS lessons make a difference to the academic attainment and emotional health of children in primary schools in England.

Where will the research be conducted?

Primary schools in Kent Local Authority are taking part in this study.

What is the duration of the research?

The project itself runs from August 2015 until December 2017. The schools that provide FRIENDS (see below) will do so from April 2016 to July 2016.

Why have I been chosen?

We are writing to you because your child may be taking part in the FRIENDS lessons and we want to let you know about the research that we will do to find out if the classes work and help the children. Year 5 Classes will be randomly chosen to either teach FRIENDS in the 2016 spring term for a period of 10 weeks (FRIENDS classes), or continue as normal (comparison classes). We will be collecting information in both FRIENDS and comparison classes. After the end of the 2016 academic year, all classes will then decide whether they wish to start/continue using FRIENDS.

What would my child be asked to?

Your child will be asked to complete a short survey about how they feel about themselves. An example of one of the items is:

I worry that I will suddenly get a scared feeling when there is nothing to be afraid of	Never	Sometimes	Often	Always

If your child needs help to do this, they will be able to get support from school staff or one of our researchers. Your child's class teacher will also complete a brief questionnaire about your child's strengths and weaknesses.

These surveys will be completed two times – Once in February 2016, and once in December 2016. They will take approximately 20-25 minutes to complete each time. In giving consent you are giving permission for your child and his/her teacher to complete these surveys.

What happens to the data collected?

The data will be analysed by our research team at the University of Manchester. We will write a report based on our analyses for the Educational Endowment Foundation. It is also likely that we will write articles for academic journals based on what we find out in the project. It is possible that we will write a book about the research. Your child's name will not be used in any of the reports that we write. Anonymised data may be included in the National Pupil Database or other official records, and included in the UK Data archive. We will access some of this anonymised data to aid with the analysis. For instance, we will see if the programme works just as well for all pupils who are eligible for free school meals.

How is the information about my child kept private?

All data provided will be treated as confidential and will be completely anonymous. Identifying information (e.g. your child's name) will only be used in order to match information about them from different sources (e.g. from teachers and your child) across different times (e.g. to match February 2016, information with December 2016 information). After this matching process is complete, all identifying information will be destroyed – We will not match pupil names to other sources of data (e.g. gender or school). All survey data will be stored on a secure, password protected computer to which only senior members of the research team have access.

What happens if I do not want my child to take part or I change my mind?

It is up to you if you want your child to take part in the research. If you are happy for your child to participate, you do not need to do anything.

If you decide that you do not wish for your child to take part then you need to either complete the form enclosed and return it to our research team at the address above, or contact me, Dr. Michael Wigelsworth, by telephone or email (details below) before 22nd February.

If you decide that your child can take part and then change your mind, you are free to withdraw your child up until July 2017 (when the final report will be written) without needing to give a reason. If you do this please rest assured that we will destroy any data collected about your child as part of the study.

Will I be paid for participating in the research?

We are not able to offer any payment or incentive for participating in this study.

Disclosure and Barring Service

Every member of our research team has undergone a Disclosure and Barring Service check at the Enhanced Disclosure level.

Who has reviewed the research project?

This project has been reviewed by the University of Manchester Research Ethics Committee 3

What if something goes wrong?

If you have any concerns about your child's wellbeing then you should contact the school in the first instance and ask to speak to his/her teacher.

You can also get independent support and advice from a charity called Young Minds. Their parent helpline number is 0808 802 5544.

What if I want to complain?

If you have any concerns or want to complain, you should contact me, Dr. Michael Wigelsworth, in the first instance (contact details below).

If you wish to make a formal complaint about the conduct of the research you can contact a Research Governance and Integrity Manager, Research Office, Christie Building, University of Manchester, Oxford Road, Manchester, M13 9PL, by emailing: **research.complaints@manchester.ac.uk** or by telephoning 0161 275 2674 or 275 8093

Contact for further information

Questions about the research or the data collectio	**Questions about FRIENDS:**
Dr. Michael Wigelsworth Manchester Institute of Education University of Manchester, Oxford Road, Manchester M13 9PL Tel: 0161 306 1763 Email: **michael.wigelsworth@manchester.ac.uk**	Salus – Head Office Greenacres Barn Pound Lane Smeeth Ashford Kent TN25 6RJ Telephone: 01303 817470 Email: **info@salusgroup.org.uk**

Also, please see our website for further details about the FRIENDS classes and background, the project design and project team.

The website can be found at: www.friendsforlifestudy.org.uk

This Project Has Been Approved by the University of Manchester's Research Ethics Committee 3

The FRIENDS Programme: An evaluation of academic and emotional health outcomes

Consent form for Parents/ Guardians

An information sheet is attached to this form. Please read it carefully before making a decision about taking part.

If you are willing for your child to take part then you do not need to do anything at the moment.

If you decide your child should not take part, then you need to complete the opt-out slip below and return it to:

Dr. Michael Wigelsworth,
Manchester Institute of Education,
University of Manchester,
Oxford Road,
Manchester,
M13 9PL.

Alternatively, Dr. Wigelsworth can be contacted by telephone on 0161 306 1763 or email at michael.wigelsworth@manchester.ac.uk. If you do not wish your child to participate please let us know before 22nd February.

Finally, please also remember that if you do decide your child can take part, you are free to change your mind at any point in the study and without giving your reasons for doing so.

To: Dr Wigelsworth

I **do not** wish for my child to participate in the research. My details are as follows:

My name	
My child's name	
Name of my child's school	

Signed: _____ Date: _____

Appendix F: Pupil Surveys (RCADS25 & PSWQ-C)

Pupil Name:	**«Forename» «Surname»**
School:	**«School_Name»**

Dear teacher/pupil

Please tear off this first page and destroy

once you have finished

All About Me!

How are you? How do you feel? This is what we would like you to tell us.

Please read each question carefully. What answer comes to your mind first?

Chose the box that fits your answer best.

Here is an example:

I am happy	Never	Sometimes	Often	Always

This person is often happy!

Remember: This is not a test. There are no wrong answers. You do not have to show your answers to anybody.

Nobody who knows you will look at your questionnaire once you have finished with it.

Please answer **all** of the questions.

Turn over the page to start…

How I feel

Please read every question carefully and circle the answer that fits you best. Please ask a teacher if you get stuck.

1	I feel sad or empty	Never	Sometimes	Often	Always
2	I worry when I think I have done poorly at something	Never	Sometimes	Often	Always
3	I would feel afraid of being on my own at home	Never	Sometimes	Often	Always
4	Nothing is much fun anymore	Never	Sometimes	Often	Always
5	I worry that something awful will happen to someone in my family	Never	Sometimes	Often	Always
6	I am afraid of being in crowded places (like shopping centers, the movies, buses, busy playgrounds)	Never	Sometimes	Often	Always
7	I worry what other people think of me	Never	Sometimes	Often	Always
8	I have trouble sleeping	Never	Sometimes	Often	Always
9	I feel scared if I have to sleep on my own	Never	Sometimes	Often	Always
10	I have problems with my appetite	Never	Sometimes	Often	Always
11	I suddenly become dizzy or faint when there is no reason for this	Never	Sometimes	Often	Always
12	I have to do some things over and over again (like washing my hands, cleaning or putting things in order)	Never	Sometimes	Often	Always
13	I have no energy for things	Never	Sometimes	Often	Always
14	I suddenly start to tremble or shake when there is no reason for this	Never	Sometimes	Often	Always
15	I cannot think clearly	Never	Sometimes	Often	Always

16	I feel worthless	Never	Sometimes	Often	Always
17	I have to think of special thoughts (like numbers or words) to stop bad things from happening	Never	Sometimes	Often	Always
18	I think about death	Never	Sometimes	Often	Always
19	I feel like I don't want to move	Never	Sometimes	Often	Always
20	I worry that I will suddenly get a scared feeling when there is nothing to be afraid of	Never	Sometimes	Often	Always
21	I am tired a lot	Never	Sometimes	Often	Always
22	I feel afraid that I will make a fool of myself in front of people	Never	Sometimes	Often	Always
23	I have to do some things in just the right way to stop bad things from happening	Never	Sometimes	Often	Always
24	I feel restless	Never	Sometimes	Often	Always
25	I worry that something bad will happen to me	Never	Sometimes	Often	Always

My worries

Please read every question carefully and circle the answer that fits you best. Please ask a teacher if you get stuck.

1	My worries really bother me	Never	Sometimes	Often	Always
2	I don't really worry about things	Never	Sometimes	Often	Always
3	Many things make me worry	Never	Sometimes	Often	Always
4	I know I shouldn't worry, but I just can't help it	Never	Sometimes	Often	Always
5	When I am under pressure, I worry a lot	Never	Sometimes	Often	Always
6	I am always worrying about something	Never	Sometimes	Often	Always
7	I find it easy to stop worrying when I want	Never	Sometimes	Often	Always
8	When I finish one thing, I start to worry about everything else	Never	Sometimes	Often	Always
9	I never worry about anything	Never	Sometimes	Often	Always
10	I've been a worrier all my life	Never	Sometimes	Often	Always
11	I notice that I have been worrying about things	Never	Sometimes	Often	Always
12	Once I start worrying, I can't stop	Never	Sometimes	Often	Always
13	I worry all the time	Never	Sometimes	Often	Always
14	I worry about things until they are done	Never	Sometimes	Often	Always

You are doing really well, just a few more questions on the next page.

Thank-you very much!

There are no more questions. Please hand this to your teacher.

Appendix G: Power and Sample Size (PASS)

It is worth noting that although the following PASS calculations are based on detecting changes in attainment, this is a theoretically plausible *distal* consequence of the intervention. This is likely to mean the study is overpowered to detect the *proximal* effects of anxiety, depression, and worry. However, there are a small number of factors that mitigate this risk. For instance, the ICC for attainment is greater for that than anxiety, and there is strong cause to ensure subgroup differences (i.e. heightened levels of baseline anxiety) are included in the analysis.

As noted earlier, there is limited extant data on the impact of FRIENDS on attainment; therefore arriving at a precise effect size to use in our power calculations is difficult. Although a number of reviews have examined the effects CBT based intervention with children (e.g. Ishikawa, Okajima, Matsuoka, & Sakano, 2007; Cartwright-Hatton, Roberts, Chitsabesan, Fothergill, & Harrington, 2004), these have not included attainment as an outcome.

A further consideration is changes to the KS2 examination currently being introduced. In the summer term of 2016, children in year 2 and year 6 will undertake revised SATS examinations, reflecting the new national curriculum (Department for Education, 2014). The new examinations will be marked externally, cover reading, grammar and punctuation (including spelling), maths, and science (selected sample only). Scaled scores will be provided, and can be compared to national averages. For the trial cohort, KS1 examination (i.e. pre-test covariate) is based on the 'old' system, whereas the KS2 examination (i.e. post-test) will be based on the 'new' system. Although data will become available regarding the correlation between these two systems, this will not be available until after the trial has started and the protocol published. Therefore, power calculations do not include controlling for academic pre-test data (i.e. KS1).

Therefore, we have powered the study for a minimal detectable effect size (MDES) of 0.10. This represents a pragmatic limit in terms of a practical demonstration of effect, and is close to prior discussion relating to upper thresholds of available implementation resources:

All calculations assume*: N= 28 per cluster (Department for Education, 2015a); ICC (class level) = 0.17*[24]*, Power=0.8, Alpha=.05,*[25] *proportion of single form entry = 51%*[26]*, overall* 10% attrition

MDES	Number of pupils	Number of classes	Approx. number of schools (of which are single form entry)
0.10	3,300	110	77 (55)

[24] Estimate drawn from the largest academic ICC (KS2 results in writing) from the PATHS trial
[25] 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.
[26] Based on approximate data of Kent school sizes, provided by Salus

Appendix H: Usual Practice Survey

The FRIENDS Programme: An evaluation of academic and emotional health outcomes

Usual practice and school culture survey

School Name	
Person completing form	
Position in school	

This survey is designed to determine the usual practice of your school with regard to emotional health and related areas. There are three parts. Part 1 asks about named initiatives, programmes and interventions that your school may already be implementing (e.g. SEAL). Part 2 is about specific activities that address a pupil's emotional health. Part 3 asks general questions about your perception of the school climate and ethos. The survey provides an important baseline that allows us to establish your 'usual practice'; please answer as honestly as possible. Because the questions cover practice across the whole of Key Stage 2 in your school, **we strongly advise you to consult with relevant teaching staff in completing the survey**, particularly in Section 3.

The survey should take no more than 25 minutes to complete. All data will be treated as anonymous and confidential. Thank-you for cooperation.

PART 1 - ABOUT THE NAMED INITIATIVES BEING IMPLEMENTED IN YOUR SCHOOL

1. Which of the following named initiatives, programmes and interventions relating to social and emotional learning are your school <u>currently</u> implementing in Key Stage 2?

	Not implementing	Just getting started	Well underway	Fully embedded
SEAL whole school resources (e.g. termly themes, assemblies, posters etc.)				
SEAL whole class lessons for KS2 ('Yellow and Green SEAL')				
SEAL small group work for KS2 ('Silver SEAL')				
Family SEAL ('Gold SEAL')				
SEAL Staff development activities ('Purple SEAL')				

Targeted Mental Health in Schools				
National Healthy Schools Programme				
(Jenny Moseley's) Circle Time				
Circle of Friends				
Nurture Groups				
Inclusion Development Programme – BESD strand				
Achievement for All				
Place2Be				
Seasons for Growth				
Friends Forever				
Progress Programme (Antidote)				
Behaviour for Learning (BfL)				
Restorative Justice				
'Imported' social-emotional learning curriculum (e.g. Second Step, Incredible Years)				
Building Learning Power				
Philosophy for Children				
1 to 1 counselling				
Yoga Therapy				
Mindfulness				
Drama Therapy				
Gardening Therapy				
Other (please specify) _____				

Part 2 – ABOUT THE SPECIFIC STRATEGIES USED TO PROMOTE SOCIAL AND EMOTIONAL LEARNING IN YOUR SCHOOL

Please indicate which of the following artefacts, resources, or activities are used in Key Stage 2 classes in your school (tick any that apply, and please indicate to whom they are accessible):

❑ Posters or other visual displays that promote emotional awareness, understanding and vocabulary about feelings (e.g. happy, sad faces)

- ❑ Accessible to all pupils in a given class
- ❑ Used specifically with pupils in need of additional support

❑ School practices or physical spaces that support meditative strategies (i.e. have quiet time - think and breathe)

- ❑ Accessible to all pupils in a given class
- ❑ Used specifically with pupils in need of additional support

❑ Posters or other resources that indicate rules or guidelines for problem solving and/or self-help behaviour strategies (e.g. 'take things step by step to achieve your goals')

- ❑ Accessible to all pupils in a given class
- ❑ Used specifically with pupils in need of additional support

❑ Posters, other resources or systems to encourage pupils to reward their *own* behaviours ('well done for trying hard' – e.g. stickers given by teachers don't count, but encouragement of self-praise does)

- ❑ Accessible to all pupils in a given class
- ❑ Used specifically with pupils in need of additional support

❑ Posters or other resources or school practices that support social networks – e.g. playground buddies or friendship circles to encourage children to make friendship groups

- ❑ Accessible to all pupils in a given class
- ❑ Used specifically with pupils in need of additional support

❑ Narratives that implicitly or explicitly promote social and emotional learning and friendships (e.g. a short story in which a central character learns about the importance of understanding other people's perspectives) (tick one only; if both apply please tick 'all pupils'):

- ❑ Accessible to all pupils in a given class
- ❑ Used specifically with pupils in need of additional support

Part 3 –THE CLIMATE IN YOUR SCHOOL

This section focuses on the climate in your school. Please read each statement and indicate your level of agreement.

In my school…	Disagree strongly 1	Disagree 2	Agree 3	Agree strongly 4
1. Pupils show little concern for one another	1	2	3	4
2. Pupils are very friendly with one another	1	2	3	4
3. Pupils are kind and supportive of one another	1	2	3	4
4. Pupils are hard to control	1	2	3	4
5. Pupils frequently argue and fight	1	2	3	4
6. Pupils come to my school without having been taught by their families to be responsible or to abide by clear moral values	1	2	3	4
7. There are generally good relations between teachers and pupils	1	2	3	4
8. Teachers like the pupils and treat them with respect	1	2	3	4
9. Pupils do not respect their teachers	1	2	3	4
10. Teachers are interested in what pupils do outside of school	1	2	3	4
11. There is a great deal of co-operative effort among staff members	1	2	3	4
12. Teachers are supportive of one another	1	2	3	4
13. This school seems like a big family, everyone is so close and cordial	1	2	3	4
14. You can count on staff members to help out anywhere, anytime	1	2	3	4
15. Teachers frequently consult with and help one another	1	2	3	4
16. Teachers keep to themselves	1	2	3	4
17. Staff fall into conflicting cliques	1	2	3	4
18. Teachers provide an intellectually stimulating and challenging learning environment for their pupils	1	2	3	4
19. The teachers are very talented	1	2	3	4
20. Teachers are continually learning and seeking new ideas	1	2	3	4
21. It is an intellectually stimulating place for teachers	1	2	3	4

Appendix I: Observation schedule

Observation schedule

Date (dd/mm/yy)	
School name	
Project Officer	
FRIENDS lesson number	
Observation start time	
Observation end time	
Name of observer	
Number of pupils in class	

General contextual notes	
Contextual information that might be relevant to the conduct of the lesson (e.g. other things happening in the class or school). Remember 'sensitising' concepts.	
Descriptive Comments	***Interpretations***
Number of adults supporting class: ____	

1. **Fidelity/adherence**

Rate the extent to which the implementer delivers the lesson with fidelity to the FRIENDS guidance:

1a. Coverage of lesson objectives

- *To what extent does the project officer cover the general and specific objectives of the lesson?*

None									*All*

1b. Adherence to lesson structure and sequence

- *To what extent does the implementer follow the structure and sequence of activities outlined in the lesson guidance? E.g. introduction, core activities, closure.*

None									*All*

Did the session complete all of the objectives as listed in the lesson guidance?

Yes
If no, please list sections omitted

1c. Core components

- *How closely does the implementer adhere to the guidance when teaching the core activities of the lesson? e.g. content, suggested mode of delivery.*

None									*All*

Fidelity notes	
Descriptive Comments	***Interpretations***
1a.	
1b.	
1c.	

2. Adaptations

Adaptation (addition, omission and enhancement) notes		
Activity number	***Descriptive Comments***	***Interpretations***

3. Quality

Rate the quality of delivery of the lesson:

3a. Preparedness

- *How well prepared is the implementer for the lesson?*

Not at all									*Completely*

3b. Interest and enthusiasm

- *Rate the implementer's interest and enthusiasm in his/her delivery of the lesson*

Not interested									*Fully involved*

3c. Clarity of expression

- *How clearly does the implementer explain key concepts and activities in the lesson?*

Does not attempt									*Fully*

3d. Implementer responsiveness as required

- *How well does the implementer respond to pupil queries/ meet the needs of all of the class if it is required?*

None									*All*

How well does the PO relate to the rest of the class (e.g. quality of interaction)?

Poor									*Excellent*

Quality notes	
Descriptive Comments	***Interpretations***
3a.	
3b.	
3c.	
3d.	

4. Participant responsiveness

Rate children's engagement with and responsiveness to the lesson

4a. Pupil engagement in core activities

- *Rate the extent to which children in the class actively participate in the lesson activities (e.g. joining in role plays, answering questions).*

None									*Fully*

4b. Pupil interest levels

- *Rate the level of <u>sustained</u> interest and attentiveness among children in the class* ***during the lesson.***

None									*Fully*

4c. Pupil interest levels

- Rate the level of sustained interest and attentiveness among children in the class **to the Project Officer**.

None									*Fully*

Fully

4d. Pupil learning

- *Rate the extent to which the learning objectives have been met.*

None									*All*

Participant responsiveness notes	
Descriptive Comments	***Interpretations***
4a.	
4b.	
4c.	
4d.	

5. Reach

Approximately what proportion of the class are present throughout the lesson?

None									*All*

Participant reach and withdrawal notes	
Descriptive Comments	***Interpretations***

6. Overall implementation quality

Provide a summative rating of implementation quality

Extremely poor									*Extremely good*

Overall implementation quality notes, contributing factors and justification	
Descriptive Comments	***Interpretations***

Any other notes	
Descriptive Comments	***Interpretations***

Appendix J: Details of analysis code

This is an example of how to run the multilevel models presented in this report. Comments are preceded by "#". Multilevel imputation requires the use of REALCOM-Impute in combination with MLwiN and the R package R2MLwiN.

Note: indentation is not required; it is only displayed as such in certain parts of this exemplar code to improve readability.

 # Install and load required package

install.packages("R2MLwiN")

library(R2MLwiN)

 # Specify the path to MLwiN 2.36 to run the models.

options(MLwiN_path = "C:/Program Files (x86)/MLwiN v2.36/i386")

 # All subsequent models require a dataframe called "data"

 # estoptions can include the option "debugmode = TRUE" to retrieve multiply-imputed datasets generated by REALCOM-Impute and generate pooled estimates within MLwiN before importing back to R

 # Empty model

empty <- runMLwiN(zks2 ~ cons + (cons | sid) + (cons | cid) + (cons | pid), estoptions = list(EstM = 0, resi.store = F), data = data)

 # ITT model

Itt <- runMLwiN(zks2 ~ cons + zks1 + factor(allocation) + (cons | sid) + (cons | cid) + (cons | pid), estoptions = list(EstM = 0, resi.store = F), data = data)

 # ITT model with minimisation variables (minaps, minrcads)

minim <- runMLwiN(zks2 ~ cons + zks1 + factor(allocation) + factor(minaps) + factor(minrcads)+ (cons | sid) + (cons | cid) + (cons | pid), estoptions = list(EstM = 0, resi.store = F), data = alldata)

 # Subgroup analysis

 # Raised SDQ internalising scores (intcat1)

subgroup1 <- runMLwiN(zks2 ~ cons + zks1 + factor(allocation) + factor(intcat1) + zks1*factor(allocation) + factor(intcat1)*factor(allocation) + (cons | sid) + (cons | cid) + (cons | pid), estoptions = list(EstM = 0, resi.store = F), data = alldata)

 # FSM-eligible (everfsm2)

subgroup2 <- runMLwiN(zks2 ~ cons + zks1 + factor(allocation) + factor(everfsm2) + zks1*factor(allocation) + factor(everfsm2)*factor(allocation) + (cons | sid) + (cons | cid) + (cons | pid), estoptions = list(EstM = 0, resi.store = F), data = alldata)

 # Calculate minimum detectable effect size (MDES) with $1 - \beta = 0.8$ and $\alpha = 0.05$

```
MDES <- function(one.tail = TRUE, vpcempty = 0, varexp2 = 0, varexp1 = 0,
                 proprandom = 0, nrandom = 0, avsize = 0) {
  if(one.tail == TRUE) {
    mdes1 <- 2.5*(sqrt(((vpcempty*(1 - varexp2))/
                        (proprandom*(1 - proprandom)*nrandom)) +
                       ((1 - vpcempty)*(1 - varexp1))/
                       ((proprandom*(1 - proprandom))*avsize*nrandom)))
    print((paste("One-tailed MDES", " = ", round(mdes1, digits = 3))))
  }
  else {
    mdes2 <- 2.8*(sqrt(((vpcempty*(1 - varexp2))/
                        (proprandom*(1 - proprandom)*nrandom))+
                       ((1 - vpcempty)*(1 - varexp1))/
                       ((proprandom*(1 - proprandom))*avsize*nrandom)))
    print(paste("Two-tailed MDES", " = ", round(mdes2, digits = 3)))
  }
}
```

 # Calculate post-hoc MDES

 # Estimate a model with baseline only to increase power

```
baseline <-runMLwiN(zks2 ~ cons + zks1 + (cons | sid) + (cons | cid) + (cons | pid), estoptions = list(EstM
= 0, resi.store=F), data = alldata)
```

 # Estimate 2-tailed MDES using results from empty model and baseline only model

```
MDES(one.tail = F,
     vpcempty = as.numeric((empty@RP[1] + empty@RP[2]) / (
       empty@RP[1] + empty@RP[2] + empty@RP[3])),
     varexp1 = as.numeric((empty@RP[3] - baseline@RP[3])/empty@RP[3]),
     varexp2 = as.numeric(((empty@RP[1] + empty@RP[2]) -
       (baseline@RP[1] + baseline@RP[2])) /
       (empty@RP[1] + empty@RP[2])),
```

```
    proprandom = 62/122, nrandom = 122, avsize = 3010/122)
# Estimate 2-tailed MDES using results from empty model only (no baseline measure)
MDES(one.tail = F,
    vpcempty = as.numeric((emptyks2imp@RP[1]+emptyks2imp@RP[2])/(
      emptyks2imp@RP[1]+emptyks2imp@RP[2]+emptyks2imp@RP[3])),
    proprandom = 62/122, nrandom = 122, avsize = 3010/122)
```

Appendix K: Methodological notes

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

An ITT analysis was conducted for H1, H2, H3 (a-e). This was done in accordance to intention-to-treat principles, e.g. ignoring noncompliance, protocol deviations and other events that take place after randomisation (Gupta, 2011). This analysis was carried out through fitting 3-level (schools, classes, pupils) hierarchical models to account for the nested nature of the data. Allocation to the FRIENDS intervention was done at the class-level, which means that a multilevel model must be fitted to account for the different sources of variation. The school-level was included as not doing so would have potentially biased the results, since two classes within the same school are more likely to be similar than two classes in different schools. From a statistical point of view, ignoring the school level would spuriously inflate the class variance and hence overestimate the effect of FRIENDS.

All complete-case analysis models were fitted using MLwiN Version 2.36, while the imputation models were fitted using REALCOM-Impute and then plugged back into MLwiN for the pooling. The models presented in this report for the MI correspond to the pooled results.

First, we implemented an initial unconditional (empty) model to ascertain the amount of variance attributable to schools, classes and pupils, followed by models with treatment allocation (e.g. FRIENDS vs. comparison) included at the class level. Models for H1, H2 and H3a-c have normally-distributed outcomes and hence the ITT models have the following algebraic form:

$$y_{t(ijk)} = \beta_{0ijk} + \beta_1 Friends_{jk} \qquad \textbf{Eq. 1}$$
$$\beta_{0ijk} = \beta_0 + v_{00k} + u_{0jk} + e_{ijk}$$
$$where:$$
$$v_{00k} \sim N(0, \sigma_v^2)$$
$$u_{0jk} \sim N(0, \sigma_u^2)$$
$$e_{ijk} \sim N(0, \sigma_e^2)$$

Where $y_{t(ijk)}$ is the standardised outcome (primary or secondary) at post-test (time "t") of the i-th pupil in the j-th class of the k-th school; β_0 is the intercept or overall average; β_1 is the effect of the allocation of the j-th class in the k-th school to the FRIENDS intervention. The second line of equation corresponds to the random part of the multilevel model, where v_{00k} is the effect uniquely attributable to the k-th school, which follows a normal distribution with mean 0 and variance σ_v^2; u_{0jk} is the unique effect of the j-th class of the k-th school, which also follows a normal distribution with mean 0 and variance σ_u^2; and finally e_{ijk} represents pupils' heterogeneity, which is also assumed to be normally distributed with a mean of 0 and a variance σ_e^2.

For the case of models for H3d and H3e, the distribution of SDQ scores had a highly skewed distribution, which made it impossible to implement a transformation to normally-distributed scores. Therefore, SDQ scores were dichotomised as per current SDQ guidelines[27] into scores of 0 for "close to average" and 1 for "slightly raised". This binary outcome was modelled with a multilevel logistic regression, which had the following form:

[27] http://www.sdqinfo.com/py/sdqinfo/b3.py?language=Englishqz(UK)

$$logit\left(\pi_{ijk}\right) = ln\left(\frac{\pi_{ijk}}{1-\pi_{ijk}}\right) = \beta_0 + X_{ij}\beta + v_{00k} + u_{0jk} \quad \textbf{Eq. 2}$$

$$\beta_{0ijk} = \beta_0 + v_{00k} + u_{0jk}$$
$$where:$$
$$v_{00k} \sim N(0, \sigma_v^2)$$
$$u_{0jk} \sim N(0, \sigma_u^2)$$

In the case of equation 2, $ln\left(\frac{\pi_{ijk}}{1-\pi_{ijk}}\right)$ corresponds to the natural logarithm of the odds a pupil i in class j and school k to have a slightly raised SDQ score. $X_{ij}\beta$ corresponds to a vector of covariates, which are defined as in equations 1 and 5. Estimated coefficients are in the so-called log-odds scale, which can transformed back to odds ratio using the inverse of the natural logarithm "exp". Odds ratios have no unit of measure and hence are customarily used as measures of effect size.

Considering equation 1, the expected value of the outcome (primary or secondary) of a pupil in a class allocated to the FRIENDS intervention ($\hat{f} = [\hat{y}|Friends = 1]$) would be as follows:

$$E(\hat{f}) = \beta_0 + \beta_1 \quad \textbf{Eq. 3}$$

While in comparison, the expected value of the outcome for a pupil in a class under usual practice ($\hat{c} = [\hat{y}|Friends = 0]$) would be:

$$E(\hat{c}) = \beta_0 \quad \textbf{Eq. 4}$$

Given that $y_{t(ijk)}$ and $y_{t-1(ijk)}$ are standardised scores, and $Friends_{jk}$ is a binary indicator (1=class allocated to FRIENDS and 0=class allocated to usual practice), all β coefficients are also standardised.

Subgroup analysis (H4 (H4a-b))

The protocol for the ITT analysis was followed. In addition, further models for each hypothesis H4a-b) were constructed and include risk status (binary indicator) as a main effect and a cross-level interaction term. An intervention effect at the subgroup level will be noted if the coefficients associated with the interaction terms noted above are statistically significant. These will subsequently be converted to Hedge's g, as per EEF reporting standards. These models are built up from equation 1 (ITT models), and thus have the following form:

$$y_{t(ijk)} = \beta_{0ijk} + \beta_1 Friends_{jk} + \beta_2 y_{t-1(ijk)} + \beta_3 Friends_{jk} * y_{t-1(ijk)} + \beta_4 subgroup_{ijk} + \beta_5 Friends_{jk} * subgroup_{ijk} \quad \textbf{Eq. 5}$$

Equation 4 omits the random part of the model, since it remains unchanged from equation 1. β_1 is the effect of the FRIENDS intervention. β_2 is the effect of the standardised outcome (primary or secondary) at pre-test (time "t-1") of the i-th pupil in the j-th class of the k-th school $y_{t-1(ijk)}$. β_3 is the interaction effect between allocation to FRIENDS and the standardised outcome at pre-test. β_4 is the effect of the subgroup of interest (FSM eligibility or raised SDQ scores). β_5 is the interaction effect between allocation to FRIENDS and the subgroup. Given that allocation was done at the class-level and the lagged outcomes and risk factors were measured at the pupil-level, β_3 and β_5 are cross-level interaction effects. Similar to equation 3, the expected value of the outcome for a child at risk or FSM eligible in a class allocated to FRIENDS ($\hat{r} = [\hat{y}|Friends = 1]$) would be:

$$E(\hat{r}) = \beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 \quad \textbf{Eq. 6}$$

Meanwhile, the expected value for a child at risk or FSM eligible in a class allocated to usual practice ($\hat{r} = [\hat{y}|Friends = 0]$) would be as follows:

$$E(\hat{r}) = \beta_0 + \beta_2 + \beta_4 \quad \textbf{Eq. 7}$$

Finally, the expected value for a child not at risk (nor FSM eligible) in a class allocated to FRIENDS would be equation 3 and the expectation for a child not at risk (nor FSM eligible) in a class allocated to usual practice would be equation 4.

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

This analysis was carried out on the FRIENDS trial group only. This model uses scores for dosage, fidelity/adherence, quality, reach, and participant responsiveness from observations.

This was conducted through the construction of 3-level (project officer, class, pupil) hierarchical models[28] to account for nested nature of dataset using MLWin Version 2.36. The algebraic form of this model is akin to the previous models (ITT and Subgroup).

$$\begin{aligned} y_{t(ijk)} &= \beta_{0ijk} + \beta_1 y_{t-1(ijk)} + \beta_2 fidelity_k + \beta_3 quality_k \\ &\quad + \beta_4 engagement_k + \beta_5 reach_k \\ \beta_{0ijk} &= \beta_0 + v_{00k} + u_{0jk} + e_{ijk} \\ where&: \\ v_{00k} &\sim N(0, \sigma_v^2) \\ u_{0jk} &\sim N(0, \sigma_u^2) \\ e_{ijk} &\sim N(0, \sigma_e^2) \end{aligned} \quad \textbf{Eq. 8}$$

Equation 8 is a standard 3-level model with random intercepts for project officers (v_{00k}), classes (u_{0jk}) and pupils (e_{ijk}). As in previous equations, $y_{t(ijk)}$ represents combined KS2 scores and $y_{t-1(ijk)}$ represents standardised KS1 average point scores.

Implementer characteristics

This analysis was carried out on the FRIENDS trial group only. This model uses scores for emotional self-efficacy (emself), teaching self-efficacy (teself) and views on social and emotional learning (views). As with the previous model, this was conducted through the construction of 3-level (project officer, class, pupil) hierarchical models to account for nested nature of dataset using MLWin Version 2.36. The algebraic form of this model is akin to equation 8.

$$y_{t(ijk)} = \beta_{0ijk} + \beta_1 y_{t-1(ijk)} + \beta_2 emself_k + \beta_3 teself_k + \beta_4 views_k \quad \textbf{Eq. 9}$$

Equation 9 has the same structure than equation 8 as it is a standard 3-level model with random intercepts for project officers (v_{00k}), classes (u_{0jk}) and pupils (e_{ijk}). The random part was omitted as it follows the same form of equation 8. Also, as in previous equations, $y_{t(ijk)}$ represents combined KS2 scores and $y_{t-1(ijk)}$ represents standardised KS1 average point scores.

Effect size calculation

In all cases, effect sizes are reported using Hedge's g (Cohen's d bias corrected) as per EEF specifications (Tymms, 2004) and 95% confidence intervals for the effect sizes were estimated as

[28] An alternative parameterisation was attempted in which schools were specified as fixed effects (instead of random effects, i.e. a second higher level above class) and robust standard errors were estimated. For the sake of parsimony, the 3-level specification was kept.

described by Fritz et al. (2012). Following the notation provided above in equations 1-7, effect size estimates are calculated as follows for binary indicators (e.g. FRIENDS vs usual practice):

$$g\big(Friends_{ijk} = 1\big) = \frac{\beta_1}{\sigma_e} \quad \textbf{Eq. 10}$$

While for standardised continuous predictors (e.g. the lagged outcome), the effect size is calculated as follows:

$$g\big(y_{t-1(ijk)}\big) = \frac{2 * \beta_2}{\sigma_e} \quad \textbf{Eq. 11}$$

In equation 10, the ES estimate is effectively a comparison between pupils in classes allocated to FRIENDS and pupils in usual practice classes. Meanwhile, equation 11 provides a slightly different comparison, as there are no two distinguishable groups to compare; in this case, the comparison is between pupils who scores one standard deviation below the mean of the lagged outcome and pupils who scored one standard deviation above the mean of the lagged outcome.

Confidence intervals for ES estimates can be calculated based on their theoretical sampling variance as follows:

$$s_g^2 = \frac{n_a + n_b}{n_a n_b} + \frac{g^2}{2(n_a + n_b)} \quad \textbf{Eq. 12}$$

$$and\ then:$$

$$95\%CI(g) = g \pm 1.96 s_g$$

In equation 12, n_a and n_b are the sample sizes of the groups being compared, g is the estimated effect size (equations 8 and 9), s_g^2 and s_g are the sampling variance and standard deviation (respectively) of the estimated effect size, and 1.96 is the z-score for the 97.5 percentile of the standard normal distribution (and -1.96 is the z-score for the 2.5 percentile).

For the case of models for H3d and H3e (equation 2), the effect size estimates correspond to the Odds ratios (OR), which are obtained from taking the exponential of the estimated coefficients (logit). Confidence intervals are obtained in the same way, by taking the exponential of the OR confidence intervals.

For ES estimations (equations 10 and 11), the denominator is the standard deviation at the pupil level (σ_e) of the models after controlling for covariates. The ITT (Equation 1), Subgroup (Equation 5) and Implementation models are nested, which means that each simpler model is a restricted form of a more complex model. The ITT model contains fewer variables than the Subgroup analysis, which in turn contains fewer variables than the implementation model. One consequence of this is that goodness of fit measures between models can be compared directly using the likelihood ratio test. Nevertheless effect size estimates are not directly comparable between models, because the models control for a different number of covariates and hence the (conditional) pupil-level standard deviation (σ_e) will be different for each model.

Appendix L: Comparison between multilevel models fitted with complete cases and multiply-imputed datasets

KS2 Combined (complete cases)

Empty model							
Variable*	**Coef.**	**Std. Err.**	**p**	**ES****	**var(ES)**	**CI 95% (ES)****	
Intercept	-0.01713	0.04149	0.6798	--	--	--	--
ITT analysis							
Intercept	-0.013	0.039	0.743	--	--	--	--
KS1 score	**0.768**	**0.013**	**0.000**	**2.496**	**0.001**	**2.424**	**2.569**
FRIENDS	0.004	0.037	0.908	0.007	0.001	-0.065	0.078
ITT analysis with minimisation							
Intercept	0.031	0.062	0.621	--	--	--	--
KS1 score	**0.773**	**0.013**	**0.000**	**2.510**	**0.001**	**2.437**	**2.583**
FRIENDS	0.020	0.035	0.575	0.032	0.001	-0.039	0.104
factor(minaps)2	-0.108	0.065	0.096	-0.175	0.001	-0.250	-0.101
factor(minaps)3	**-0.208**	**0.073**	**0.004**	**-0.337**	**0.002**	**-0.415**	**-0.259**
factor(minrcads)2	**0.099**	**0.050**	**0.048**	**0.161**	**0.001**	**0.087**	**0.235**
factor(minrcads)3	0.054	0.053	0.307	0.088	0.002	0.008	0.168
Subgroup analysis (FSM)							
Intercept	0.037	0.041	0.357	--	--	--	--
KS1 score	**0.719**	**0.018**	**0.000**	**2.352**	**0.001**	**2.283**	**2.422**
FRIENDS	-0.009	0.040	0.827	-0.014	0.001	-0.086	0.057
KS1*FRIENDS	**0.071**	**0.025**	**0.005**	**0.116**	**0.001**	**0.045**	**0.188**
FSM eligible	**-0.169**	**0.039**	**0.000**	**-0.276**	**0.002**	**-0.357**	**-0.195**
FRIENDS*FSM	0.031	0.057	0.587	0.050	0.003	-0.058	0.159
Subgroup analysis (SDQ internalising)							
Intercept	0.001	0.040	0.984	--	--	--	--
KS1 score	**0.720**	**0.018**	**0.000**	**2.358**	**0.001**	**2.288**	**2.427**
FRIENDS	0.010	0.038	0.802	0.016	0.001	-0.056	0.087
KS1*FRIENDS	**0.076**	**0.026**	**0.003**	**0.125**	**0.001**	**0.054**	**0.197**
raised SDQ internalising	**-0.120**	**0.059**	**0.042**	**-0.197**	**0.004**	**-0.319**	**-0.075**
raised SDQ*FRIENDS	-0.002	0.084	0.979	-0.004	0.007	-0.172	0.165

Notes: p-values in bold are significant at 0.05. These estimates are the results from complete-case analysis.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2; non-raised SDQ (internalising).

** Effect sizes are estimated using formulae described by Tymms (2004). 95% confidence intervals for effect sizes are estimated as described by Fritz et al. (2012)

KS2 Combined (multiple imputation)

Empty model							
Variable*	**Coef.**	**Std. Err.**	**p**	**ES****	**var(ES)**	**CI 95% (ES)****	
Intercept	-0.03	0.04334	0.49	--	--	--	--
ITT analysis							
Intercept	-0.016	0.041	0.691	--	--	--	--
KS1 score	**0.759**	**0.013**	**0.000**	**2.419**	**0.001**	**2.348**	**2.490**
FRIENDS	0.004	0.038	0.909	0.007	0.001	-0.065	0.078
ITT analysis with minimisation							
Intercept	0.025	0.065	0.697	--	--	--	--
KS1 score	**0.763**	**0.013**	**0.000**	**2.432**	**0.001**	**2.360**	**2.503**
FRIENDS	0.018	0.037	0.631	0.028	0.001	-0.043	0.099
factor(minaps)2	-0.096	0.068	0.160	-0.153	0.001	-0.227	-0.079
factor(minaps)3	**-0.196**	**0.075**	**0.009**	**-0.312**	**0.002**	**-0.390**	**-0.234**
factor(minrcads)2	0.090	0.052	0.082	0.143	0.001	0.069	0.217
factor(minrcads)3	0.044	0.054	0.422	0.069	0.002	-0.011	0.149
Subgroup analysis (FSM)							
Intercept	0.032	0.043	0.446	--	--	--	--
KS1 score	**0.720**	**0.019**	**0.000**	**2.308**	**0.001**	**2.239**	**2.376**
FRIENDS	-0.006	0.041	0.887	-0.009	0.001	-0.081	0.062
KS1*FRIENDS	0.050	0.027	0.061	0.080	0.001	0.009	0.152
FSM eligible	**-0.166**	**0.041**	**0.000**	**-0.266**	**0.002**	**-0.347**	**-0.185**
FRIENDS*FSM	0.026	0.059	0.656	0.042	0.003	-0.066	0.151
Subgroup analysis (SDQ internalising)							
Intercept	-0.005	0.041	0.904	--	--	--	--
KS1 score	**0.731**	**0.018**	**0.000**	**2.335**	**0.001**	**2.267**	**2.404**
FRIENDS	0.008	0.039	0.845	0.012	0.001	-0.059	0.084
KS1*FRIENDS	0.049	0.026	0.064	0.077	0.001	0.006	0.149
raised SDQ internalising	-0.101	0.059	0.089	-0.161	0.004	-0.283	-0.039
raised SDQ*FRIENDS	-0.025	0.090	0.783	-0.039	0.007	-0.208	0.129

Notes: p-values in bold are significant at 0.05. These estimates are the pooled results from 10 multiply-imputed datasets.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2; non-raised SDQ (internalising).

** Effect sizes are estimated using formulae described by Tymms (2004). 95% confidence intervals for effect sizes are estimated as described by Fritz et al. (2012)

KS2 Maths (complete cases)

Empty model							
Variable*	**Coef.**	**Std. Err.**	**p**	**ES****	**var(ES)**	**CI 95% (ES)****	
Intercept	-0.03288	0.04447	0.4597	--	--	--	--
ITT analysis							
Intercept	-0.053	0.045	0.239	--	--	--	--
KS1 score	**0.723**	**0.014**	**0.000**	**2.167**	**0.001**	**2.102**	**2.233**
FRIENDS	0.002	0.039	0.953	0.003	0.001	-0.068	0.075
ITT analysis with minimisation							
Intercept	-0.030	0.069	0.668	--	--	--	--
KS1 score	**0.726**	**0.014**	**0.000**	**2.178**	**0.001**	**2.113**	**2.244**
FRIENDS	0.024	0.035	0.498	0.036	0.001	-0.036	0.107
factor(minaps)2	-0.108	0.072	0.132	-0.162	0.001	-0.236	-0.088
factor(minaps)3	**-0.232**	**0.082**	**0.005**	**-0.348**	**0.002**	**-0.426**	**-0.270**
factor(minrcads)2	**0.157**	**0.051**	**0.002**	**0.235**	**0.001**	**0.161**	**0.309**
factor(minrcads)3	0.062	0.055	0.256	0.093	0.002	0.013	0.173
Subgroup analysis (FSM)							
Intercept	-0.004	0.047	0.933	--	--	--	--
KS1 score	**0.681**	**0.020**	**0.000**	**2.052**	**0.001**	**1.989**	**2.115**
FRIENDS	-0.014	0.043	0.743	-0.021	0.001	-0.093	0.050
KS1*FRIENDS	**0.059**	**0.029**	**0.039**	**0.089**	**0.001**	**0.018**	**0.161**
FSM eligible	**-0.164**	**0.043**	**0.000**	**-0.247**	**0.002**	**-0.328**	**-0.166**
FRIENDS*FSM	0.038	0.062	0.544	0.057	0.003	-0.052	0.165
Subgroup analysis (SDQ internalising)							
Intercept	-0.030	0.046	0.515	--	--	--	--
KS1 score	**0.683**	**0.020**	**0.000**	**2.049**	**0.001**	**1.986**	**2.112**
FRIENDS	-0.001	0.042	0.989	-0.001	0.001	-0.072	0.071
KS1*FRIENDS	**0.072**	**0.029**	**0.013**	**0.108**	**0.001**	**0.036**	**0.179**
SDQ raised (internalising)	**-0.227**	**0.065**	**0.000**	**-0.340**	**0.004**	**-0.463**	**-0.218**
FRIENDS*raised SDQ	0.059	0.093	0.522	0.089	0.007	-0.080	0.258

Notes: p-values in bold are significant at 0.05. These estimates are the results from complete-case analysis.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2; non-raised SDQ (internalising).

** Effect sizes are estimated using formulae described by Tymms (2004). 95% confidence intervals for effect sizes are estimated as described by Fritz et al. (2012)

KS2 Maths (multiple imputation)

Empty model							
Variable*	**Coef.**	**Std. Err.**	**p**	**ES****	**var(ES)**	**CI 95% (ES)****	
Intercept	-0.063	0.04701	0.179	--	--	--	--
ITT analysis							
Intercept	-0.052	0.045	0.251	--	--	--	--
KS1 score	**0.716**	**0.014**	**0.000**	**2.115**	**0.001**	**2.051**	**2.179**
FRIENDS	0.001	0.040	0.970	0.002	0.001	-0.069	0.074
ITT analysis with minimisation							
Intercept	-0.034	0.070	0.627	--	--	--	--
KS1 score	**0.719**	**0.015**	**0.000**	**2.125**	**0.001**	**2.061**	**2.190**
FRIENDS	0.021	0.037	0.564	0.031	0.001	-0.040	0.103
factor(minaps)2	-0.098	0.072	0.173	-0.145	0.001	-0.219	-0.071
factor(minaps)3	**-0.215**	**0.082**	**0.009**	**-0.318**	**0.002**	**-0.395**	**-0.240**
factor(minrcads)2	**0.149**	**0.053**	**0.005**	**0.221**	**0.001**	**0.147**	**0.295**
factor(minrcads)3	0.054	0.055	0.331	0.079	0.002	-0.001	0.159
Subgroup analysis (FSM)							
Intercept	-0.006	0.047	0.895	--	--	--	--
KS1 score	**0.683**	**0.020**	**0.000**	**2.027**	**0.001**	**1.965**	**2.090**
FRIENDS	-0.013	0.043	0.767	-0.019	0.001	-0.090	0.053
KS1*FRIENDS	0.039	0.030	0.189	0.058	0.001	-0.014	0.129
FSM eligible	**-0.157**	**0.043**	**0.000**	**-0.233**	**0.002**	**-0.314**	**-0.152**
FRIENDS*FSM	0.039	0.062	0.528	0.058	0.003	-0.051	0.167
Subgroup analysis (SDQ internalising)							
Intercept	-0.031	0.046	0.496	--	--	--	--
KS1 score	**0.691**	**0.019**	**0.000**	**2.048**	**0.001**	**1.985**	**2.111**
FRIENDS	0.001	0.041	0.974	0.002	0.001	-0.069	0.073
KS1*FRIENDS	0.038	0.029	0.192	0.056	0.001	-0.016	0.127
SDQ raised (internalising)	**-0.199**	**0.069**	**0.004**	**-0.295**	**0.004**	**-0.417**	**-0.173**
FRIENDS*raised SDQ	0.021	0.092	0.822	0.031	0.007	-0.138	0.199

Notes: p-values in bold are significant at 0.05. These estimates are the pooled results from 10 multiply-imputed datasets.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2; non-raised SDQ (internalising).

** Effect sizes are estimated using formulae described by Tymms (2004). 95% confidence intervals for effect sizes are estimated as described by Fritz et al. (2012)

KS2 Reading (complete cases)

Empty model							
Variable*	**Coef.**	**Std. Err.**	**p**	**ES****	**var(ES)**	**CI 95% (ES)****	
Intercept	-0.00525	0.03891	0.8926	--	--	--	--
ITT analysis							
Intercept	-0.015	0.039	0.708	--	--	--	--
KS1 score	**0.718**	**0.015**	**0.000**	**2.068**	**0.001**	**2.005**	**2.132**
FRIENDS	-0.009	0.041	0.830	-0.013	0.001	-0.084	0.059
ITT analysis with minimisation							
Intercept	0.032	0.065	0.622	--	--	--	--
KS1 score	**0.723**	**0.015**	**0.000**	**2.082**	**0.001**	**2.019**	**2.146**
FRIENDS	0.002	0.042	0.961	0.003	0.001	-0.069	0.074
factor(minaps)2	-0.119	0.067	0.076	-0.171	0.001	-0.246	-0.097
factor(minaps)3	**-0.169**	**0.074**	**0.022**	**-0.244**	**0.002**	**-0.322**	**-0.167**
factor(minrcads)2	**0.074**	**0.056**	**0.191**	**0.106**	**0.001**	**0.032**	**0.180**
factor(minrcads)3	0.059	0.060	0.328	0.084	0.002	0.004	0.164
Subgroup analysis (FSM)							
Intercept	0.029	0.041	0.482	--	--	--	--
KS1 score	**0.669**	**0.021**	**0.000**	**1.937**	**0.001**	**1.877**	**1.998**
FRIENDS	-0.011	0.045	0.799	-0.016	0.001	-0.088	0.055
KS1*FRIENDS	**0.072**	**0.030**	**0.017**	**0.104**	**0.001**	**0.032**	**0.175**
FSM eligible	**-0.142**	**0.045**	**0.001**	**-0.206**	**0.002**	**-0.287**	**-0.125**
FRIENDS*FSM	-0.012	0.065	0.851	-0.018	0.003	-0.126	0.091
Subgroup analysis (SDQ internalising)							
Intercept	-0.009	0.040	0.824	--	--	--	--
KS1 score	**0.678**	**0.021**	**0.000**	**1.950**	**0.001**	**1.889**	**2.011**
FRIENDS	-0.013	0.044	0.759	-0.019	0.001	-0.091	0.052
KS1*FRIENDS	**0.076**	**0.030**	**0.012**	**0.109**	**0.001**	**0.038**	**0.181**
SDQ raised (internalising)	-0.043	0.068	0.523	-0.063	0.004	-0.184	0.059
FRIENDS*raised SDQ	-0.002	0.097	0.985	-0.003	0.007	-0.171	0.166

Notes: p-values in bold are significant at 0.05. These estimates are the results from complete-case analysis.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2; non-raised SDQ (internalising).

** Effect sizes are estimated using formulae described by Tymms (2004). 95% confidence intervals for effect sizes are estimated as described by Fritz et al. (2012)

KS2 Reading (multiple imputation)

Empty model							
Variable*	**Coef.**	**Std. Err.**	**p**	**ES****	**var(ES)**	**CI 95% (ES)****	
Intercept	-0.035	-0.035	-0.03	--	--	--	--
ITT analysis							
Intercept	-0.015	0.040	0.698	--	--	--	--
KS1 score	**0.709**	**0.016**	**0.000**	**2.003**	**0.001**	**1.941**	**2.065**
FRIENDS	-0.011	0.043	0.802	-0.015	0.001	-0.087	0.056
ITT analysis with minimisation							
Intercept	0.028	0.067	0.674	--	--	--	--
KS1 score	**0.714**	**0.016**	**0.000**	**2.016**	**0.001**	**1.954**	**2.078**
FRIENDS	-0.001	0.044	0.979	-0.002	0.001	-0.073	0.070
factor(minaps)2	-0.109	0.071	0.124	-0.154	0.001	-0.228	-0.080
factor(minaps)3	**-0.165**	**0.077**	**0.032**	**-0.233**	**0.002**	**-0.310**	**-0.155**
factor(minrcads)2	0.071	0.061	0.248	0.100	0.001	0.026	0.174
factor(minrcads)3	0.052	0.062	0.397	0.074	0.002	-0.006	0.154
Subgroup analysis (FSM)							
Intercept	0.026	0.042	0.531	--	--	--	--
KS1 score	**0.666**	**0.021**	**0.000**	**1.889**	**0.001**	**1.830**	**1.949**
FRIENDS	-0.012	0.046	0.787	-0.018	0.001	-0.089	0.054
KS1*FRIENDS	**0.059**	**0.030**	**0.049**	**0.084**	**0.001**	**0.012**	**0.155**
FSM eligible	**-0.143**	**0.045**	**0.001**	**-0.202**	**0.002**	**-0.283**	**-0.121**
FRIENDS*FSM	-0.002	0.064	0.976	-0.003	0.003	-0.111	0.106
Subgroup analysis (SDQ internalising)							
Intercept	-0.010	0.040	0.795	--	--	--	--
KS1 score	**0.677**	**0.021**	**0.000**	**1.915**	**0.001**	**1.855**	**1.976**
FRIENDS	-0.011	0.044	0.810	-0.015	0.001	-0.086	0.057
KS1*FRIENDS	**0.062**	**0.030**	**0.038**	**0.087**	**0.001**	**0.016**	**0.159**
SDQ raised (internalising)	-0.046	0.073	0.525	-0.066	0.004	-0.187	0.056
FRIENDS*raised SDQ	0.001	0.099	0.991	0.002	0.007	-0.167	0.170

Notes: p-values in bold are significant at 0.05. These estimates are the pooled results from 10 multiply-imputed datasets.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2; non-raised SDQ (internalising).

** Effect sizes are estimated using formulae described by Tymms (2004). 95% confidence intervals for effect sizes are estimated as described by Fritz et al. (2012)

PSW (complete cases)

Empty model							
Variable*	**Coef.**	**Std. Err.**	**p**	**ES****	**var(ES)**	**CI 95% (ES)****	
Intercept	-0.0003	0.029	0.992	--	--	--	--
ITT analysis							
Intercept	-0.026	0.037	0.484	--	--	--	--
FRIENDS	0.051	0.047	0.269	0.052	0.001	-0.019	0.124
ITT analysis with minimisation							
Intercept	0.132	0.058	0.023	--	--	--	--
FRIENDS	0.029	0.042	0.489	0.030	0.001	-0.042	0.101
factor(minaps)2	-0.009	0.059	0.881	-0.009	0.001	-0.083	0.065
factor(minaps)3	0.005	0.065	0.943	0.005	0.002	-0.073	0.082
factor(minrcads)2	**-0.177**	**0.054**	**0.001**	**-0.180**	**0.001**	**-0.254**	**-0.106**
factor(minrcads)3	**-0.269**	**0.058**	**0.000**	**-0.274**	**0.002**	**-0.354**	**-0.194**
Subgroup analysis (FSM)							
Intercept	-0.017	0.042	0.690	--	--	--	--
FRIENDS	0.018	0.053	0.733	0.018	0.001	-0.053	0.090
FSM eligible	-0.035	0.064	0.585	-0.036	0.002	-0.117	0.045
FRIENDS*FSM	0.144	0.090	0.109	0.148	0.003	0.039	0.256
Subgroup analysis (SDQ internalising)							
Intercept	-0.073	0.038	0.058	--	--	--	--
FRIENDS	0.058	0.051	0.254	0.060	0.002	-0.036	0.155
SDQ raised (internalising)	**0.533**	**0.100**	**0.000**	**0.546**	**0.004**	**0.423**	**0.668**
FRIENDS*SDQ internalising	-0.088	0.141	0.531	-0.090	0.007	-0.259	0.078

Notes: p-values in bold are significant at 0.05. These estimates are the results from complete-case analysis.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2; non-raised SDQ (internalising).

** Effect sizes are estimated using formulae described by Tymms (2004). 95% confidence intervals for effect sizes are estimated as described by Fritz et al. (2012)

PSW (multiple imputation)

Empty model							
Variable*	**Coef.**	**Std. Err.**	**p**	**ES****	**var(ES)**	**CI 95% (ES)****	
Intercept	0.005	0.030	0.873	--	--	--	--
ITT analysis							
Intercept	-0.024	0.037	0.514	--	--	--	--
FRIENDS	0.059	0.047	0.210	0.059	0.001	-0.012	0.131
ITT analysis with minimisation							
Intercept	0.129	0.059	0.029	--	--	--	--
FRIENDS	0.043	0.043	0.320	0.043	0.001	-0.028	0.115
factor(minaps)2	-0.0004	0.062	0.995	-0.0004	0.001	-0.075	0.074
factor(minaps)3	-0.008	0.069	0.909	-0.008	0.002	-0.085	0.069
factor(minrcads)2	**-0.179**	**0.057**	**0.002**	**-0.182**	**0.001**	**-0.256**	**-0.108**
factor(minrcads)3	**-0.265**	**0.058**	**0.000**	**-0.269**	**0.002**	**-0.349**	**-0.189**
Subgroup analysis (FSM)							
Intercept	-0.015	0.043	0.730	--	--	--	--
FRIENDS	0.024	0.054	0.661	0.024	0.001	-0.047	0.095
FSM eligible	-0.032	0.068	0.635	-0.033	0.002	-0.113	0.048
FRIENDS*FSM	0.125	0.090	0.162	0.127	0.003	0.018	0.236
Subgroup analysis (SDQ internalising)							
Intercept	-0.075	0.038	0.048	--	--	--	--
FRIENDS	0.056	0.049	0.246	0.058	0.002	-0.038	0.154
SDQ raised (internalising)	**0.496**	**0.100**	**0.000**	**0.509**	**0.004**	**0.386**	**0.631**
FRIENDS*SDQ internalising	-0.030	0.138	0.829	-0.031	0.007	-0.199	0.138

Notes: p-values in bold are significant at 0.05. These estimates are the pooled results from 10 multiply-imputed datasets.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2; non-raised SDQ (internalising).

** Effect sizes are estimated using formulae described by Tymms (2004). 95% confidence intervals for effect sizes are estimated as described by Fritz et al. (2012)

RCADS (complete cases)

Empty model							
Variable*	**Coef.**	**Std. Err.**	**p**	**ES****	**var(ES)**	**CI 95% (ES)****	
Intercept	-0.004	0.031	0.895	--	--	--	--
ITT analysis							
Intercept	-0.005	0.042	0.911	--	--	--	--
FRIENDS	0.001	0.056	0.983	0.001	0.001	-0.070	0.073
ITT analysis with minimisation							
Intercept	0.200	0.063	0.001	--	--	--	--
FRIENDS	-0.022	0.052	0.671	-0.023	0.001	-0.094	0.049
factor(minaps)2	-0.017	0.062	0.786	-0.017	0.001	-0.092	0.057
factor(minaps)3	-0.032	0.066	0.624	-0.033	0.002	-0.110	0.044
factor(minrcads)2	**-0.183**	**0.061**	**0.003**	**-0.188**	**0.001**	**-0.263**	**-0.114**
factor(minrcads)3	**-0.376**	**0.066**	**0.000**	**-0.388**	**0.002**	**-0.468**	**-0.307**
Subgroup analysis (FSM)							
Intercept	0.007	0.046	0.882	--	--	--	--
FRIENDS	-0.060	0.062	0.336	-0.061	0.001	-0.133	0.010
FSM eligible	-0.031	0.064	0.621	-0.032	0.002	-0.113	0.048
FRIENDS*FSM	**0.237**	**0.090**	**0.009**	**0.245**	**0.003**	**0.136**	**0.354**
Subgroup analysis (SDQ internalising)							
Intercept	-0.0424	0.0415	0.306	--	--	--	--
FRIENDS	-0.0116	0.0587	0.843	-0.012	0.003	-0.114	0.090
SDQ raised (internalising)	**0.4363**	**0.0974**	**0.000**	**0.453**	**0.001**	**0.380**	**0.525**
FRIENDS*raised SDQ	0.1848	0.1380	0.181	0.192	0.005	0.056	0.327

Notes: p-values in bold are significant at 0.05. These estimates are the results from complete-case analysis.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2; non-raised SDQ (internalising).

** Effect sizes are estimated using formulae described by Tymms (2004). 95% confidence intervals for effect sizes are estimated as described by Fritz et al. (2012)

RCADS (multiple imputation)

Empty model							
Variable*	**Coef.**	**Std. Err.**	**p**	**ES****	**var(ES)**	**CI 95% (ES)****	
Intercept	0.009	0.030	0.764	--	--	--	--
ITT analysis							
Intercept	0.006	0.041	0.891	--	--	--	--
FRIENDS	0.007	0.056	0.904	0.007	0.001	-0.064	0.078
ITT analysis with minimisation							
Intercept	0.214	0.061	0.000	--	--	--	--
FRIENDS	-0.008	0.051	0.870	-0.009	0.001	-0.080	0.063
factor(minaps)2	-0.023	0.063	0.713	-0.024	0.001	-0.098	0.050
factor(minaps)3	-0.052	0.066	0.429	-0.053	0.002	-0.131	0.024
factor(minrcads)2	**-0.193**	**0.061**	**0.001**	**-0.198**	**0.001**	**-0.272**	**-0.124**
factor(minrcads)3	**-0.372**	**0.064**	**0.000**	**-0.382**	**0.002**	**-0.463**	**-0.302**
Subgroup analysis (FSM)							
Intercept	0.011	0.045	0.816	--	--	--	--
FRIENDS	-0.054	0.061	0.381	-0.055	0.001	-0.127	0.016
FSM eligible	-0.016	0.064	0.804	-0.016	0.002	-0.097	0.064
FRIENDS*FSM	**0.216**	**0.089**	**0.015**	**0.222**	**0.003**	**0.114**	**0.331**
Subgroup analysis (SDQ internalising)							
Intercept	-0.038	0.040	0.341	--	--	--	--
FRIENDS	-0.013	0.056	0.820	-0.013	0.003	-0.115	0.088
SDQ raised (internalising)	0.436	0.091	0.000	0.453	0.001	0.381	0.525
FRIENDS*raised SDQ	0.148	0.140	0.289	0.154	0.005	0.019	0.289

Notes: p-values in bold are significant at 0.05. These estimates are the pooled results from 10 multiply-imputed datasets.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2; non-raised SDQ (internalising).

** Effect sizes are estimated using formulae described by Tymms (2004). 95% confidence intervals for effect sizes are estimated as described by Fritz et al. (2012)

SDQ Internalising (complete cases)

Empty model						
Variable*	**Coef.**	**Std. Err.**	**p**	**OR****	**CI 95% (OR)****	
Intercept	-2.319	0.119	0.000	--	--	--
ITT analysis						
Intercept	-2.291	0.153	0.000	--	--	--
FRIENDS	-0.058	0.202	0.773	0.943	0.635	1.402
ITT analysis with minimisation						
Intercept	-1.893	0.248	0.000	--	--	--
FRIENDS	-0.046	0.202	0.821	0.955	0.643	1.419
factor(minaps)2	-0.329	0.254	0.197	0.720	0.437	1.186
factor(minaps)3	**-0.654**	**0.294**	**0.026**	**0.520**	**0.292**	**0.925**
factor(minrcads)2	-0.162	0.253	0.524	0.851	0.518	1.398
factor(minrcads)3	-0.090	0.270	0.740	0.914	0.539	1.552
Subgroup analysis (FSM)						
Intercept	-2.535	0.170	0.000	--	--	--
FRIENDS	0.006	0.231	0.978	1.006	0.639	1.584
FSM eligible	**0.827**	**0.196**	**0.000**	**2.286**	**1.555**	**3.359**
FRIENDS*FSM	-0.157	0.288	0.587	0.855	0.486	1.505
Subgroup analysis (SDQ internalising)						
Intercept	-2.363	0.164	0.000	--	--	--
FRIENDS	0.021	0.238	0.928	1.022	0.641	1.628
raised SDQ internalising	**2.151**	**0.231**	**0.000**	**8.596**	**5.468**	**13.513**
FRIENDS*raised SDQ	-0.392	0.335	0.241	0.676	0.351	1.302

Notes: p-values in bold are significant at 0.05. These estimates are the results from complete-case analysis.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2.

** Odds ratio is a conventional measure of effect size for logistic regression models. However, it is not directly comparable with other effect size measures.

SDQ internalising (multiple imputation)

Empty model						
Variable*	**Coef.**	**Std. Err.**	**p**	**OR****	**CI 95% (OR)****	
Intercept	-2.325	0.120	0.000	--	--	--
ITT analysis						
Intercept	-2.298	0.155	0.000	--	--	--
FRIENDS	-0.056	0.198	0.776	0.945	0.641	1.394
ITT analysis with minimisation						
Intercept	-1.846	0.237	0.000	--	--	--
FRIENDS	-0.054	0.197	0.785	0.948	0.645	1.393
factor(minaps)2	-0.366	0.246	0.138	0.694	0.428	1.124
factor(minaps)3	**-0.699**	**0.292**	**0.017**	**0.497**	**0.280**	**0.882**
factor(minrcads)2	-0.160	0.247	0.518	0.852	0.525	1.384
factor(minrcads)3	-0.119	0.265	0.652	0.887	0.528	1.492
Subgroup analysis (FSM)						
Intercept	-2.265	0.164	0.000	--	--	--
FRIENDS	0.040	0.213	0.853	1.040	0.685	1.579
FSM eligible	**0.709**	**0.187**	**0.000**	**2.032**	**1.408**	**2.932**
FRIENDS*FSM	-0.151	0.265	0.568	0.860	0.512	1.444
Subgroup analysis (SDQ internalising)						
Intercept	-2.349	0.165	0.000	--	--	--
FRIENDS	0.019	0.223	0.933	1.019	0.658	1.578
raised SDQ internalising	**1.982**	**0.233**	**0.000**	**7.261**	**4.597**	**11.468**
FRIENDS*raised SDQ	-0.350	0.339	0.302	0.705	0.363	1.370

Notes: p-values in bold are significant at 0.05. These estimates are the pooled results from 10 multiply-imputed datasets.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2.

** Odds ratio is a conventional measure of effect size for logistic regression models. However, it is not directly comparable with other effect size measures.

SDQ Externalising (complete cases)

Empty model						
Variable*	**Coef.**	**Std. Err.**	**p**	**OR****	**CI 95% (OR)****	
Intercept	-2.252	0.117	0.000	--	--	--
ITT analysis						
Intercept	-2.270	0.158	0.000	--	--	--
FRIENDS	0.033	0.218	0.878	1.034	0.674	1.585
ITT analysis with minimisation						
Intercept	-1.844	0.238	0.000	--	--	--
FRIENDS	-0.008	0.206	0.969	0.992	0.663	1.484
factor(minaps)2	-0.008	0.239	0.973	0.992	0.622	1.583
factor(minaps)3	**-0.841**	**0.289**	**0.004**	**0.431**	**0.245**	**0.760**
factor(minrcads)2	-0.330	0.247	0.181	0.719	0.443	1.166
factor(minrcads)3	-0.100	0.259	0.698	0.904	0.544	1.503
Subgroup analysis (FSM)						
Intercept	-2.571	0.178	0.000	--	--	--
FRIENDS	-0.012	0.252	0.962	0.988	0.603	1.620
FSM eligible	**0.920**	**0.199**	**0.000**	**2.510**	**1.699**	**3.708**
FRIENDS*FSM	0.124	0.287	0.665	1.132	0.645	1.987
Subgroup analysis (SDQ internalising)						
Intercept	-2.093	0.147	0.000	--	--	--
FRIENDS	0.075	0.207	0.717	1.078	0.718	1.619
raised SDQ internalising	0.913	0.252	0.000	2.491	1.520	4.081
FRIENDS*raised SDQ	-0.414	0.372	0.266	0.661	0.318	1.371

Notes: p-values in bold are significant at 0.05. These estimates are the results from complete-case analysis.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2.

** Odds ratio is a conventional measure of effect size for logistic regression models. However, it is not directly comparable with other effect size measures.

SDQ Externalising (multiple imputation)

Empty model						
Variable*	**Coef.**	**Std. Err.**	**p**	**OR****	**CI 95% (OR)****	
Intercept	-2.276	0.117	0.000	--	--	--
ITT analysis						
Intercept	-2.281	0.154	0.000	--	--	--
FRIENDS	0.008	0.213	0.970	1.008	0.664	1.531
ITT analysis with minimisation						
Intercept	-1.869	0.255	0.000	--	--	--
FRIENDS	-0.018	0.210	0.932	0.982	0.650	1.483
factor(minaps)2	0.005	0.250	0.984	1.005	0.616	1.640
factor(minaps)3	**-0.827**	**0.272**	**0.002**	**0.437**	**0.257**	**0.745**
factor(minrcads)2	-0.267	0.261	0.307	0.766	0.459	1.278
factor(minrcads)3	-0.077	0.257	0.764	0.926	0.560	1.531
Subgroup analysis (FSM)						
Intercept	-2.266	0.150	0.000	--	--	--
FRIENDS	-0.060	0.217	0.784	0.942	0.616	1.442
FSM eligible	**0.817**	**0.189**	**0.000**	**2.264**	**1.563**	**3.279**
FRIENDS*FSM	0.123	0.263	0.641	1.131	0.675	1.894
Subgroup analysis (SDQ internalising)						
Intercept	-2.090	0.144	0.000	--	--	--
FRIENDS	0.050	0.197	0.800	1.051	0.715	1.545
raised SDQ internalising	**0.835**	**0.237**	**0.000**	**2.304**	**1.449**	**3.665**
FRIENDS*raised SDQ	-0.400	0.362	0.270	0.671	0.330	1.363

Notes: p-values in bold are significant at 0.05. These estimates are the pooled results from 10 multiply-imputed datasets.

* Reference categories: Usual practice; non-FSM eligible; minaps1; minaps2.

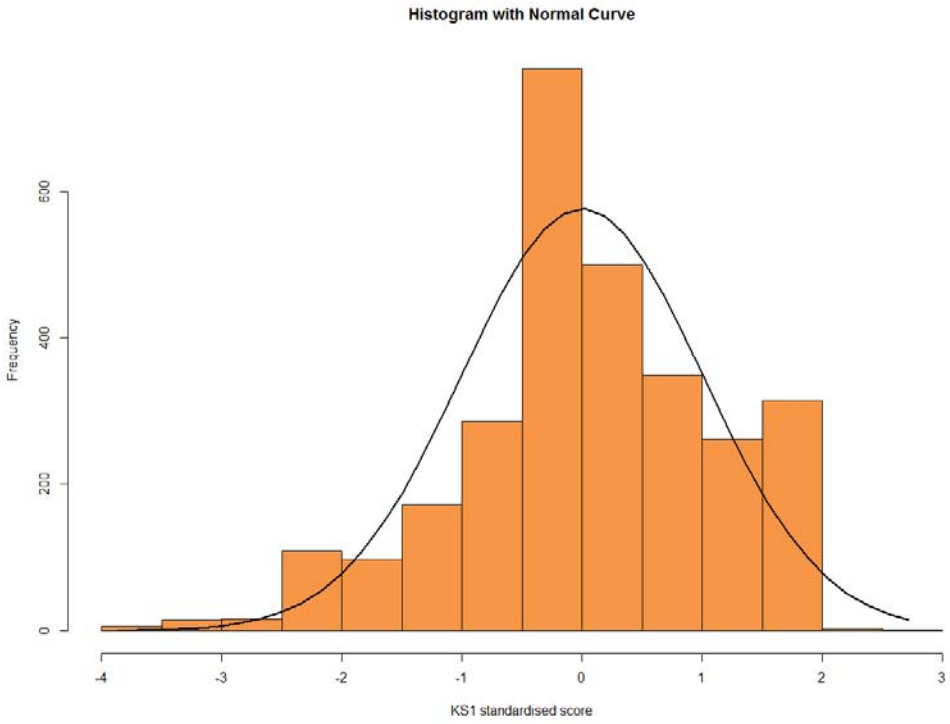
** Odds ratio is a conventional measure of effect size for logistic regression models. However, it is not directly comparable with other effect size measures.

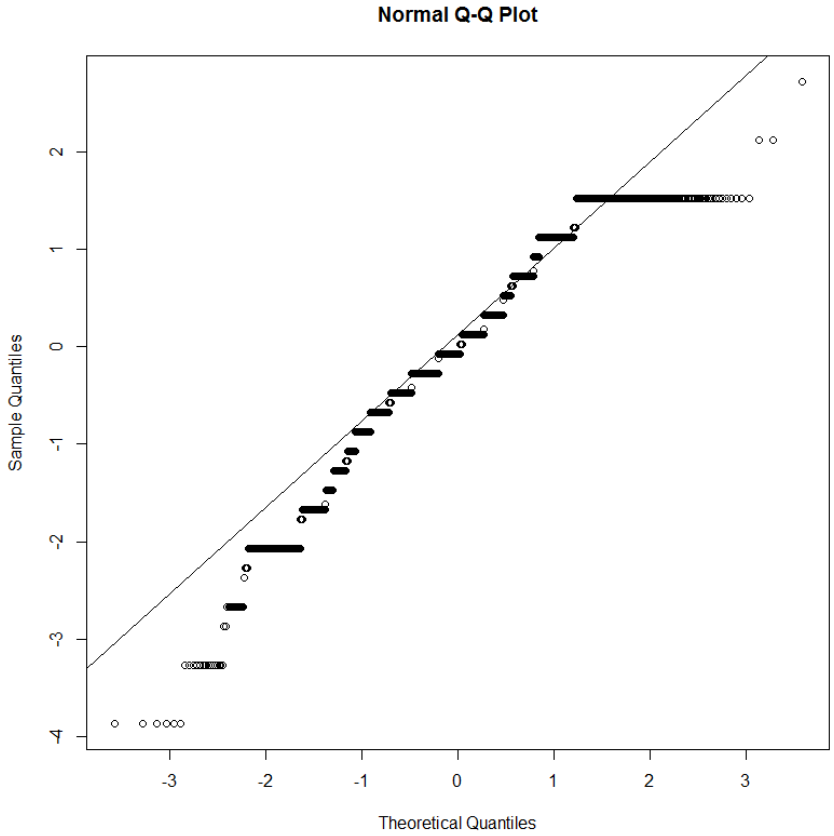
Appendix M: Histograms and QQ plots

As can be seen below, data distribution was mostly normal, with an expected positive skew for measures of mental health (PSW, RCADS).

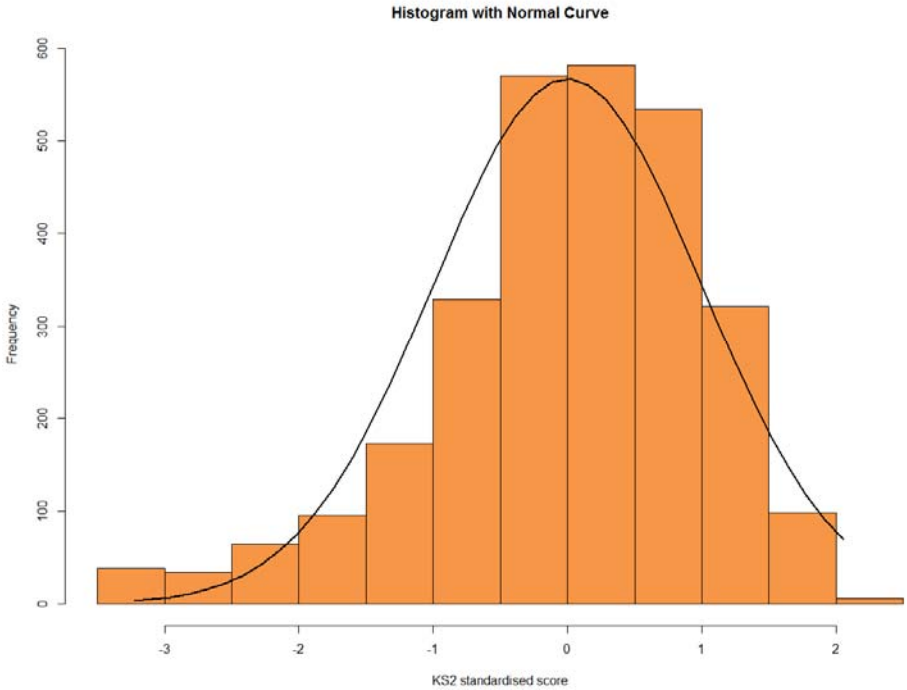
Histograms

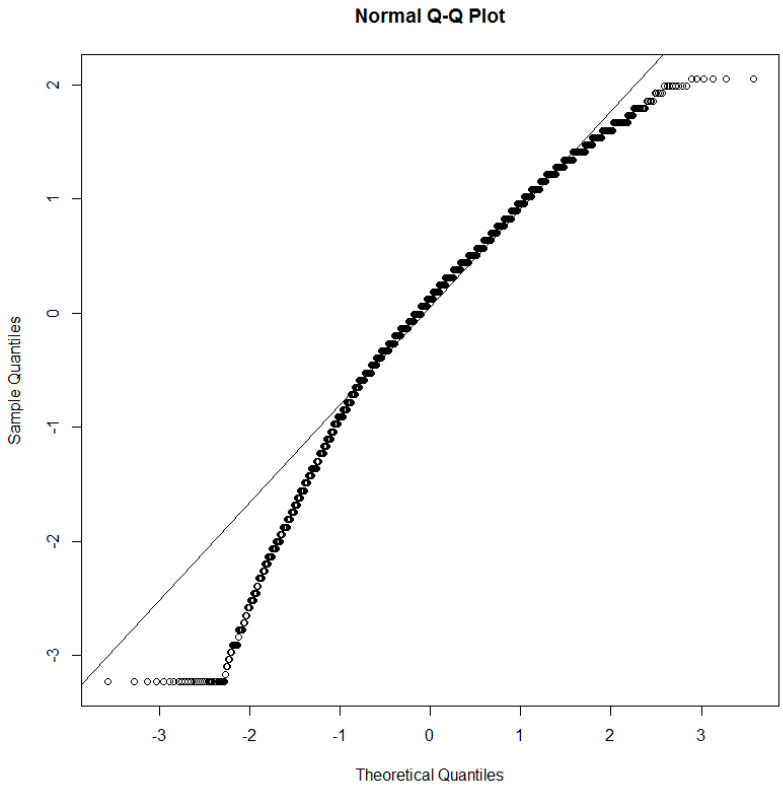
KS1



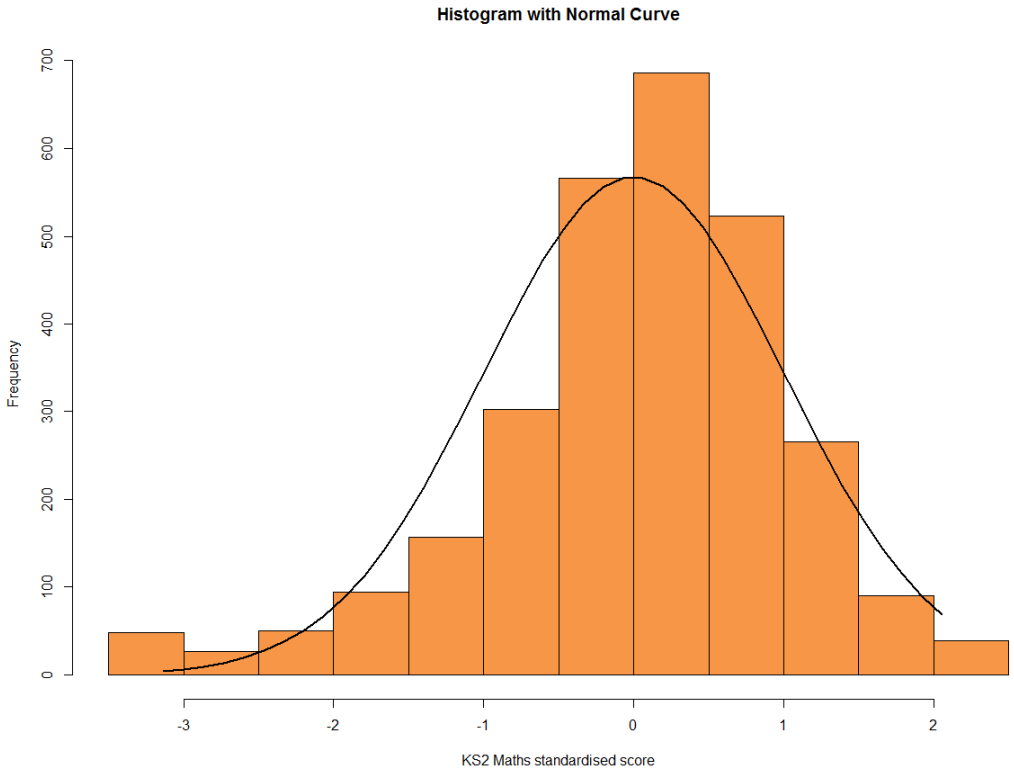

Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
-4
-3
-2
-1
0
1
2
-3
-2
-1
0
1
2
3

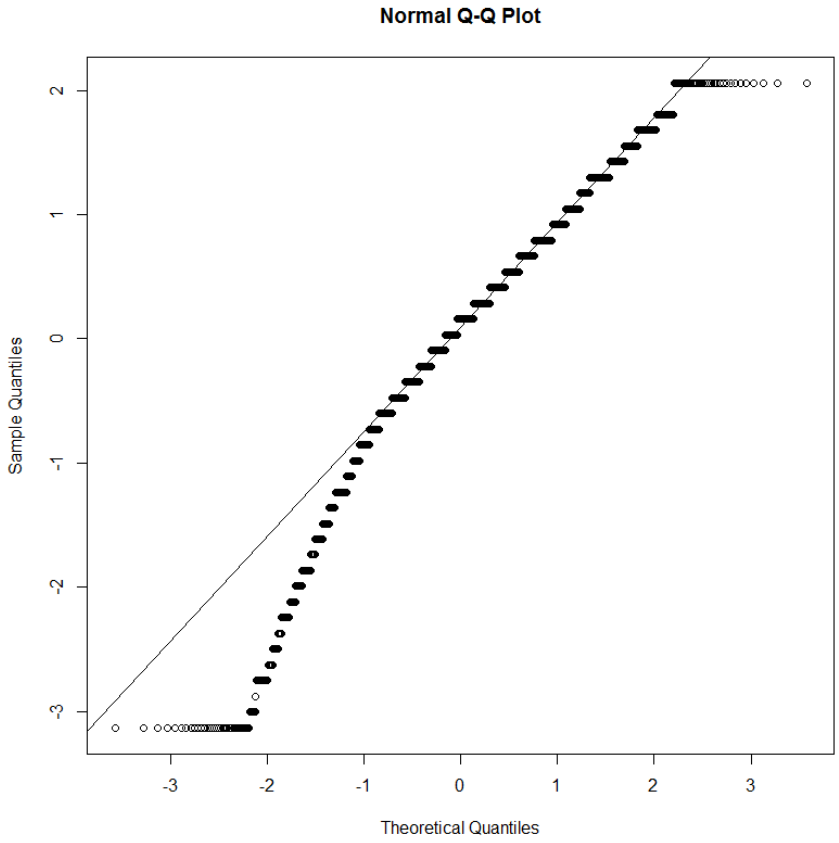
KS2


Histogram with Normal Curve
Frequency
KS2 standardised score
0
100
200
300
400
500
600
-3
-2
-1
0
1
2

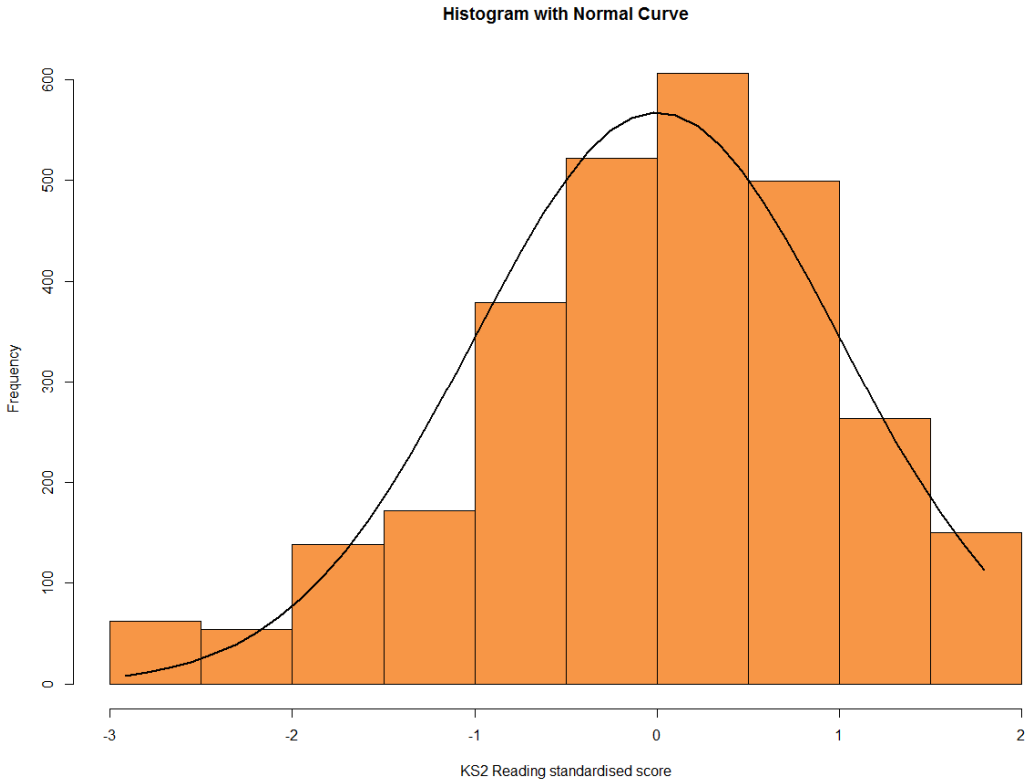

Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
-3
-2
-1
0
1
2
3

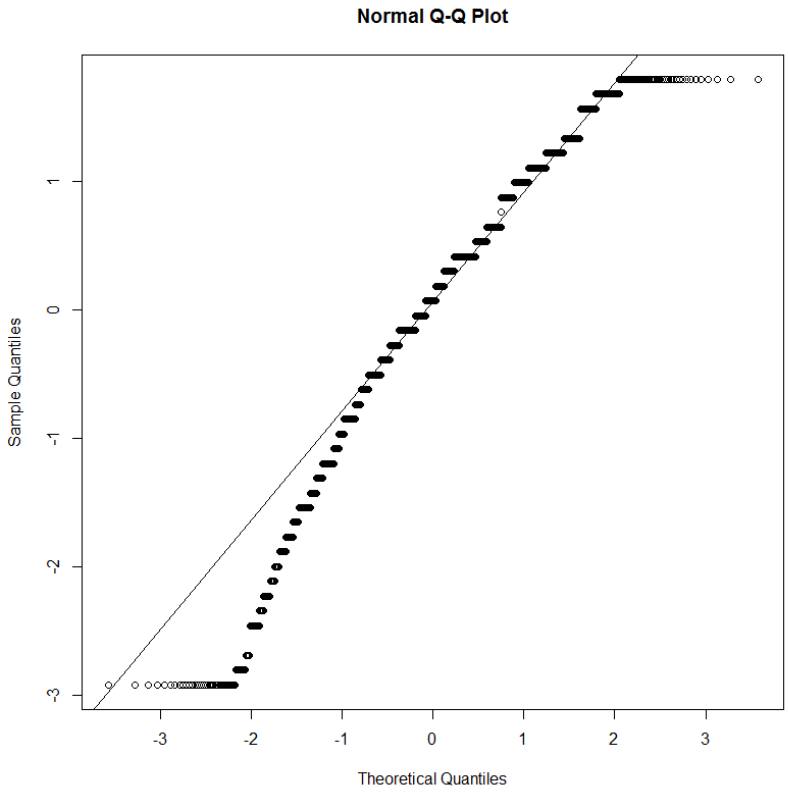
Maths


Histogram with Normal Curve
Frequency
KS2 Maths standardised score
0
100
200
300
400
500
600
700
-3
-2
-1
0
1
2

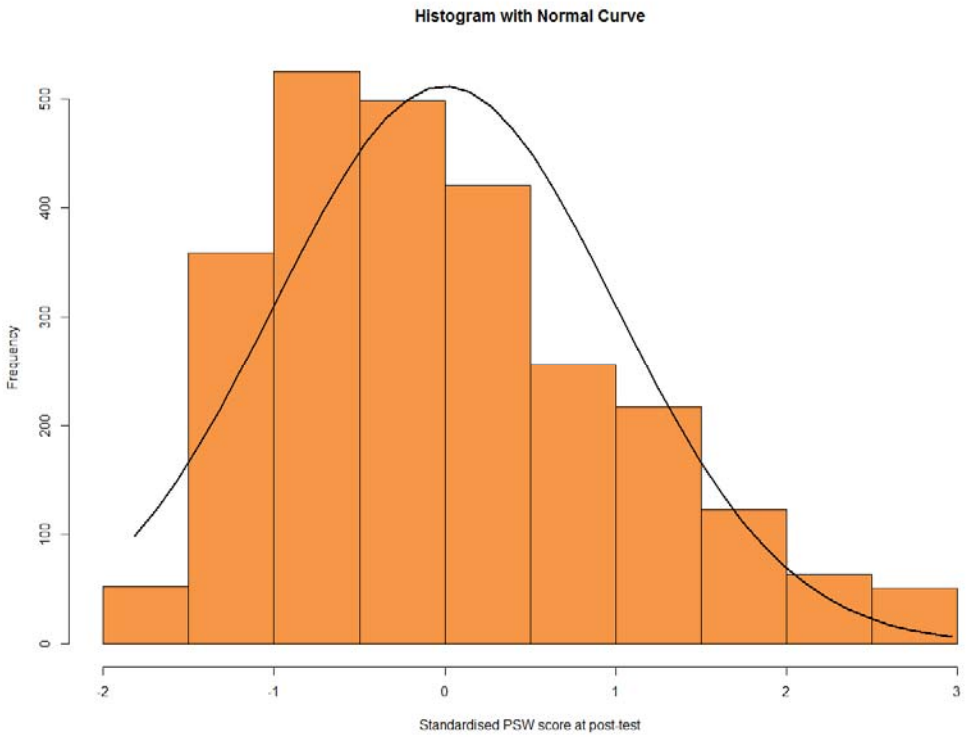

Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3

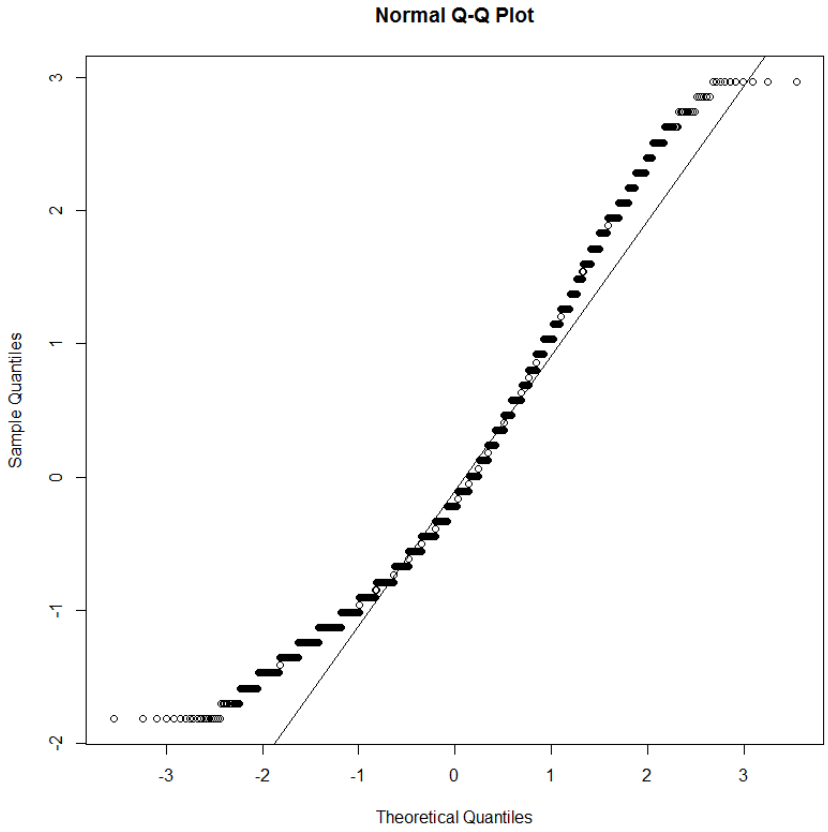
Reading


Histogram with Normal Curve
Frequency
0
100
200
300
400
500
600
-3
-2
-1
0
1
2
KS2 Reading standardised score

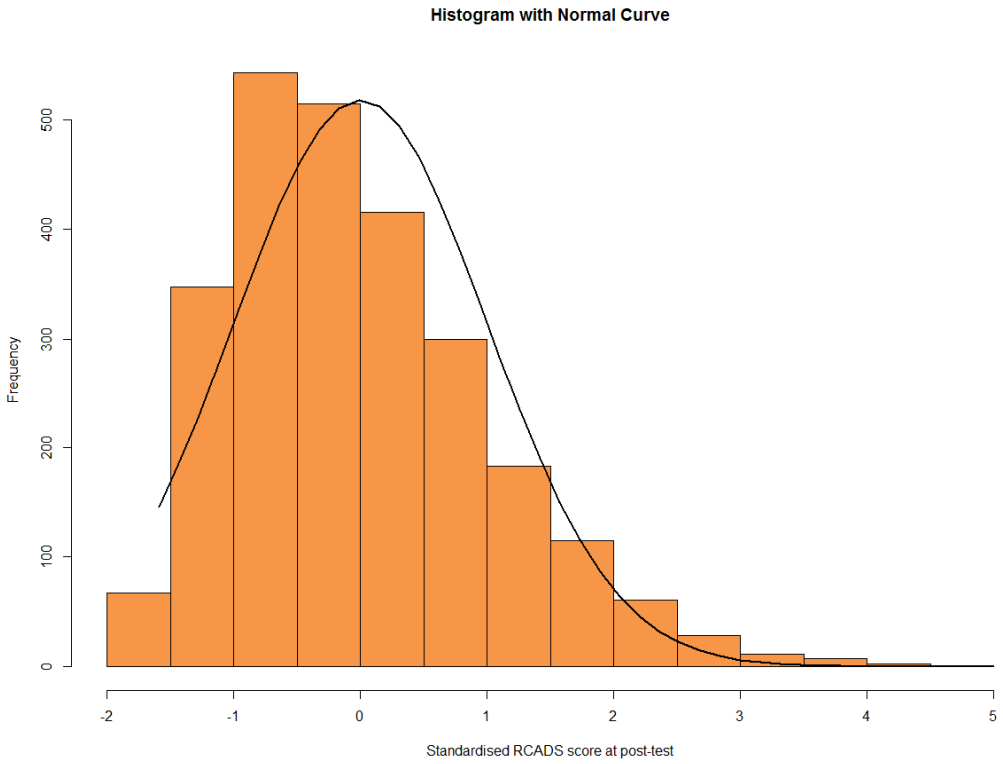

Normal Q-Q Plot
Sample Quantiles
-3
-2
-1
0
1
-3
-2
-1
0
1
2
3
Theoretical Quantiles

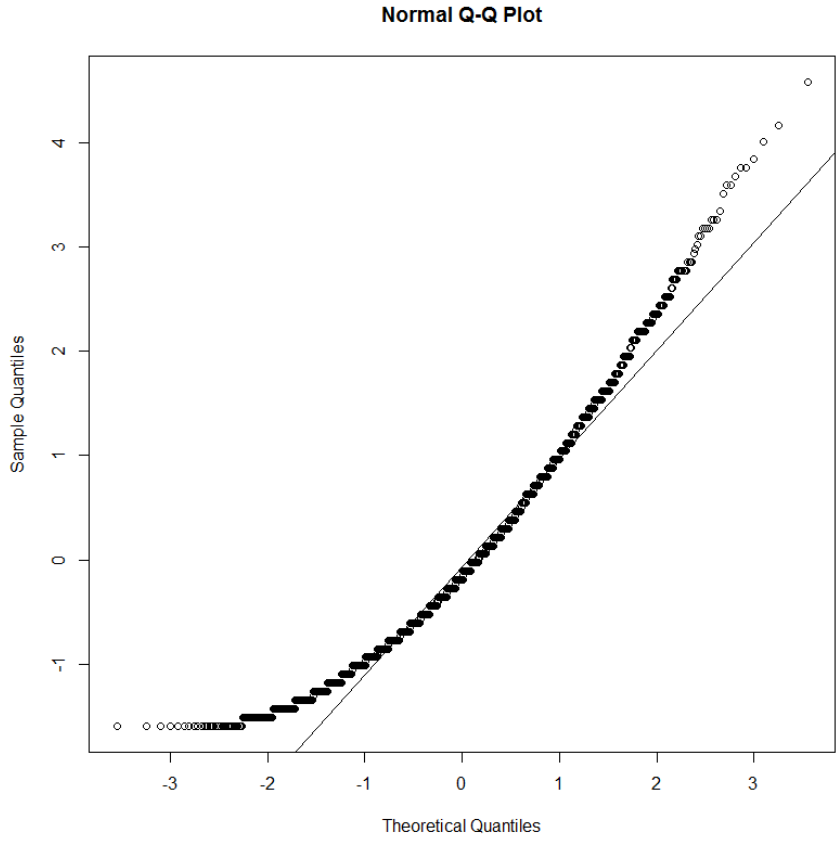
PSW post-test





RCADS at post-test




Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
-3
-2
-1
0
1
2
3
-1
0
1
2
3
4

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