



Texting Students and Study Supporters (Project SUCCESS)

Evaluation Report

May 2020

Berenice Scandone, Robert Wishart, Julia Griggs, Neil Smith, Helen Burridge, Katriina Lapanjuuri, Peter Hall, Tom Chadwick and Phoebe Averill

NatCen
Social Research that works for society



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.





This project was funded by the Education Endowment Foundation (EEF) as part of a joint initiative with J.P. Morgan to explore how to improve outcomes for disadvantaged 16- to 18-year-old students who achieve below a grade 4 in GCSE English or maths.

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:

-  Jonathan Kay
 Education Endowment Foundation
 5th Floor, Millbank Tower
 21–24 Millbank
 SW1P 4QP
-  0207 802 1653
-  jonathan.kay@eefoundation.org.uk
-  www.educationendowmentfoundation.org.uk

Contents

About the evaluator	3
Executive summary	4
Introduction	6
Methods.....	15
Impact evaluation results	28
Implementation and process evaluation	39
College staff time spent on Project Success	53
Conclusion	55
References	59
Appendix A: EEF cost rating.....	61
Appendix B: Security classification of trial findings.....	62
Appendix C: Effect size estimation (secondary outcome)	64

About the evaluator

The project was independently evaluated by a team from NatCen Social Research:

Berenice Scandone, Robert Wishart, Julia Griggs, Neil Smith, Helen Burrridge, Katriina Lapanjuuri, Peter Hall, Tom Chadwick and Phoebe Averill.

The lead evaluator was Neil Smith.

Contact details:

Neil Smith

NatCen Social Research

35 Northampton Square

London EC1V 0AX

Email: Neil.Smith@natcen.ac.uk

Tel: 020 7250 1866

Executive summary

The project

Project Success was developed by the Behavioural Insights Team (BIT) and aimed to use text messages to improve GCSE English and maths re-sit pass rates by prompting students to attend classes and exams, engage with study materials, and form better study habits, either through direct contact with the learner or through prompting a dialogue with a nominated study supporter such as a family member. The text messages were targeted at further education college students aged between 16 and 18 years and who were re-sitting maths or English. Over the course of the academic year, weekly text messages (a total of 36 for English or 37 for maths) were sent to students or their study supporters (or both) via the BIT Promptable text messaging service. This project was funded by the Education Endowment Foundation (EEF) as part of a joint initiative with J.P. Morgan to explore how to improve outcomes for disadvantaged 16- to 18-year-old students who achieve below a grade 4 in GCSE English or maths.

The evaluation included 3,779 students across 31 further education (FE) colleges in England. The efficacy trial used a four-armed, multi-site, randomised controlled design with individual random assignment to each trial arm. The four trial arms were: student received text messages, study supporter received text messages, both student and study supporter received text messages, and control (no text messages). The efficacy trial investigated the extent to which the receipt of text messages (either by the student, a study supporter, or both) improved students' college attendance and GCSE maths or English re-sit results. The trial also explored whether any effect of receiving text messages varied according to the student's gender or whether they had ever been eligible for free school meals. The primary outcome was obtaining a pass grade (4–9) in GCSE English or maths upon re-sitting, with lesson attendance being assessed as a secondary outcome. Where students were re-sitting both English and maths, the subject used for the intervention was randomly selected; the same subject was analysed as the primary outcome.

Alongside the impact evaluation, a mixed-methods implementation and process evaluation (IPE) was carried out. This included: a short, online diagnostic survey of students developed by BIT and NatCen, which was completed as part of the recruitment process; an observation of a tutor workshop; and interviews with college project and subject leads, study supporters, and students. The trial started in September 2017 and concluded in October 2019 with the intervention being delivered throughout the academic year 2017/2018.

Table 1: Key conclusions

Key conclusions	
1.	There is no evidence that the Project Success intervention had any impact on the GCSE English or maths re-sit pass rate for further education college students.
2.	There is no evidence that the Project Success intervention had any impact on the attendance of further education college students re-sitting GCSE English or maths.
3.	The intervention did not have a differential impact on the GCSE re-sit pass rate by gender or by eligibility for free school meals (at the end of KS4). The subject being re-examined or the number of re-sits being taken also did not lead to differential effects from the intervention.
4.	The use of mobile phone technology was perceived as a highly appropriate, effective, and low risk means of engaging with the target student cohort, though mobile phone use was less popular among study supporters.
5.	There were significant limitations to the programme's ability to engage those who may need it the most as it was the highly motivated students that were more engaged with their studies and with college generally who were more likely to sign up to the intervention.

EEF security rating

These findings have a high security rating. This was an efficacy trial, which tested whether the intervention worked under everyday conditions in a large number of schools. The trial was well-designed, well-powered, and relatively few pupils who started the trial were not included in the final analysis. The pupils that received the intervention were similar to those assigned to the control group in terms of prior attainment. The trial lost one padlock due to moderate threats to validity caused by concerns about self-selection of pupils, lack of information on concurrent interventions, and evidence of contamination between groups.

Additional findings

The intervention had no impact on college attendance, which is hypothesised to improve attainment (Miller et al., 2016 and Groot et al., 2017). Though the content of text messages was not specifically designed to encourage attendance, this may in part explain the null findings of the primary analysis. In support of the logic model, text messages sent either to students or their study supporters encouraged revision by functioning as reminders and planning tools, and by providing useful web resources. Motivational texts were perceived to be less useful. Most students stated that receiving BIT text messages prompted increased engagement in revision outside of college, but the extent of additional revision reported varied considerably among participants. Conversely, increased engagement in revision outside of college did not translate into increased engagement in maths and English classes. There are limitations to the analysis of attendance data relating to data completeness and quality.

College-wide engagement with the texting intervention was low, potentially missing a further opportunity to maximise the impact of the regular delivery of motivational messages. Text messages may not have maximised the opportunity to engage with and motivate the students or study supporters, with content being generic and students becoming unengaged with material once they had detected the automated system in place. Content delivery times were not universally suitable to all students or study supporters, with individual preferences for times or frequencies likely to predict the level of engagement with the content and consequent motivation




A key limitation of the evaluation design was the reliance on a binary outcome of pass/fail at GCSE re-sit. This did not allow the impact evaluation to inspect for more finely graded improvements in student outcomes meaning a considerably large step-change in attainment was required in order to detect a difference in academic performance between study arms. Months of progress that the treatment provided over the control group has been estimated by converting relative risk ratios to Hedge's g (see Methodology). Statistical uncertainty around the results mean that the estimates are consistent with very small positive or negative impacts—with the best estimate being zero months impact for all of the primary outcomes. The results of this trial may not be fully generalisable because the student survey at sign-up suggests that those who opted into the trial were more likely to be highly motivated and generally more engaged with their college studies.

Cost

The costs per pupil per year over three years for delivering each of the three arms of Project Success are £13 for text messages to students or study supporters only, or £15 for text messages to both.

Impact

Table 2: Summary of impact on primary outcomes

Outcome/ Group	Relative risk ratio (95% confidence interval)	Hedges' g (95% confidence interval)	Months progress	EEF security rating	Pupils (N)	P Value	EEF cost rating
Pass GCSE re-sit in English or maths: texts to students	1.03 [0.86; 1.26]	0.01 [-0.05; 0.07]	0		1,877 [939; 938]	0.740 ^a	£ £ £ £ £
Pass GCSE re-sit in English or maths: texts to supporters	1.04 [0.84; 1.27]	0.01 [-0.05; 0.08]	0		1,872 [934; 938]	0.691 ^a	£ £ £ £ £
Pass GCSE re-sit in English or maths: texts to students and study supporters	0.92 [0.73; 1.14]	-0.03 [-0.09; 0.04]	0		1,873 [935; 938]	0.343 ^a	£ £ £ £ £
Pass GCSE re-sit in English or maths: ever eligible for FSM at the end of KS4 (sub- sample)	0.91 [0.65; 1.29]	-0.03 [-0.12; 0.08]	0	N/A	1,056 [0.65; 1.29]	0.606	£ £ £ £ £

^a 95% confidence intervals for the Relative Risk Ratios have been Bonferroni adjusted for the primary analysis. Therefore, a p-value of less than 0.0167 would indicate a statistically significant difference with a type-one error rate of 0.05.

Introduction

Background

Current government policy requires all students aged 16 to 18 who do not hold a GCSE grade 9 to 4 in maths or English to continue studying these subjects. Students with maths or English GCSE attainment at grade 3 are required to enrol in a GCSE qualification (in maths and/or English as appropriate) to achieve at least a grade 4. In 2016/2017, 41.5% of students left secondary school without a 9–4 for English and maths (DfE, 2017). However, only one out of four post-16 re-sit students achieved a passing grade (DfE, 2018). Around half the students who fail in maths and/or English at age 16 go on to re-sit the qualification at a further education (FE) college (DfE 2016). These students are more likely than those in schools or sixth form colleges to have lower prior attainment, and considerably less likely to achieve a pass at re-sits (Impetus-PEF, 2017). FE colleges also experience challenges in providing personalised support to students due to structural issues affecting the sector such as substantial budget cuts and high staff turnover (Frontier Economics, 2017; McNally and Wyness, 2017; Wolf, 2015). Identifying effective, low-cost and scalable interventions to improve FE students English and maths re-sit attainment is thus a key question for post-16 education policy and practice.

Previous research indicates that texting family and peers can have a positive impact on students' attainment and attendance. A recent efficacy trial of a school-level intervention conducted in England, developed by the Behavioural Insights Team (BIT) and designed to improve pupil outcomes at Key Stages 3 and 4 by texting parents/carers, found a small positive impact on maths (0.067 SD) and English (0.033 SD), and a reduction in absenteeism (-0.054 SD) (Miller et al., 2016). The intervention involved sending 30 text messages to parents of Year 7, 9, and 11 students from 36 secondary schools to inform them about dates of upcoming tests, whether homework was submitted on time, and what their children were learning at school. The trial was a two-armed, multi-site cluster randomised controlled trial (RCT) involving a sample of 15,697 students from 29 schools. Overall, the findings have moderate security as 19% of the schools dropped out of the trial and there were important differences between pupils in the treatment and control group, with the latter more likely to be eligible for free school meals. The research team surveyed parents about their engagement in their child's learning through phone interviews with over 1,900 parents of intervention pupils. The odds of parents who received the text messages talking to their child about studying for an upcoming test were almost three times greater than the odds of parents who did not receive text messages, thus showing that prompts from schools may have helped parents take a more active role in their children's education.

A randomised controlled trial conducted in the U.S. found that high-school students whose parents received weekly, one-sentence, individualised messages from teachers about the student's performance and behaviour in school were less likely to be absent and to drop out of a class (Kraft and Rogers, 2015). The evaluation employed a blocked randomised trial design with multiple treatment arms. Students and parents were assigned to one of three conditions—positive information (n = 146), improvement information (n = 136), or control (n = 153), blocking on the first class taken by each student. Messages decreased the probability a student was absent by 2.5 percentage points (p = 0.011) and that they dropped out of a class by 6.1 percentage points (p = 0.046), and reduced the percentage of students who failed to earn course credit from 15.8% to 9.3%—a 41% reduction. The mechanisms through which the messages affected student success were assessed through teacher and student surveys and phone interviews with parents. While no evidence was found that texts increased the occurrence of supportive conversations overall, students in the treatment group reported that their parents spoke to them more often about things they should work on to improve in school.

Finally, two recent trials conducted by BIT assessed the impact of an intervention involving sending text messages to study supporters selected by students on student attendance and attainment (Groot et al., 2017). This intervention builds on the above discussed evidence on the effectiveness of texting parents and—recognising that students at post-16 institutions may no longer live at home, may have other supportive adults in their lives, and may be seeking more independence and not necessarily want their parents to be closely involved with their education—takes a different approach where students could choose their own 'study supporter'. Researchers found a positive effect of the intervention with the effect size of the intervention amounting to an increase of 0.10 standard deviations in attendance. Yet, the trial did not assess the behaviours that study supporters engaged in after receiving the text messages. It is thus possible that the positive effect of the intervention was due to study supporters simply passing on the information contained in the messages rather than actively engaging in supportive behaviours such as helping with assignments or providing emotional support.

Project Success extended the scope of previous interventions to test different approaches to the programme: (1) texts targeted at the student, (2) texts targeted at a study supporter such as a family member or peer identified by the student, or (3) texts targeted at both. The quantitative analysis estimated the effect size of each intervention compared to the single 'business-as-usual' control group. This enabled an exploration of the impact of directly engaging students as well as promoting positive relationships with supportive peers or adults, thus helping to assess the mechanisms eventually leading to improved attainment. The primary outcome was the percentage of students that pass their GCSE mathematics or English re-sit exams after one year. Students' attendance was considered as a secondary outcome measure.

The evaluation is a four-armed, multisite, randomised controlled efficacy trial. This design was chosen because the risk of contamination was perceived to be low when the evaluation was designed. The multisite design also yields greater statistical power, allowing the three variants of the intervention to be tested against control with a relatively small sample of colleges. A mixed-methods implementation and process evaluation (IPE) was also conducted alongside the impact assessment. The IPE involved a survey of eligible students and in-depth interviews with college project and subject leads, students, and study supporters. This approach enabled us to bring together the views of all groups involved in the programme to provide a rich understanding of delivery across colleges, with a focus on: reach of the intervention among its target group; buy-in from stakeholders; fidelity of implementation; usual practice at participating colleges; responsiveness among students and study supporters; and perceptions of intermediate outcomes.

Intervention

Overview and definition

The intervention evaluated is the Texting Students and Study Supporters programme (known as Project Success) developed by the Behavioural Insights Team (BIT). The intervention consisted of 36 or 37¹ text messages sent to students re-sitting GCSE English or maths, or sent to named study supporters, over the course of an academic year.

Why: rationale

Two recent trials conducted by the Behavioural Insights Team (BIT) in England have found an increase in FE college students' attendance through a programme of text messages sent to a nominated study supporter (Groot et al., 2017). The text messages encouraged interaction between the student and their supporter, specifically in relation to the students' learning and upcoming assessments. This intervention builds upon previous trials by including text messages sent directly to students as well as to study supporters. The evaluation was therefore able to assess the impact of directly engaging students, as well as through interactions with supportive peers or adults, thus helping to test the mechanisms underlying eventual improvements in students' outcomes.

Who: recipients

Project Success was targeted at FE college students re-sitting English or maths GCSEs. In line with legislative requirements for re-sits and with the intervention aim to improve English or maths re-sit outcomes, the primary outcome of interest was GCSE re-sit results in one of these two subjects.

Students were eligible to take part in the trial if they were:

- enrolled at a participating FE college in England in September 2017; and
- due to re-sit GCSE maths and/or English in the 2017/18 academic year.

Study supporters were identified by students who were asked to select someone over 16 who cared about them and their learning, and whom they trusted and felt comfortable talking to about college. Students were also advised that study supporters did not need to be experts in English or maths but should not be a member of their class. Study supporters were typically family members or peers. The supporter's role was described to students by college staff and a short informational video.

¹ Maths students have one additional exam and therefore receive an additional reminder.

What: materials

Depending on which trial arm they were assigned to, students and/or study supporters were sent weekly text messages. These messages contained either motivational content, information on course content, academic resources such as practice websites, notifications about deadlines, details of extra tutorial sessions, or exam dates (see Supplementary Appendix). If both the student and study supporter were receiving messages, the messages would also encourage them to speak with one another about the content. In the supporter-only arm, the supporter was encouraged to speak with the student. Recruitment and sign up processes were facilitated by the production of a study website. Colleges were instructed not to send any motivational text messages during the intervention implementation period.

What: procedures

Following the recruitment of colleges to Project Success, English and maths tutors with responsibility for teaching GCSE re-sit students at participating colleges were invited to attend an informational workshop. Workshops were delivered by one or two BIT staff members at all participating colleges in the months preceding the 2017/2018 academic year. College project and subject leads were responsible for ensuring all maths and English GCSE tutors attended the workshop, although this was not always possible. Attendance varied across sessions. No information on tutor attendance was routinely collected as part of the evaluation, however, it is estimated that most sessions were attended by more than half of tutors and in some instances, attendance was poorer (around a quarter of tutors). Attendee numbers ranged from around four to 20 tutors. The purpose of the workshops was to introduce all maths and English GCSE tutors to the project and to their responsibilities by providing background information (what BIT is, rationale for the intervention and evaluation) and details of the project (tasks, timelines, and responsibilities). The workshop also provided a space for college staff to provide initial feedback on the project and suggestions on what the messages could cover, including what students struggled most with, and to ask any queries about the intervention.

Project Success was advertised to students in participating colleges using posters (see Supplementary Appendix). Students identified by college project leads as eligible to receive BIT text messages took part in an online recruitment process delivered by English or maths tutors with responsibility for teaching GCSE re-sit students at each college. As part of the student recruitment process, students were shown an online video designed by BIT to inform them about the intervention and the trial. Students signed up for the intervention and evaluation using an online survey developed by BIT and NatCen and hosted by BIT. This included information on taking part and use of personal data, and a question for sign-up followed by a set of questions on students' demographic characteristics and their interest, motivation, and engagement in maths, English, and college. For students agreeing to take part in the programme, the survey also gave prompting questions to help them identify a suitable supporter (for example, 'Who encourages and motivates you the most?', 'Who do you talk to about your goals?', 'Who is a role model to you?') and collected information on selected supporters and students' relation with them, and mobile phone numbers for both the student and study supporter. For those who opted out, it explored main motivations for doing so.

Study supporters entered the programme indirectly, via a student who nominated them, but were given the option to opt out. BIT advised tutors to split the sign-up process into two components: (1) one section to describe the project, get students to think about potential supporters, and talk to their supporters about it, and (2) a second element to be delivered a few days later, giving students time in class to sign up for the programme. When nominating their study supporter, students were asked to provide their phone number to enable BIT to deliver the text messaging programme to them. After signing up, students were encouraged to notify the person they had nominated to be their study supporter about the project. In the first text message, study supporters were provided with information about the programme, informed that they had been chosen by the student as their supporter, and given the option to opt out of receiving the text messages (by replying 'STOP'). They were reminded again of how to opt out at the midpoint of the year. If at any point they texted STOP or provided any other response indicating an intent to withdraw, BIT unsubscribed them. BIT did not foresee any risk for study supporters to participate. No further data was collected on the supporters and their activities were not analysed. Their data was only used for the purpose of delivering the messages.

It was intended that all students and/or their study supporters would receive 36 or 37² weekly text messages over the course of the academic year. Text messages were developed by BIT staff with the support of college maths and English tutors. Texts were drafted as follows:

- BIT staff completed a 30-minute phone call with tutor leads. These were the maths and English tutors nominated at each college to be the main point of contact with BIT.
- Tutor leads shared their Schemes of Work with BIT staff.
- BIT drafted a master schedule of messages using the Schemes of Work (identifying topics and tasks that were consistent across colleges), tutor ideas, and the behavioural science literature. The schedule was then tailored to fit the details of each college (again using each Scheme of Work). Where colleges had multiple schemes of work for different classes teaching the same subject, BIT identified topics that were overarching across all classes and scheduled the message to be sent after all students had encountered the topic; each college received a single suite of messages per subject (one for maths and one for English).
- BIT staff shared the draft schedule of messages with tutor leads for feedback. Tutor leads were instructed to share these with all tutors for input. BIT received input from most, but not all, colleges.
- Tutor leads shared the draft schedule of messages with feedback with BIT staff.
- Where appropriate, BIT staff amended messages accordingly and shared the final draft of messages with tutor leads.
- All schedules were signed off at the start of each term and additional updates were sent to enable tutors to make any last-minute amendments.
- Each week, tutor leads received an automated email with the message scheduled to go out that week. Tutors were instructed to review the message and send amendments to BIT staff within a couple of working days where necessary. BIT received an average of seven responses per week to text message emails from tutor leads.
- BIT staff were responsible for sending the text messages (through *Promptable*). Messages were scheduled onto the system and were amended according to tutor feedback each week.

Who: implementers

The intervention was implemented and delivered by BIT. BIT recruited colleges to take part in the project, liaised with college project leads, provided information and instructions for tutors and advertising material for students, designed the online survey used for student sign-up, and drafted the suite of text messages. BIT also sent the text messages via its texting platform and handled all opt-outs and other incoming messages.

Each college had a nominated project lead who took responsibility for the intervention at their institution and they acted as the key contact for BIT. The project leads were responsible for cascading information to subject leads and to English or maths tutors with responsibility for teaching GCSE re-sit students at each college.

The intervention was designed to have two (or more) subject leads at each college. Each subject lead, one for English and the other for maths, was responsible for providing a planned schedule of work at the start of the academic year and for reviewing and editing the text messages on a weekly basis.

English or maths tutors also played a role in the intervention by informing their students about Project Success and facilitating and implementing the sign-up process.

² Maths students have one additional exam and therefore receive an additional reminder.

How: mode of delivery

Text messages were sent to students and study supporters via BIT's texting platform *Promptable* (<https://promptable.com>).

Where: settings

The intervention was conducted in Further Education colleges located in England. College eligibility was determined by two factors:

- Institution status on EduBase: Further Education Corporations were eligible, but Sixth Form Corporations were not. Specialist Designated Colleges (for example, land-based providers) were to be omitted unless the overall target of 31 colleges could not be recruited from Further Education Corporations.
- The number of eligible students: Colleges with at least 100 eligible students were targeted in the first instance. The rationale for this decision being that the greater the number of participating students, the greater the statistical power in the analysis of student outcomes.

When and how much: dosage

Students and/or named study supporters were sent 36 or 37 weekly messages³ over the course of the 2017/2018 academic year. The first messages were sent a few days after the student had been randomised in early November 2017 and continued until the GCSE (re-sit) exams in June 2018. Messages were sent weekly on a Thursday evening. The timing of the texts was varied by up to an hour each week but they were always sent after college hours (from five pm onwards) to minimise the risk that students in different intervention groups might talk to one another about the texts. Timing was selected to motivate students for the week ahead and not be during the weekend or normal college hours. Tutors confirmed that the timing should be outside of college hours. Timing was also informed by prior iterations of the intervention conducted (Groot et al., 2017).

Tailoring

The intervention was delivered in the same way in all participating colleges and to all participating students (although the nature of delivery, that is, whether text messages were sent to the student, their study supporter, or both, varied according to which randomisation group the student was assigned to). College tutors at participating colleges were able to tailor the text messages in line with their own schedule of work. This was intended to ensure the messages were fully relevant for students in terms of the college timetable and local curriculum. However, the content of the messages could not be tailored to individual students or groups of students.

Issues that occurred during the project**Recruitment**

Sign-up of students to the intervention was at times difficult to obtain and was very low in certain colleges. The main motivation for students opting out of receiving BIT text messages at recruitment stage was the desire not to take part in research. Recruitment of students and study supporters for the IPE also proved difficult, which impeded in-depth analysis of student-study supporter dyads as originally planned.

Randomisation

Some students who should have been excluded from the evaluation were randomised in error. In total, 18 pupils were randomised when they were not re-sitting a GCSE in English or maths. A further 26 students had already achieved a pass (grade 9–4) in the subject they were re-sitting in a previous academic year. These students have been classified as 'randomised in error' and are not included in the randomised sample. These students were identified after randomisation but prior to analysis.

³ Maths students have one additional exam, and therefore receive an additional reminder text.

Implementation

A few students and study supporters reported, during interviews conducted as part of the IPE, receiving BIT text messages less often than intended as part of the intervention. It is not clear why this happened, and whether it was due to participants not recalling the reception of messages or messages not being received. Interviews with students and study supporters also highlighted instances where students and study supporters stopped receiving text messages, either as they had changed number and did not notify BIT or for unknown reasons. Approximately two-fifths (37%) of students in the 'texts to students' and 'texts to study supporters' trial arms received all the texts intended for them. Just one-fifth of students in the 'students and study supporters' trial arm received the full schedule of texts. A full breakdown of the proportion of texts received by students in different trial arms can be found in the Impact Evaluation chapter.

Contamination

As randomisation was at pupil level rather than college level, there are several ways which contamination could have occurred between students in different trial arms.

Although BIT encouraged students to nominate study supporters outside of their GCSE peer group, students could have nominated a fellow student who may also have been participating in Project Success. Therefore, it is possible a student who was assigned to the control group was receiving text messages as a study supporter, which would have affected their validity as a member of the control group.

There was also a risk of students passing on content from the text messages they received themselves or from interaction with their study supporter to students in the control group. This may have affected outcomes for the control group and therefore the evaluation's ability to detect an impact.

Contamination is explored as part of the process evaluation. In particular, interviews with students and with project and subject leads were used as an opportunity to discuss behaviour towards students/peers, and in the case of students, whether information was shared within/outside the GCSE cohort.

Data collected as part of the sign-up survey shows that 275 students who signed up for Project Success provided the same telephone number for student and study supporter. This may have led to contamination between the student-only/student-and-supporter and the supporter-only arms, as these students received BIT text-messages intended for their supporters. Two hundred and forty-one students who signed up were also nominated by others as supporters, which may have led to contamination between the student-only/student-and-supporter and the supporter-only arms and to spill-over to the control group. In interviews conducted as part of the IPE, college project and subject leads reported that no discussion took place in class about the text messages, although a few students mentioned discussing these with peers.

Intervention logic model

The Project Success logic model set out in Figure 1 (below) was developed in conjunction with representatives from BIT during the set-up phases of the project. The logic model includes the sequence of activities at the level of developer, college tutor, student, and study supporter. No modifications were made during the trial.

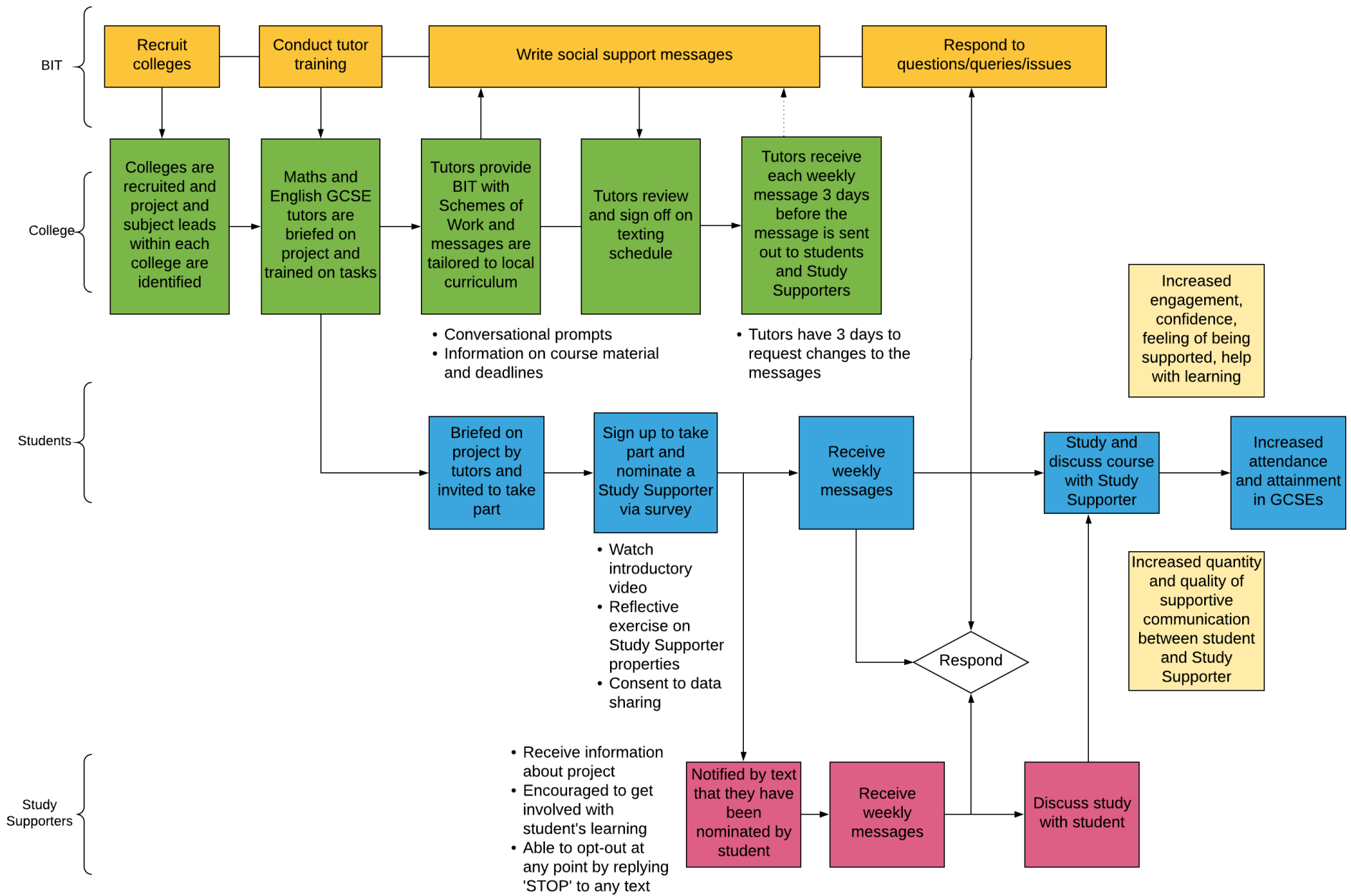


Figure 1: Project Success logic model

Evaluation objectives

The efficacy trial aimed to answer the following principal research questions:

1. To what extent does the receipt of text messages (either by the student, a study supporter, or both) improve students' college attendance and GCSE maths or English re-sit results—compared to those who do not receive text messages?
2. To what extent, if at all, do impacts differ by a student's gender and whether they have ever been eligible for free school meals (as a measure of disadvantage)?

The implementation and process evaluation (IPE) was designed to explore how the intervention was implemented and delivered, and understand how, why, and for whom it works.

This element of the study aimed to address the following research questions:

1. How was the intervention implemented?
2. How was the intervention interpreted and delivered across colleges and trial arms?
3. To what extent did contextual variation affect fidelity?
4. What adaptations (if any) were put in place?
5. To what extent did students and study supporters engage with and act upon the messages they received?
6. What were the barriers to delivery and how were these addressed?
7. What facilitated successful delivery?
8. What was the cost of delivery?

The evaluation protocol⁴ for Project Success and the statistical analysis plan (SAP)⁵ can be found on the Education Endowment Foundation's website.

Ethics and trial registration

Ethical review and trial registration

Ethical approval for the impact and process evaluation was obtained from NatCen's Research Ethics Committee in May 2017. At college recruitment stage, a Memorandum of Understanding (MoU) was signed by each college expressing their agreement to take part in the trial and all evaluation tasks. The MoU set out the roles and responsibilities of the three parties (the college, BIT, and NatCen) and formalised the college's commitment to take part in the trial (see Supplementary Appendix). Students were informed of what participation in the intervention and evaluation would involve, how their data would be used, and their rights to withdraw (see below) through the online survey used for sign-up. By agreeing to take part in Project Success, students also agreed to take part in the trial and the processing of their data (see Supplementary Appendix). Study supporters' phone numbers were provided by students in the sign-up survey and they had the opportunity to opt out of receiving BIT text messages at any time over the course of the intervention by replying to the texts. No further data was collected on the supporters and their phone number was only used for the purpose of delivering the messages.

The trial was registered with International Standard Randomised Controlled Trial Number (ISRCTN). The trial number is ISRCTN70011940.⁶

Data protection

Data processing roles and legal basis

The National Centre for Social Research (NatCen) was the single data controller and was a data processor in conjunction with the Behavioural Insights Team for this project. As data controller, NatCen was responsible for deciding the purpose and legal basis for processing data (under Article 6 of GDPR). For this project, NatCen's assessment is

⁴ https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Protocols/Post_16_-_Project_SUCCESS_AMENDED.pdf

⁵ [https://educationendowmentfoundation.org.uk/public/files/Projects/Post_16_-_Texting_Students_\(Project_SUCCESS\)_SAP.pdf](https://educationendowmentfoundation.org.uk/public/files/Projects/Post_16_-_Texting_Students_(Project_SUCCESS)_SAP.pdf)

⁶ <http://www.isrctn.com/ISRCTN70011940>

that the evaluation fulfilled one of its core business purposes (undertaking research, evaluation, and information activities) and is therefore in its 'legitimate interest'. NatCen has considered and balanced any potential impact on the data subjects' rights and find that its activities will not do the data subject any unwarranted harm.

Use of data

The research team at NatCen had access to student, study supporter, and college project and subject lead contact details, recordings of interviews, transcripts, charted data, student sign-up information (including contact details), and students' exam results. McGowan Transcriptions had access to recordings and transcriptions from all interviews. McGowan Transcriptions are an approved supplier and compliant with all NatCen information security policies.

All data collected was used for research purposes only. Data gathered from interviews with students, study supporters, and college project and subject leads—alongside data from the student survey—was used to inform the process study element of the project. Participation in all IPE activities took place on a voluntary basis. Student data from the National Pupil Database was used for the impact study element. Students had the opportunity to object to their data being used for the evaluation by getting in contact with NatCen's team via email. All data was anonymised before being analysed and archived.

At the end of the research, anonymised student NPD data will be shared with the Education Endowment Foundation's archive manager, FFT Education, and stored in the Office for National Statistics.⁷ At this point, the EEF becomes data controller and the EEF's archive manager becomes data processor. All personal information, and any other data held on the project, will be securely deleted from NatCen's server within six months from report publication. BIT will delete data within three months of publication.

Privacy notice

A privacy notice was published on the study page on NatCen's website and issued to all colleges.⁸ The privacy notice explained the legal basis for data collection, provided information on who had access to the data, and provided information on how the data would be used (see Supplementary Appendix).

Project team

Project team at BIT

The intervention manager, Bibi Groot, was assisted by Kimberly Bohling (Senior Research Advisor) and Sara Halkiopoulos (Research Advisor), with oversight from Michael Sanders (Head of Research and Evaluation). Samantha Dodd, Miranda Jackman, and Patrick Taylor provided support on recruitment and message development. Paul Calcraft (Tech/Product Lead, Promptable) managed programming and delivery of all text messages. Todd Rogers (Professor of Public Policy at the Harvard Kennedy School) originally developed the Study Supporter intervention and provided advice on implementation.

Evaluation team at NatCen

The evaluation was managed by staff from the evaluation team and Children and Families team at NatCen. The trial manager was Julia Griggs (Research Director), assisted by Peter Hall, Katriina Lapanjuuri, and Berenice Scandone (Senior Researchers). The management team were supported by Phoebe Averill, Helen Burridge, and Tom Chadwick (Researchers). Neil Smith (Head of Analysis) led the randomisation and oversaw the completion of the project including the impact evaluation, which was conducted by Robert Wishart (Senior Researcher, Analyst).

⁷ ONS archive ingestion is currently planned for early 2020.

⁸ <http://www.natcen.ac.uk/taking-part/studies-in-field/evaluation-of-project-success/privacy-notice/>

Methods

Trial design

Table 3: Trial design

Trial design, including number of arms		Four-armed, multi-site randomised control trial
Unit of randomisation		Student
Stratification variable (s) (if applicable)		College
Primary outcome	Variable	English or maths attainment
	Measure (instrument, scale, source)	GCSE English or maths re-sit, passed (grades 9–4) or failed (3–1), NPD
Secondary outcome(s)	Variable(s)	Attendance at English or maths lessons
	Measure(s) (instrument, scale, source)	College registers ⁹
Baseline for primary outcome	Variable	No baseline outcome
	Measure (instrument, scale, source)	No baseline outcome
Baseline for secondary outcome(s)	Variable	No baseline outcome
	Measure (instrument, scale, source)	No baseline outcome

The evaluation of Project Success used a four-armed, multi-site, randomised controlled efficacy trial with individual random assignment. The four trial arms were:

- student received text messages;
- study supporter received text messages;
- student and study supporter both received text messages; and
- control (no text messages, business as usual).

This design was chosen because the risks of spill-over or contamination were perceived to be low when the evaluation was designed and this design yields the greatest statistical power. In practice, contamination did occur (see Impact Evaluation).

The primary outcome is obtaining a pass grade (9–4) in GCSE English or maths upon re-sitting. Where students were re-sitting both English or maths, the subject used for the intervention was randomly selected; the same subject was analysed as the primary outcome. Attendance at English or maths lessons was analysed as a secondary outcome.

⁹ The number of possible classes varies between colleges, but typically had approximately 75 potential classes.

Participant selection

In total, 3,823 students were initially randomised to one of the intervention arms or a business as usual control group on a 1:1:1:1 basis. However, it later emerged that some students had not met the eligibility criteria (classified as 'randomised in error'. This occurred for two reasons: first, they were not originally attempting to re-sit GCSE maths or English but instead were entered into a different award in the same subject(s) (n = 18); second, some students had already achieved a level four or above in the GCSE of the subject they were re-sitting (n = 26). Consequently, the as-randomised sample has a total sample size of 3,779 students.

Outcome measures

Baseline measures

There are no baseline measures in this study. To be eligible for the intervention, students must have achieved a GCSE grade of less than four, meaning that KS4 results would not vary substantially for the students eligible for the trial. Almost all students would have achieved a grade three, as students achieving a grade one or two would have likely been encouraged to attempt a different qualification. Therefore, it was perceived that it would not be appropriate to treat KS4 attainment as a continuous measure and it would also provide limited gains in statistical power.

At protocol, it was anticipated that BKSB/ ForSkills scores could be used as a measure of baseline academic achievement. Further discussion with colleges revealed that this would only be available for a small minority of students. KS2 attainment is recorded differently for English and maths and would not therefore provide a consistent baseline measure.¹⁰

Primary outcome

The primary outcome is achieving a passing grade (9–4) in GCSE English or maths. This outcome is most appropriate as the intervention targets students who need to re-sit GCSE English or maths (having achieved below grade four at the end of KS4). Students re-sitting both English and maths were randomly selected to receive the intervention for one subject and only the outcome in this subject is analysed for such students. Students re-sitting different subjects are pooled into their intervention group, rather than being analysed separately. For example, students re-sitting maths in the 'texts to students' intervention arm are analysed alongside students re-sitting English.

As anticipated in the Statistical Analysis Plan (Smith and Ashworth, 2018) attainment is assessed as a binary outcome as the distribution of grades is clustered around the pass/fail threshold (assessed on a 1–9 scale where achieving 4 or above is defined as a pass). It is therefore not possible to robustly analyse attainment as a continuous outcome.

Attainment data in English (YPMAD_GCSE_GRADE_ENG) and maths (YPMAD_GCSE_GRADE_MATHS) was sourced from the 'Young Person's Matched Administrative Data' (YPMAD) chronological dataset from the National Pupil Database (NPD) for the 2017/2018 academic year.¹¹ In a small number of cases where this information was missing, the outcome was sourced from the KS4 dataset for 2017/2018 (GCSE_ENG_94 and GCSE_MATH_94). As outlined in the SAP, pupils who did not attempt the exam (that is, no grade is recorded, but they were successfully linked to the NPD) are coded as failing. Trial data was linked to NPD pupil records using seven data identifiers: first name, last name, date of birth, Unique Learner Number (ULN), Unique Reference Number (UPN), school name, and pupil postcode. All personal data was removed before the matched dataset was transferred to the ONS Secure Research Service for analysis.

¹⁰ No baseline attainment was specified in the SAP. In theory, it would be possible to use a KS2 attainment measure for English and Maths at randomisation. However, this was not feasible in reality because of the long time period between finalising the recruitment and receipt of KS2 scores from an NPD data request. Using KS2 at randomisation would have delayed the start of the intervention by around 3 to 6 months meaning the intervention period would cover only a small proportion of the academic year. It is possible to use KS2 attainment in the final model, but this covers long period of education between leaving primary school at 7 and re-sitting GCSEs at 17.

¹¹ This dataset does not distinguish between results from a November resit against resits at any other point in the academic year.

Secondary outcomes

Previous studies of similar interventions (Miller et al., 2016; Groot et al., 2017; Kraft and Rogers, 2015) identify attendance as an interim outcome. These studies found that communication through texts improved attendance and this in turn is hypothesised to improve attainment. Therefore, attendance at college is considered an interim output of the intervention and is assessed as a secondary outcome.

Attendance in the 2017/2018 academic year is assessed as a secondary outcome, collected directly from college registers over the entire academic year. The outcome is assessed as the number of English or maths classes as a proportion of all possible classes, to ensure comparability across colleges.

Sample size

All power calculations were conducted in PowerUp! (Dong and Maynard, 2013), with the MDES transformed into a percentage impact using the probability of success in control. This was estimated by dividing the effect size by the pooled standard deviation to estimate the proportion anticipated to achieve success in the intervention group (see Smith and Ashworth, 2018, p. 6). Separate power calculations were estimated for the effect of the intervention on FSM pupils by collapsing the three interventions into a single trial arm. However, the FSM analysis is exploratory only.

Protocol

At protocol (Griggs, 2017), it was assumed that each of the 30 FE colleges would have recruited 125 eligible students (approximately 31 per trial arm), totalling approximately 3,750 students. It was assumed that there would be no college-level attrition, but student-level attrition of 25%. This assumption was based on the 2016 Statistical First Release (SFR) for level 1 and 2 attainment in English and maths by students aged 16–18, which indicated that on average 81% of GCSE maths entrants and 88% of GCSE English entrants in FE colleges re-sit GCSE exams. However, as the data covered the academic year 2014/2015, these percentages preceded the recent government requirements for all students who obtain a GCSE grade D (grade 3) at KS4 to re-sit GCSEs in post-16 settings. Therefore, a slightly higher level of attrition was anticipated as the 2017/2018 cohort includes students that would not have otherwise volunteered to re-sit their GCSEs.

The pass rate was assumed to be 30% in the control group (Joint Council for Qualifications, 2018), 80% power, a type-one error rate of 0.05,¹² and two-tailed significance testing. The level of clustering was anticipated to be low, with an ICC of 0.10 and no anticipated treatment heterogeneity.

The sample size calculations conducted at protocol were estimated on the basis of being able to detect an impact of a 6.8 percentage point increase in the GCSE pass rate, from an estimated 30% to 37%.

Statistical analysis plan

The assumptions used in the SAP were similar to those at protocol. The control group pass rate was assumed to be 30%, 80% power, a type-one error rate of 0.05, two-tailed significance testing, an ICC of 0.10, and no treatment effect heterogeneity.

The sample size calculations were adjusted at SAP stage to account for an additional college which the developers were able to recruit. The randomised sample includes 3,823 students (including 44 pupils randomised in error) and the calculations were also updated to reflect this. If the pass rate for the control group was 30%, the analysis would be powered to detect a statistically significant difference between the intervention and control groups if the pass rate in the intervention group was 36.8% or higher.

The SAP also included separate power calculations estimated for pupils eligible for free school meals (everFSM_6_P). It was assumed that 20% of pupils in the trial would have been eligible for free school meals. This assumption is based on the proportion of pupils eligible for FSM being higher than the national average amongst pupils re-sitting GCSEs in English or maths. Adjusting for this resulted in an MDE of 10.7 percentage points.

¹² Bonferroni adjusted to account for multiple treatment arms (0.0167 for each trial arm).

Analysis

The power calculations at analysis followed similar assumptions: 80% power, a type-one error rate of 0.05, two-tailed significance testing, and no treatment effect heterogeneity. The ICC was estimated as 0.05 with a control group pass rate of 21.7%.

After accounting for those not meeting the eligibility criteria ($n = 44$) and students missing re-sit data in the NPD extract ($n = 33$), the final analytic sample contained 31 colleges with 3,746 students, yielding an MDE of 6.4 percentage points.

At analysis stage, a total of 35.8% of pupils were ever eligible for free school meals by the end of KS4, yielding an MDE of 10.7 percentage points. As the student sample could include students aged up to 25, `everFSM_6_P` is not consistently recorded for all pupils in the sample and therefore an alternative definition was selected at analysis: eligibility for FSM at the end of KS4.

Randomisation

Randomisation was conducted as a multi-site block randomisation, with students allocated to trial arms on a 1:1:1:1 basis. The block variable was FE colleges to ensure that there were equal students in each trial arm in each college. It had been agreed at the outset that pupils re-sitting GCSEs in both English and maths would only have one subject assessed as an outcome for the trial. The outcome for these pupils was selected at random. This approach was taken to prevent within-participant spill-over effects. For example, a person receiving a reminder text about English might also act upon the reminder to do further work on their maths. Therefore, the maths result for that pupil should not be used as an observation in the control group as this could bias the impact estimate.

Randomisation to each of the four treatment arms was conducted in Stata MP, Version 14.1, using the `randomize` command, with a maximum of 100 randomisations used to achieve balance on the subject taken. To ensure the randomisation was replicable, a stable seed was set using a random number from the website `random.org`. The randomisation was conducted in October 2017 by an independent, blinded analyst within the evaluation team. The full randomisation syntax is provided in the supplementary appendix.

Statistical analysis

Primary analysis

The primary analysis was conducted on an intention-to-treat (ITT) basis. The attainment measure—obtaining a pass or fail in a GCSE re-sit of English or maths—was assessed as a binary outcome.¹³ Following EEF guidance (EEF, 2018), the outcome was therefore estimated as a relative-risk ratio (RRR). The RRR was estimated from a multilevel logistic regression, with students at level one, nested within colleges at level two. A multilevel logistic regression was selected as an appropriate technique to account for the clustering of students within colleges. Separate models will be estimated for each treatment arm against control as follows:

$$\text{logit}(\pi)_{ij} = \beta_{0ij} + \beta_1 \text{treatment}_{ij} + e_{ij} + r_j$$

Where treatment_{ij} is a binary indicator of treatment allocation, e_{ij} is the student-level error term, and r_j is the random effect at college level. This was estimated using the `melogit` command in Stata 14.

No assessment of relative effects (that is, comparing intervention arms against each other) will be undertaken as it is unlikely that the trial would have enough power to detect such an effect. No alternative model specifications are considered.

Secondary analysis

The distributions of attendance (see supplementary appendix) indicated a heavily skewed distribution making the multilevel linear regression model proposed in the analysis plan inappropriate to estimate the effect of the intervention.

¹³ Students who do not re-sit the exam are classified as a 'fail', in line with the intention-to-treat approach.

In a deviation from the trial analysis plan, the impact was estimated using negative binomial regression. This is most appropriate for count data with the distributions observed. The model was analysed as follows:

$$nbinomial(\pi)_{ij} = \beta_{0ij} + \beta_1 treatment_{ij} + e_{ij} + r_j$$

Separate models were estimated for each intervention arm against control. This was estimated using the `menbreg` command in Stata 16. No baseline measure of attendance is available as not all students were in attendance at the same college in 2017/2018 as during 2016/2017 and students in the evaluation are of a range of ages. Most children are aged 16–19, though there were students as old as 25 involved in the intervention. It is therefore unlikely that a baseline measure of attendance would be consistently recorded for all students in the evaluation.

Analysis in the presence of non-compliance

As was discussed in the Statistical Analysis Plan (Smith and Ashworth, 2018), it is difficult to establish in advance a clear threshold of the dosage required for the intervention to be effective. The SAP defined compliance as ‘receiving any less than all’ texts intended for the individual student and/or their study supporter.

In practice, relatively few individuals were sent all texts intended for them (students only 37%, supporters only 37%, students and study supporters 19%) and, as Gerber and Green (2012) argue, ‘classifying partially treated as untreated leads to a violation of the exclusion restriction when partial treatment affects the outcome’. Therefore, the analysis has also been conducted using a range of different thresholds to test the sensitivity to this assumption, as suggested by the SAP. In addition to receiving 100% of texts, analysis is also conducted for 70%, 80%, and 90% of intended texts sent to the students and/or their study supporter. This data was provided by BIT using *Promptable*. The proportion of students defined as ‘compliant’ for different thresholds is explored in greater detail in the impact evaluation chapter.

The SAP assumed that non-compliance was likely to be limited to the intervention arms (one-sided non-compliance) as it was perceived that students in the control arm would be unlikely to seek out supportive texts in a way that mirrors the intervention. This is also reflected in the evaluation design, where students were randomised within colleges.

However, the data provided by the delivery partner on how many texts were sent to students and/or their study supporters indicates that some students in the control arm were sent texts (7% of control arm: 11% of students only, 11% of study supporters only, see Impact Evaluation for more information). In a limited number of cases ($n = 275$), students self-nominated by providing the same telephone number for themselves and their study supporter. It is also possible that students may have cross-nominated peers who were allocated to a different intervention arm, exposing them to text messages they should not have received.¹⁴

The revised definition of compliance is therefore ‘receiving a sufficient proportion of texts’ (for different thresholds, 70%, 80%, 90%, and 100%) and that the student and their nominated study supporter did not receive any texts that violated their allocation. It is worth considering that although the data indicates message delivery, this does not mean that students or their study supporters read the content of the message.

The contamination of trial arms is a limitation of the evaluation and should be considered when assessing the security of the evaluation’s findings.

The contamination also means that the intervention suffers from two-sided non-compliance, rather than the one-sided non-compliance outlined in the SAP. Consequently, the analytical approach outlined therein is no longer suitable. Instead, non-compliance with the intervention was analysed using an Instrumental-Variables (IV) approach (Angrist and Imbens, 1995), where random assignment was used as an instrument for receipt of the intervention.

As specified in the EEF Statistical Analysis Guidance (EEF, 2018), tests for endogeneity were conducted to accompany the IV-regression (Wooldridge, 1995). These consistently indicated—across trial arms and for different dosage thresholds—that the instrumented variable was exogenous rather than endogenous. The results of these tests are presented in the supplementary appendix.

¹⁴ The available data was not recorded in such a way to test this hypothesis.

Consequently, non-compliance was analysed using a multilevel logistic regression:

$$\text{logit}(\pi)_{ij} = \beta_{0ij} + \beta_1 \text{compliant}_{ij} + e_{ij} + r_j$$

Where compliant_{ij} is a binary indicator of whether the student was sent a sufficient number of texts for the dosage threshold and that they (or their nominated study supporter) did not receive texts that violated their allocation. All other terms are the same as the primary analysis model.

The SAP also suggested that it might be possible to estimate the ‘average causal response with variable treatment Intensity’ (Angrist and Pischke, 2009, section 4.5.3). This would require a sufficient number of students across the distribution of texts sent. However, the distribution of texts sent is heavily skewed, preventing the creation of meaningful groups with a sufficient sample size to conduct this analysis.

Missing data analysis

The Statistical Analysis Plan (Smith and Ashworth, 2018) outlined conditions under which missing data analysis would be conducted. As the primary outcome was sourced from the National Pupil Database (NPD) it was anticipated that the proportion of missing data would be low. Following EEF guidance (EEF, 2018), no missing data analysis was proposed if the proportion of students missing primary outcome data was less than 5%.

The proportion of students missing outcome data was very low. In total, 33 of 3,823 (<1%) students randomised were missing the primary outcome for their relevant subject. Therefore, no missing data analysis was conducted.

Subgroup analyses

There were four subgroups for this evaluation, which were:

- ever eligible for free school meals by end of KS4;
- gender;
- subject (English or maths); and
- studying for a single re-sit (maths or English) or re-sitting both English and maths.

In a deviation from the evaluation protocol (Griggs, 2017) and the SAP, the definition of free school meals has changed—pupils must be in primary or secondary education to be eligible for free school meals. As the student sample could include students aged up to 25, everFSM_6_P is not consistently recorded for all pupils in the sample in the year they took part in the intervention. Therefore, an alternative definition was selected: eligibility for FSM at the end of KS4, or in the six preceding years.

The subgroup analysis for FSM eligibility combined the samples from each treatment arm to form a single treatment group. The first model used just the sample of those eligible for free school meals, using the same model specification as the primary analysis. A second model including a binary indicator of free school meal eligibility, interacted with treatment indicator will test for a differential impact on FSM recipients. The specifications are as follows:

$$\text{logit}(\pi)_{ij} = \beta_{0j} + \beta_1 \text{treatment}_{ij} + e_{ij} + r_j \text{ (FSM sample only)}$$

$$\text{logit}(\pi)_{ij} = \beta_{0j} + \beta_1 \text{treatment}_{ij} + \beta_2 \text{everFSM}_{ij} + \beta_3 \text{everFSM} * \text{treatment}_{ij} + e_{ij} + r_j$$

Where everFSM_{ij} is a binary indicator of whether the student was ever eligible for free school meals at the end of KS4 or in the previous six years. Note that $\beta_3 \text{everFSM} * \text{treatment}_{ij}$ indicates the interaction term between random allocation and FSM eligibility.

The remaining three subgroup analyses used the interaction method described above, resulting in the following model specifications:

$$\text{logit}(\pi)_{ij} = \beta_{0j} + \beta_1 \text{treatment}_{ij} + \beta_2 \text{gender}_{ij} + \beta_3 \text{gender} * \text{treatment}_{ij} + e_{ij} + r_j$$

$$\text{logit}(\pi)_{ij} = \beta_{0j} + \beta_1 \text{treatment}_{ij} + \beta_2 \text{subject}_{ij} + \beta_3 \text{subject} * \text{treatment}_{ij} + e_{ij} + r_j$$

$$\text{logit}(\pi)_{ij} = \beta_{0j} + \beta_1 \text{treatment}_{ij} + \beta_2 N_subjects_{ij} + \beta_3 N_subjects * \text{treatment}_{ij} + e_{ij} + r_j$$

Additional analyses and robustness checks

The SAP specified that sensitivity analysis would be conducted by expanding the primary analysis to include any covariates that were statistically significant between intervention and control arms as randomised. A single characteristic, the number of re-sits a student was taking (that is, both English and maths), was more prevalent in the 'supporters only' and 'students and supporters' interventions arms, relative to the control group (statistically significant). This was therefore included in the sensitivity analysis model.

Estimation of effect sizes

The primary outcome is binary and therefore the results from the primary analysis were estimated as relative risk ratios, in line with EEF analysis guidance (EEF, 2018). This was estimated from the `melogit` command in Stata 14, using the `or` option to present estimates as odds ratios. These were then transformed into relative risk ratios using the following formula:

$$RRR = \frac{OR}{(1 - p + (p \times OR))}$$

Where p is the expected GCSE re-sit pass rate in the control group. This was estimated using the `margins` post estimation command in Stata 14.

The odds ratios were also converted to Cohen's d as follows:

$$d = \log(OR) \times \frac{\sqrt{3}}{\pi}$$

Finally, it was converted to Hedge's g by multiplying by the correction factor J :

$$J = 1 - \left(\frac{3}{4(N^T + N^C - 2) - 1} \right)$$

Attendance is a continuous outcome and is therefore presented as a Hedge's G effect size, in accordance with EEF guidance (EEF, 2018). These were estimated using the formulae presented below, from Hedges (2007).

The Hedge's g effect size will be estimated following Hedge's (2007) formulae for the effect size d_t for designs with unequal sample sizes. The effect size, g_t is estimated as follows:

$$g_t = J \times \left(\frac{\bar{Y}_{\blacksquare\blacksquare}^T - \bar{Y}_{\blacksquare\blacksquare}^C}{S_T} \right) \sqrt{1 - \rho \left(\frac{(N - n_u^T m^T - n_u^C m^C) + n_u^T + n_u^C - 2}{N - 2} \right)}$$

Where:

- $\bar{Y}_{\blacksquare\blacksquare}^T$ and $\bar{Y}_{\blacksquare\blacksquare}^C$ are the grand means of the treatment and control groups
- ρ is the intra-cluster correlation
- N is the total number of pupils
- M is the total number of colleges, divided between the intervention group m^T and the control group m^C

The remaining terms are calculated as follows:

The correction factor J is defined as:

$$J = 1 - \left(\frac{3}{4(N^T + N^C - 2) - 1} \right)$$

The pooled standard deviation, S_T is defined as:

$$S_T = \sqrt{\frac{\sum_{i=1}^{m^T} \sum_{j=1}^{n_i^T} (Y_{ij}^T - \bar{Y}_{\blacksquare\blacksquare}^T)^2 + \sum_{i=1}^{m^C} \sum_{j=1}^{n_i^C} (Y_{ij}^C - \bar{Y}_{\blacksquare\blacksquare}^C)^2}{N - 2}}$$

And the term n_u^C is defined in the same way as n_u^T :

$$n_u^T = \frac{(N^T)^2 - \sum_{i=1}^{m^T} (n_i^T)^2}{N^T(m^T - 1)}$$

The variance term is calculated as follows:

$$V\{g_t\} = \left(\frac{N_T + N_C}{N_T N_C}\right) (1 + (\tilde{n} - 1)\rho) + \frac{[(N - 2)(1 - \rho)^2 + A\rho^2 + 2B\rho(1 - \rho)]\delta_g^2}{2(N - 2)[(N - 2) - \rho(N - 2 - B)]}$$

Where:

$$\tilde{n} = \frac{N^C \sum_{i=1}^{m^T} (n_i^T)^2}{N^T N} + \frac{N^T \sum_{i=1}^{m^C} (n_i^C)^2}{N^C N}$$

And:

$$A = A^T + A^C$$

Where A^C is calculated in the same way as A^T :

$$A^T = \frac{(N^T)^2 \sum_{i=1}^{m^T} (n_i^T)^2 + \left(\sum_{i=1}^{m^T} (n_i^T)^2\right)^2 - 2N^T \sum_{i=1}^{m^T} (n_i^T)^3}{(N^T)^2}$$

B can be calculated as follows:

$$B = n_u^T(m^T - 1) + n_u^C(m^C - 1)$$

Confidence intervals:

Finally, confidence intervals for a two-tailed test are calculated as follows:

$$\delta_g - c_{\frac{\alpha}{2}} v_g \leq \delta_g \leq \delta_g + c_{\frac{\alpha}{2}} v_g$$

Where $c_{\frac{\alpha}{2}}$ is the critical value for a type one error rate, α , given the sample size.

Estimation of ICC

Clustering was accounted for using a multilevel modelling approach. The intra-cluster correlation, ρ_s , for each model is reported, using the following formula:

$$\rho_s = \frac{\sigma_{BS}^2}{\sigma_{BS}^2 + \sigma_{WS}^2}$$

Where σ_{BS}^2 is the between-college variance and σ_{WS}^2 is the within-college variance. The ICCs were estimated using the `estat icc` post-estimation command in Stata 14. ICC estimates for all the models in this report are presented in the supplementary appendix.

Implementation and process evaluation

Research methods

The implementation and process evaluation (IPE) was designed to explore how the intervention was implemented and delivered, and, in turn, how, why, and for whom it works.

The research questions for the IPE, as reported in the protocol were:

1. How was the intervention implemented?
2. How was the intervention interpreted and delivered across colleges and trial arms?
3. To what extent did contextual variation affect fidelity?
4. What adaptations (if any) were put in place?
5. To what extent did students and study supporters engage with and act upon the messages they received?
6. What were the barriers to delivery and how were these addressed?
7. What facilitated successful delivery?
8. What was the cost of delivery?

The IPE was designed as a mixed-methods study. Fieldwork was carried out between September 2017 and August 2018 and included the research methods set out in Table 4 and explained below.

Table 4: IPE methods overview

Research methods	Data collection methods	Participants/ data sources	Data analysis methods	Research questions addressed	Implementation/logic model relevance
Observations	Observation	Tutor workshop (1)	Thematic analysis	RQ1 RQ3; RQ4	Fidelity; Adaptation
Surveys	Student sign-up survey	Students (7,004)	Descriptive statistics	RQ1	Reach
Interviews	In-depth interviews (pre-intervention)	Project leads (24)	Framework deductive/inductive charting; thematic analysis	RQ1; RQ2	Fidelity; Programme differentiation
Interviews	In-depth interviews (post-intervention)	Project leads (21)	Framework deductive/inductive charting; thematic analysis	RQ1; RQ2; RQ3; RQ4; RQ5; RQ6; RQ7	Fidelity; Adaptation; Responsiveness
Interviews	In-depth interviews	Study supporters who opted-out (8)	Framework deductive/inductive charting; thematic analysis	RQ5; RQ6	Reach
Case studies (6; case study unit = college; analytical approach = cross-case analysis)	In-depth interviews	Maths/English leads (12)	Framework deductive/inductive charting; thematic analysis	RQ1; RQ2; RQ3; RQ4; RQ5; RQ6; RQ7	Fidelity; Adaptation; Responsiveness
	In-depth interviews	Students (38)	Framework deductive/inductive charting; thematic analysis	RQ5	Responsiveness
	In-depth interviews	Study supporters (21)	Framework deductive/inductive charting; thematic analysis	RQ5	Responsiveness

Observation of a tutor workshop

The NatCen team attended one of the informational workshops provided by BIT to college English and maths tutors in July 2017. Attendance provided the evaluation team with insight into the delivery of the programme, including important information about how the programme is implemented in the earliest stages. Information gained from this workshop informed the development of research tools such as the topic guide for interviews.

BIT/NatCen sign-up and diagnostic survey

Students completed a short diagnostic survey as part of the online recruitment process. This survey covered a range of themes relating to their studies, including personal motivation, the extent to which they were engaged with their studies and were responsible learners, and their interest in the subject(s) taken.

This information was used to compare the characteristics of students agreeing to take part in the intervention with those who refused. The results of this comparison offered important contextual information for the process evaluation, specifically, a better understanding of the possible limitations of the intervention to reach to sections of its target group.

In total, 7,004 students completed the survey.

Telephone interviews with project leads

Two rounds of interviews were carried out with the project lead in participating colleges. This was the member of staff designated the point of contact for the trial at the recruitment stage (unless this role changed over the course of the intervention).

The first interview was conducted with all project leads by telephone between October and December 2017. The aim of this round of interviews was to:

- understand planned 'business as usual' in each college, specifically in relation to improving attendance and attainment in GCSE re-sit exams;
- explore college project leads' motivations for taking part in Project Success, their views on the early stages of intervention implementation—including student recruitment and tutor workshop, and expected benefits of participation; and
- discuss college management information (MI) on attendance: as the format and mode of MI data collected by colleges was likely to vary, this early scoping discussion allowed the team to confirm that the data required was available in a workable format.

Towards the end of the trial, all project leads were invited to take part in a follow-up interview in summer 2018. The aim of the follow-up interviews was to explore any issues relating to implementation, fidelity, and changes to business as usual, as well as to collect data on costs. Interviews were also used as an opportunity to facilitate collection of attendance data needed for the impact evaluation.

Telephone interviews with study supporters opting out of the intervention¹⁵

Study supporters who opted out during the first six weeks of the intervention were sent a follow-up text message asking if they would be willing to be contacted by NatCen to take part in a short telephone interview. Overall, 60 study supporters who opted out also agreed to be re-contacted. Of these, eight took part in a 15-minute telephone interview. Interviews were used to explore the reasons study supporters had elected to opt out. This included understanding the types and quality of relationships between students and study supporters who opted out, any issues with the format and content of the text messages, and supporters' level of engagement with Project Success.

College case studies

To build a broader understanding of the implementation and variation in delivery, college-based case studies were conducted. Information gathered as part of the initial interviews with project leads was used to purposively select a sample of six case study colleges. Case studies were organised as a day visit to each college in spring 2018. However, in many cases, interviews with students and study supporters were conducted by telephone after the visit had taken place.

Each case study included:

- Interviews with the maths/English lead(s) who worked with intervention developers to define the content of texts. Interviews explored views on initial implementation, enablers and barriers, adaptation, and any perceived impact on student outcomes.
- Interviews with two students in each active trial arm. The student sample included ten students per treatment arm for each college, selected to achieve a balance of GCSE subjects (that is, five English and five maths) and a mix in terms of gender and relation with the study supporter. Students were recruited via email and follow-up phone calls. Interviews explored their attitudes and impressions of the text messages, whether they felt the intervention had changed their behaviour, improved motivation, or encouraged attendance, and whether they shared the information included in the texts with their peers.
- Interviews with two study supporters in each relevant trial arm. Supporters were matched to student interviewees wherever possible. Corresponding interviews with supporters focused on the effect of texts on their own behaviour, whether they encouraged interaction with the student, what form this interaction took, as well as any perceived impacts on student behaviour (including negative effects).

Conduct of interviews

The content of each interview was based on a topic guide to ensure systematic coverage of key themes that addressed the IPE research questions. It was intended to be flexible and interactive, allowing issues of relevance to be covered

¹⁵ Overall, 112 students and 90 supporters opted out during the intervention

through detailed follow-up questioning. Evidence of intermediate outcomes included in the logic model was sought through targeted probing during interviews with students, study supporters, and college project and subject leads. In this way, it was possible to collect and analyse information on perceived intermediate outcomes (such as behavioural change) systematically. Any evidence in support of the logic model is included in the IPE findings.

Analysis

Interviews were digitally recorded and professionally transcribed. Framework, a systematic approach to qualitative data management and analysis developed by NatCen, was used to chart (collate and summarise) and analyse transcribed data. Charting in Framework involves sorting and synthesising data by theme and case thereby helping to identify and explain patterns in the data. In this project, a mixed deductive/inductive approach was employed for the charting of transcripts, with data being synthesised according to both pre-established themes as set out in the topic guides and emerging ones. Thematic analysis aimed at identifying and explaining patterns in the data in relation to these themes. Anonymised quotes from students, study supporters, and college project and subject leads were selected to illustrate key points and included throughout the IPE section.

SPSS v21 was used to analyse student survey data and syntax files were used to ensure a record was kept of how analysis was conducted.

Costs

NatCen provided a template for colleges to use over the course of the project to log information on direct costs incurred and time spent on the intervention. This cost sheet was collected from all colleges in summer 2018. Cost data was also collected from the delivery team at BIT between December 2018 and February 2019.

The total cost was calculated based on set-up and delivery costs provided by the development team at BIT and participating colleges. The calculated total cost over three years was based on the sum of running costs (for example, text message costs—assumed to occur every year for three years) multiplied by three (three years) added to the sum of set-up costs (for example, training costs—assumed to occur once every three years). The cost per pupil per year—averaged over three years—was calculated by dividing the total cost over three years by three and then dividing this number by the average number of participating students per college ($n = 92.26$).

Timeline

Table 5: Timeline

Dates	Activity	Staff responsible / leading
April to June 2017	Preparation of materials, finalisation of eligibility criteria and outcome measures agreed	BIT and NatCen
May 2017	Ethical approval for evaluation	NatCen
May 2017	IDEA Workshop	BIT and NatCen
May to July 2017	College recruitment, MOUs signed	BIT
July 2017	Protocol produced	NatCen
September 2017	Student and study supporter data collection	BIT
September 2017	Pre-randomisation interviews with college project leads	NatCen
October 2017	Multi-site individual-level randomisation	NatCen
November 2017 to June 2018	Intervention—texting students and study supporters delivered in 31 FE colleges	BIT
November to December 2017	Process evaluation—telephone interviews with study supporter opt-outs	NatCen
April to May 2018	Process evaluation—case study research in six FE colleges	NatCen
May to August 2018	Process evaluation—follow-up interviews with project leads, MI and cost data collection	NatCen
September to October 2018	Qualitative data management and analysis	NatCen
October 2018	GCSE re-sit outcome data request submitted	NatCen
September 2019	GCSE re-sit outcome NPD data delivered	DfE
	Impact analysis	NatCen
November 2019	Draft report submitted	NatCen

Impact evaluation

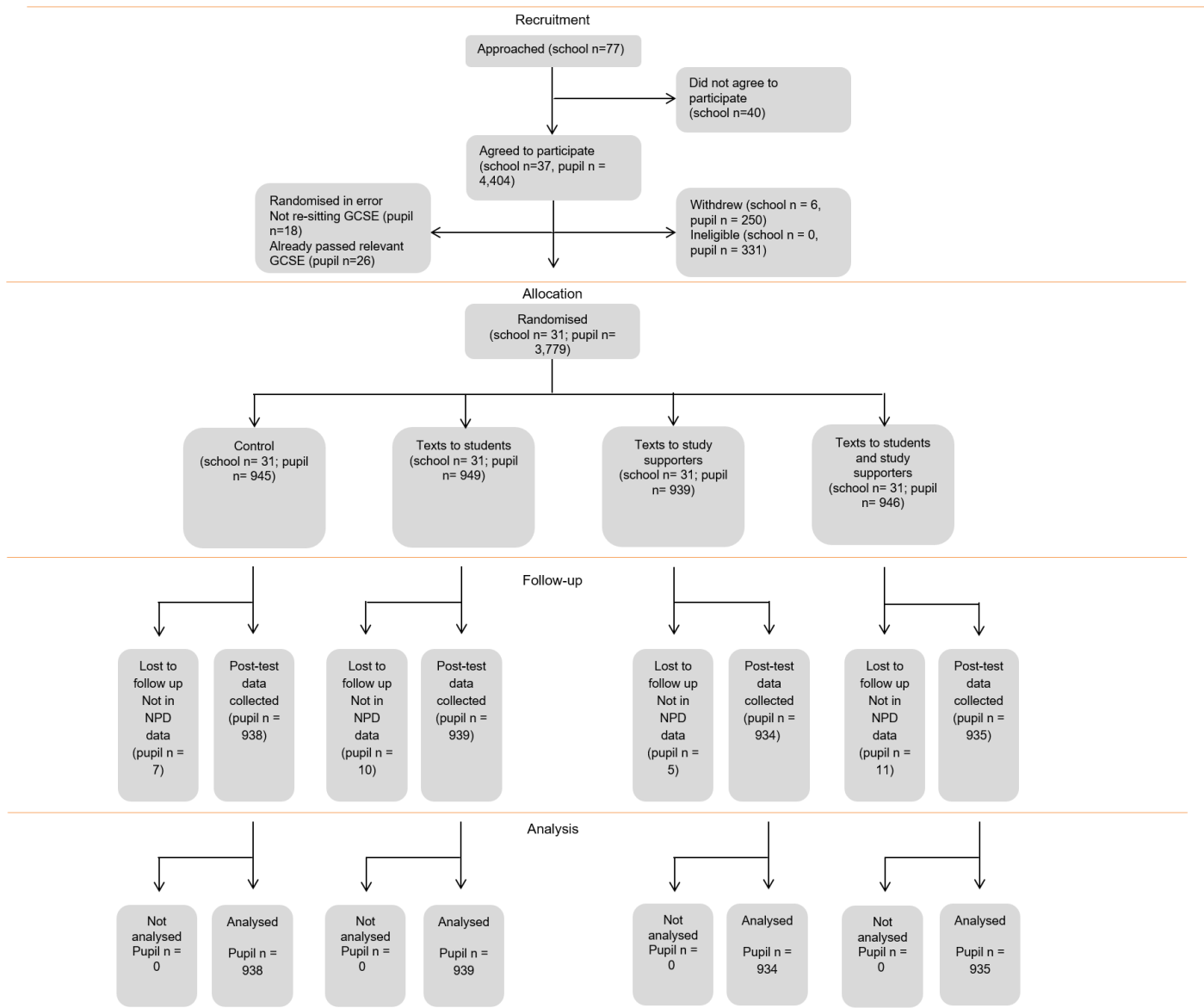
Participant flow including losses and exclusions

In total, 37 colleges were recruited by the delivery partner for the evaluation, though six colleges dropped out prior to randomisation. The remaining 31 colleges included 3,823 eligible students who had agreed to take part. These students were subsequently randomised to one of four trial arms. The distribution of students allocated to trial arms within college is reported in full in the Statistical Analysis Plan (Smith and Ashworth, 2018, p. 5). Some students were randomised in error, either because they were not enrolled to re-sit a relevant GCSE ($n = 18$) or because they had already passed a GCSE in the subject they were re-sitting ($n = 26$).¹⁶ These are indicated in the CONSORT diagram as 'randomised in error' and are not included in estimates of attrition.

Attrition was relatively low in all four trial arms with no students withdrawing after randomisation (though some students had missing outcome data). As there was no baseline data collection and student outcomes are recorded in the National Pupil Database, relatively few students' outcome data was missing ($n = 33$). Attrition was 0.7% for the control group, 1.1% for the students only group, 0.5% for the study supporters only group, and 1.2% for the students and study supporters group. The attrition (0.9%) is entirely attributed to the matching of trial data to the NPD. In these rare cases, the students could not be identified in the NPD. This could have been due to missing identifier data used to match individuals' trial data to the NPD.

¹⁶ Those not enrolled in an eligible GCSE were identified in their responses to the baseline survey. Those who had already passed their relevant GCSE were identified from the NPD data.

Figure 2: CONSORT diagram



The power calculations in Table 6 indicate relatively similar Minimum Detectable Effects (MDE) at analysis as was expected in the statistical analysis plan. Slightly smaller than anticipated sample sizes have been compensated by a lower level of clustering than anticipated. Consequently, the primary analysis is powered to detect a minimum detectable effect of a 6.4 percentage point impact, whilst the subgroup analysis is powered to detect a 10.7 percentage point impact.

Note that the power calculations are assessed for control against an individual intervention arm (and adjusted for multiple hypothesis testing); this is because the primary analysis analyses the trial arms separately. There is no substantive difference in the MDES across trial arms.

Table 6: Minimum detectable effect size at different stages

		Protocol		Randomisation		Analysis	
		Overall	FSM	Overall	FSM	Overall	FSM
MDE (%)		6.8	6.5	6.8	5.7	6.4	5.6
MDE (Hedges g equivalent)		0.15	0.14	0.15	0.12	0.15	0.14
Pre-test/post-test correlations	Level 1 (pupil)	-	-	-	-	-	-
	Level 2 (college)	-	-	-	-	-	-
Intracluster correlations (ICCs)	Level 2 (college)	0.10	0.10	0.10	0.10	0.05	0.05
Alpha		0.0167	0.05	0.0167	0.05	0.0167	0.05
Power		0.8	0.8	0.8	0.8	0.8	0.8
One-sided or two-sided?		2	2	2	2	2	2
Average cluster size		62	19	60	24	59	21
Number of colleges	Total	30	30	31	31	31	31
Number of pupils	Intervention	937	430	930	560	934	792
	Control	937	140	930	190	938	267
	Total:	1,874	570	1,860	750	1,872	1,059

Attrition

The overall rate of attrition for this trial was very low at less than 1% overall. This is likely a consequence of sourcing the primary outcome from the National Pupil Database (NPD) and not undertaking baseline data collection. Our primary outcome definition means that only those pupils with no record in the YPMAD NPD dataset are classified as missing the outcome (and hence being counted as attrition).

Table 7: Pupil level attrition from the trial (primary analysis)

		Control	Students only	Supporters only	Students and supporters	Total
Number of pupils	Randomised	945	949	939	946	3,779
	Analysed	938	939	934	935	3,746
Pupil attrition (from randomisation to analysis)	Number	7	10	5	11	33
	Percentage	0.7	1.1	0.5	1.2	0.9

Pupil and college characteristics

Table 8: Baseline balance as randomised

Pupil level (categorical)	Control		Student only		Supporter only		Student and Supporter	
	n/N (missing)	Count (%)	n/N (missing)	Count (%)	n/N (missing)	Count (%)	n/N (missing)	Count (%)
English Maths	466/945 (0) 479/945 (0)	49.3 50.7	471/949 (0) 478/949 (0)	49.6 50.4	481/939 (0) 458/939 (0)	51.2 48.8	482/946 (0) 464/946 (0)	51.0 49.0
Two-re-sits	354/945 (0)	37.5	376/949 (0)	39.6	408/939 (0)	43.5	397/946 (0)	42.0
Ever eligible for free school meals (KS4)	267/747 (198)	35.7	275/738 (211)	37.2	264/749 (200)	35.3	253/728 (218)	34.8
Female	481/939 (6)	51.2	495/939 (10)	52.7	512/933 (6)	54.9	506/941 (5)	53.8

Table 9: Baseline balance as analysed

Pupil level (categorical)	Control		Student only		Supporter only		Student and Supporter	
	n/N (missing)	Count (%)	n/N (missing)	Count (%)	n/N (missing)	Count (%)	n/N (missing)	Count (%)
English Maths	464/938 (0) 474/938 (0)	49.5 50.5	467/939 (0) 472/939 (0)	49.7 50.3	480/934 (0) 454/934 (0)	51.4 48.6	478/935 (0) 457/935 (0)	51.1 48.9
Two re-sits	349/938 (0)	37.2	374/939 (0)	39.8	407/934 (0)	43.6	392/935 (0)	41.9
Ever eligible for free school meals (KS4)	267/747 (191)	35.7	275/738 (201)	37.3	264/739 (195)	35.7	253/728 (207)	34.8
Female	479/933 (5)	51.3	489/931 (8)	52.5	510/930 (4)	54.8	502/932 (3)	53.9

Table 8 and Table 9 indicate that students' characteristics such as free school meal eligibility, gender, and subject being assessed are balanced across trial arms in both the 'as randomised' and 'as analysed' samples. Students who were re-sitting both English and maths were more likely to have been in the 'supporter only' and 'student and supporter' trial arms; they may also have had lower attainment overall prior to the evaluation and may have a greater workload relative to students re-sitting a single subject, which could potentially reduce the likelihood of a pass in either subject. This should be considered when interpreting the trial findings.

Around one third of students from each trial arm were eligible for free school meals at the end of KS4 or in the preceding six years. This reflects the attainment gap between disadvantaged and non-disadvantaged pupils in GCSE attainment when compared with the national average of 15.1% (Department for Education, 2018). There were no differences in the proportion of students eligible for free school meals between trial arms.

As this evaluation randomised at individual level, blocked by college, the characteristics of colleges do not vary between trial arms. Table 10 therefore presents the relevant characteristics of colleges in the evaluation. There was no college-level attrition between randomisation and analysis and hence these statistics are the same at both stages of the evaluation.

Table 10: Baseline balance in college characteristics

College level (categorical)	n/N (missing)	Count (%)
Institution type: colleges	31/31 (0)	100.0
Urban Rural	30/31 (0) 1/31 (0)	96.8
Ofsted Rating: Good Requires improvement	13/17 (14) 4/17 (14)	76.5 23.5

All students were studying in colleges and almost all colleges were based in urban conurbations. Ofsted rating was missing for approximately half (45.2%) of colleges. Three quarters (76.5%) of colleges with an Ofsted rating were assessed as 'good'.

Outcomes and analysis

Primary analysis

Table 11 presents the results of the primary analysis. The GCSE re-sit pass rate in English or maths ranges from 20.0% to 22.5% for different trial arms. There are no statistically significant differences in the unadjusted proportions achieving a pass grade by trial arm.

Table 11: Primary analysis—impact of the intervention on GCSE re-sits in English or maths

Group	Unadjusted pass rate				Effect size			
	Intervention group		Control group		Total n (intervention; control)	Relative Risk Ratio (95% CI)	Hedge's g (95% confidence interval)	p- value ¹⁷
n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)					
Texts to students	939 (10)	22.4 [19.7; 25.0]	938 (7)	21.7 [19.1; 24.4]	1,877 [939; 938]	1.03 [0.86; 1.26]	0.01 [-0.05; 0.07]	0.740
Texts to study supporters	934 (5)	22.5 [19.8; 25.2]	938 (7)	21.7 [19.1; 24.4]	1,872 [934; 938]	1.04 [0.84; 1.27]	0.01 [-0.05; 0.08]	0.691
Texts to students and study supporters	935 (11)	20.0 [17.4; 22.6]	938 (7)	21.7 [19.1; 24.4]	1,873 [935; 938]	0.92 [0.73; 1.14]	-0.03 [-0.09; 0.04]	0.343

The relative risk ratios indicate that the intervention does not have an impact on the proportion of students achieving a pass in GCSE maths or English on re-sit for any variation of the intervention.

Sensitivity analysis

As there is baseline imbalance between the number of re-sits between the control group and the 'supporters only' and 'students and supporters' intervention arms, sensitivity analysis was conducted including this characteristic as an independent variable. The results of the sensitivity analysis are presented in Table 12 below, producing estimates consistent with the primary analysis.

Table 12: Sensitivity analysis—impact of the intervention on GCSE re-sits in English or maths

Group	Unadjusted pass rate				Effect size			
	Intervention group		Control group		Total n (intervention; control)	Relative Risk Ratio (95% CI)	Hedge's g (95% CI)	p-value*
n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)					
Texts to students	939 (10)	22.4 [19.7; 25.0]	938 (7)	21.7 [19.1; 24.4]	1,877 [939; 938]	1.04 [0.84; 1.28]	0.01 [-0.05; 0.08]	0.657
Texts to study supporters	934 (5)	22.5 [19.8; 25.2]	938 (7)	21.7 [19.1; 24.4]	1,872 [934; 938]	1.06 [0.85; 1.29]	0.02 [-0.05; 0.08]	0.542
Texts to students and study supporters	935 (11)	20.0 [17.4; 22.6]	938 (7)	21.7 [19.1; 24.4]	1,873 [935; 938]	0.93 [0.74; 1.15]	-0.02 [-0.05; 0.08]	0.430

¹⁷ Note that 95% confidence intervals for the Relative Risk Ratios have been Bonferroni adjusted for the primary analysis. Therefore, a p-value of less than 0.0167 would indicate a statistically significant difference with a type-one error rate of 0.05.

Secondary analysis

The results of the secondary analysis of attendance at English and maths classes are displayed in Table 13. A key limitation of this analysis is the extent of missing attendance data (30.5% to 31.4% dependent on the trial arm) and in addition as there is no baseline attendance data we cannot compare with national averages to assess the level of external validity of the findings. This should be considered when interpreting the findings of the secondary analysis.

Table 13: Secondary analysis—impact of the intervention on pupil attendance

Group	Unadjusted means				Effect size		
	Intervention group		Control group		Total n (intervention; control)	Hedge's g (95% CI)	p-value
n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)				
Texts to students	654 (295)	78.4 [76.9; 79.8]	651 (294)	78.9 [77.4; 80.4]	1,301 [653; 648]	-0.01 [-0.38, 0.37]	0.432
Texts to study supporters	657 (282)	79.3 [77.8; 80.8]	651 (294)	78.9 [77.4; 80.4]	1,301 [653; 648]	0.00 [-0.50, 0.51]	0.610
Texts to students and study supporters	660 (286)	78.9 [77.3; 80.4]	651 (294)	78.9 [77.4; 80.4]	1,305 [657; 648]	0.00 [-0.49, 0.49]	0.945

Table 13 indicates that the intervention had no impact on college attendance in any of the intervention arms. This contrasts with findings from recent studies of similar texting interventions (Miller et al., 2016 and Groot et al., 2017) which both found evidence that such interventions may improve college attendance. Kraft and Rogers (2015) argue that teacher-to-parent communication, such as texts, may improve college attendance. As there is evidence that this intervention had no impact on college attendance, this would suggest that changes in the primary outcome are unlikely. Though the content of text messages was not exclusively designed to encourage attendance alone, this may in part explain the null findings of the primary analysis.

Analysis in the presence of non-compliance

The methods chapter outlined the key methodological considerations for the analysis in the presence of non-compliance. Defining a meaningful threshold of the dosage required for the intervention to be effective is challenging. After all, students who were sent the full dosage (all texts) will have had very similar exposure to the intervention as students who were sent at least 90% of texts. It is also worth noting that whilst the data allows us to analyse the texts sent and delivered, we cannot know from the data if the individual read the content of the message. As it is difficult to know what a suitable threshold is required for the intervention to be effective, this analysis is conducted with definitions of dosage with a varying threshold of the proportion of texts a student and/or their study supporter must receive to be considered non-compliant. Table 14 indicates the proportion of students in each trial arm that would be considered compliant for different thresholds of texts.

Table 14: Proportion of students receiving different thresholds of texts

Texts sent	Students only	Supporters only	Students and supporters
No texts	49 (5%)	60 (6%)	16 (2%)
At least 50%	773 (81%)	824 (86%)	815 (85%)
At least 60%	761 (80%)	811 (85%)	740 (77%)
At least 70%	745 (78%)	797 (83%)	705 (74%)
At least 80%	725 (76%)	775 (81%)	659 (69%)
At least 90%	678 (71%)	718 (75%)	594 (62%)
All texts	357 (37%)	351 (37%)	181 (19%)

As Table 14 illustrates, a small proportion of students and/or study supporters were sent all the texts that they were supposed to receive. In particular, students in the ‘students and study supporter’ arm were particularly unlikely to receive the full schedule of texts. This may have occurred as the number of texts that this arm was supposed to receive was higher, and contingent on more than one individual receiving texts (the student and the supporter, as opposed to one or the other). The analysis is therefore conducted for four dosage thresholds: 70%, 80%, 90%, and 100% of texts sent. The opt-out and delivery data indicated that 112 students and 90 study supporters opted out of receiving texts during the course of the intervention. Additionally, during the intervention, students and study supporters may have changed their phone number; 26 students and 4 study supporters were recorded as changing phone numbers during the trial and received texts to their new number accordingly. However, it is possible additional students and study supporters may have changed phone numbers without informing the intervention team.

The methods chapter also discussed the issue of contamination—where students and/or their study supporters were sent texts when they were not supposed to receive them—and hypothesises about how this could have happened. For example, in 42 cases an individual nominated themselves as their study supporter.¹⁸ ‘Compliance’ is therefore defined as a binary indicator that they were sent the defined dosage (based on the thresholds outlined above) and did not receive any erroneous texts (either to the student themselves or to the study supporter).

Table 15: Proportion of students sent texts and their allocation

Texts sent	Study arm			
	Control (%)	Students only (%)	Supporters only (%)	Students and supporters (%)
No texts	887 (93%)	48 (5%)	55 (6%)	16 (2%)
Texts to student	24 (3%)	797 (83%)	5 (1%)	51 (5%)
Texts to study supporter	32 (3%)	1 (0%)	795 (83%)	27 (3%)
Texts to student and study supporter	14 (1%)	109 (11%)	99 (10%)	863 (90%)

¹⁸ This estimate is based on the number of cases where the student nominated a study supporter with the same first name and surname.

Table 15 outlines the proportion of students in each trial arm that were sent erroneous texts or no texts. This was discovered when the evaluator accessed the texting data provided by BIT. The level of contamination varies by trial arm, with greatest prevalence for students allocated to either have texts just to the student themselves, or just to their nominated study supporter (17% for each arm respectively). Contamination was lower in the control arm and in the students and study supporters' arm (7% and 10% respectively). It is hypothesised that contamination was introduced by students in the control arm or in the students-only arm being nominated as a study supporter by students in other trial arms. Students in the supporters-only arm may have nominated themselves as a study supporter. Texts may not have been received due to students or study supporters opting out or because they changed their phone number (and did not inform the intervention team).

Table 16: Estimated impact of the intervention in the presence of non-compliance

Group	Dosage threshold (%)	Effect size		
		Total n (intervention; control)	Relative Risk Ratio (95% CI)	p-value
Students only	70	1,877 [939; 938]	1.15 [0.97; 1.36]	0.114
	80		1.16 [0.98; 1.37]	0.089
	90		1.17 [0.99; 1.39]	0.071
	100		0.90 [0.72; 1.12]	0.365
Study supporters only	70	1,872 [934; 938]	1.00 [0.83; 1.18]	0.966
	80		0.99 [0.83; 1.18]	0.937
	90		0.98 [0.81; 1.17]	0.817
	100		0.92 [0.73; 1.14]	0.450
Students and study supporters	70	1,873 [935; 938]	0.98 [0.82; 1.17]	0.849
	80		0.94 [0.78; 1.13]	0.501
	90		0.99 [0.78; 1.25]	0.920
	100		0.89 [0.67; 1.16]	0.405

Table 16 presents the estimated impacts of the intervention on students compliant with their allocation. Despite considering a range of dosage thresholds, the relative risk ratios are consistently small, suggesting that the effectiveness of the intervention was not influenced by the number of texts that a student and/or their nominated study supporter were sent (dosage).

Subgroup analyses

The Statistical Analysis Plan outlined four subgroups: free school meal eligibility, gender, re-sit subject (English or maths), and whether the student is re-sitting one subject (English or maths) or both. The subgroup analysis of the impact on FSM students is conducted in two models: firstly, on the sub-sample of FSM students, and secondly an interaction model using the full analytical sample. The FSM subgroup analysis uses eligibility for free school meals at the end of KS4, or the preceding six years.

All other subgroup analyses use the interaction model approach. *For all subgroup analyses, the intervention arms have been combined to maximise statistical power*, though this analysis is exploratory. The results of all the subgroup analyses are presented in Table 17. These results show the effect size of the intervention for the 'ever FSM' subsample. For the remaining subgroup analysis, the effect size reported is the effect size for the interaction term between the treatment indicator and the indicator for the subgroup in question (ever eligible for FSM at the end of KS4, Female, English, and re-sitting both English and maths). This effect size therefore indicates the difference in outcomes for that subgroup if they are assigned to the intervention.

Table 17: Subgroup analysis—estimated impact of the intervention on GCSE re-sits in English or maths

		Unadjusted pass rate				Effect size			
		Intervention group		Control group		Total n (intervention; control)	Relative Risk Ratio (95% CI)	Hedge's g (95% CI)	p-value
Group	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)					
FSM (subsample)		157/792 (0)	19.8 [17.0; 22.6]	57/267 (0)	21.3 [16.4; 26.3]	1,056 [792; 267]	0.91 [0.65; 1.29]	-0.03 [-0.12; 0.08]	0.606
FSM	Eligible at 15	157/792 (0)	19.8 [17.0; 22.6]	57/267 (0)	21.3 [16.4; 26.3]	2,900 [2,172; 728]	0.81 [0.56; 1.47]	-0.06 [-0.16; 0.12]	0.274
	Not eligible at 15	336/1381 (0)	24.7 [20.8; 28.7]	114/461 (0)	24.3 [22.1; 26.6]				
Gender	Female	1,501 (0)	20.7 [18.6; 22.8]	479 (0)	18.2 [14.7; 21.6]	3,726 [2,793; 933]	Female: 1.27 [0.96; 1.63] Male: 0.86 [0.67; 1.10]	Female: 0.08 [-0.01, 0.18] Male: - 0.05 [-0.12; 0.03]	Female: 0.088 Male: 0.235
	Male	1,292 (0)	22.8 [20.5; 25.1]	454 (0)	25.6 [21.5; 29.6]				
Subject	Maths	1,383 (0)	15.4 [13.5; 17.3]	474 (0)	16.7 [13.3; 20.0]	3,746 [2,808; 938]	0.92 [0.67; 1.22]	-0.02 [-0.11, 0.06]	0.561
	English	1,425 (0)	27.6 [25.3; 30.0]	464 (0)	26.9 [22.9; 31.0]				
Number of re- sits	One	1,635 (0)	24.4 [21.0; 27.9]	589 (0)	25.1 [23.0; 27.2]	3,746 [2,808; 938]	0.95 [0.69; 1.27]	-0.02 [-0.11, 0.08]	0.723
	Two	1,173	16.8 [14.7; 18.9]	349 (0)	17.2 [13.2; 21.2]				

The pass rate for GCSE re-sits in English or maths was lower for pupils who had been eligible for free school meals by the end of KS4 relative to their peers, in both intervention and control, though this was not statistically significant. This reflects the known attainment gap between disadvantaged students and their peers. The subgroup analysis indicates that the intervention had no effect on pupils eligible for free school meals the end of KS4, or the preceding six years. Furthermore, the intervention effect did not differ for these pupils relative to their peers.

Male students had a marginally higher pass rate than female students, though this was only statistically significant in the control arm. Male students receiving the intervention were no more likely to pass or fail their re-sit relative to males in the control group. The relative risk ratio indicates that the female pass rate was 27% higher (or a 5.6 percentage point increase in the pass rate) if they received the intervention, though this was not statistically significant.

Students re-sitting GCSE English were significantly more likely to pass than students re-sitting GCSE maths. However, the intervention did not have a differential impact on students sitting English relative to maths. The pass rate was also higher for students re-sitting a single subject, relative to students re-sitting both, though this is only statistically significant in the intervention arm. There are several possible explanations for this. For example, students with poorer initial attainment were more likely to fail both GCSEs at the end of KS4, or the additional workload required to re-sit two

subjects may have diverted efforts that would have otherwise been focused on a single subject. The results from Table 17 indicate that the intervention did not have a differential impact on students whether they were re-sitting a single subject or both English and maths.

Implementation and process evaluation

This chapter outlines key findings from the IPE of Project Success. It brings together the views and experiences of those involved in the evaluation, including college project and subject leads, students, and study supporters. Findings are organised thematically and discussed sequentially, moving from the early stages of intervention buy-in and sign-up, to implementation fidelity and usual practice across colleges, to the perceived outcomes of Project Success and suggestions for improvement. We firstly report on sign-up rates among students, key characteristics of those taking part, motivations for opting out, and any differences between them and those who signed up in terms of attitudes to maths, English, and college. We then delve into the range of aspects emerging from interviews as affecting buy-in of different groups (students, study supporters, college project and subject leads) within colleges, including the availability of information about the intervention, key features of this and of the sign-up process, and student motivation. We also present findings on implementation fidelity and aspects identified by participants that may have damaged this, and report on 'business as usual' within colleges to provide a better understanding of the activities and resources that students in the control group were exposed to. Finally, we discuss perceived outcomes of the intervention as identified by interview participants and the extent to which these provide support to the Logic Model (Fig. 1), and summarise formative findings on how the intervention may be improved.

Characteristics and attitudes of students signing up to Project Success

This section presents findings on:

- the sign-up rates to Project Success;
- the profile of students who opted into the intervention and any differences between them and those who opted out in relation to attitudes to maths, English, and their college; and
- the main reasons why students opted out of receiving the programme.

It provides contextual information on the characteristics of students being evaluated and offers insights on the attractiveness of the intervention for its target group as well as possible limitations in its reach.

Sign-up rates and students' profile

A total of 7,004 students were identified by college project leads as meeting the eligibility criteria for Project Success and took part in the sign-up survey designed by BIT and administered by college staff. Of these, 515 were ineligible as they were taking English and/or maths GCSEs but not as re-sits. These students were not randomised and their data has been excluded from both the impact and process evaluations. Of the 6,489 students who were eligible to participate in the programme,¹⁹ a majority opted in to receive BIT text messages (62.8% compared to 37.2% who opted out).

Of those who signed up, 20 were initially registered for maths and/or English GCSE re-sits but were later moved to a lower qualification. These students are recorded in the impact evaluation as not having sat the re-sit exam as per the specification in the Statistical Analysis Plan (Smith and Ashworth, 2018). Data collected from them through the sign-up survey was included in analysis on intervention take-up and students' profile and attitudes conducted as part of the IPE.

Among those who participated in the trial, 27.1% of students were re-sitting GCSE English, 30.9% GCSE maths, and 42% both English and maths. Most of them took their last GCSEs in the previous year although a considerable number did so in the years before (62.9% compared to 37.1%).

Students participating in the evaluation were almost equally split between females and males (51.7% and 45.9% respectively), with 1.3% identifying themselves as non-binary.²⁰ The large majority lived with parents, relatives, or

¹⁹ This is the total number of students who were eligible to take part in Project Success across all recruited colleges, including those who had then dropped out prior to randomisation.

²⁰ These numbers may differ from those included in the impact analysis as they are based on student self-reporting of gender as part of the diagnostic survey as opposed to NPD data.

guardians (93.3%), while smaller proportions lived with their partners (3.2%), on their own (1.2%), or with other students (0.3%).

Motivations for opt-out

The main reason selected by students for opting out of receiving BIT text messages was related to the trial rather than the intervention itself, that is, not wanting to participate in research (with 44.7% of those opting out selecting this option, 1080 students). This was followed by a notable 31.6% of students (764) not wanting to receive messages about college. It is also worth observing that 14.5% (351) selected 'I can't think of anyone to choose', 4.9% (118) 'I am afraid to ask people to be my study supporter', and 3.9% (95) 'I don't have a mobile phone myself', thus raising questions about the reach of the intervention and its inclusivity.

Differences in attitudes between students who opted in and those who opted out

Drawing on data collected through the sign-up survey, self-reported attitudes to maths, English, and college of students who opted in to the intervention were compared to those of students who opted out at sign-up stage. Analysis involved bi-variate statistics (cross-tabulations) in SPSS. Statistically significant differences were found between the two groups in views and attitudes around studying and college, and around maths and English in particular. Compared to the latter, students who signed up to receive BIT text messages were significantly more likely to:

- believe if they put in enough effort, they could succeed in their English GCSEs (78% agreeing or strongly agreeing compared to 69%) and maths GCSEs (75% compared to 68%);
- state that if they had different teachers, they would try harder in their English GCSEs (29% agreeing or strongly agreeing compared to 25%) and maths GCSEs (28% compared to 24%);
- think if they wanted to, they could do well in their English GCSEs (74% agreeing or strongly agreeing compared to 64%) and maths GCSEs (70% compared to 62%);
- report they work hard at college (80% agreeing or strongly agreeing compared to 72%) and complete their college work regularly (81% compared to 75%);
- mention they wanted to learn skills they could use in a job to help others (84% agreeing or strongly agreeing compared to 71%); and
- report higher levels of belonging at college (33% compared to 27%).

On the other hand, they were less likely to state they 'just don't find English GCSE interesting' (36% compared to 43%) and that they 'just don't find maths GCSE interesting' (40% compared to 44%).

Overall, those who signed up to the intervention appeared, therefore, to be already more confident, motivated, and engaged—both in English or maths and more generally in college—than those who opted out. This suggests a limitation to the programme's ability to engage those who may need it the most.

Intervention buy-in

Overall, findings from the process evaluation indicate that project and subject leads, students, and study supporters held positive views of Project Success. Nevertheless, interviewees highlighted several factors which influenced their own and other individuals' buy-in to the intervention. These included the availability of information about Project Success, key features of the intervention, student motivation, and aspects of the sign-up process. These aspects played out in different ways across colleges and among interview participants. Project and subject leads generally reported positive views about the intervention and their experiences of delivering it. Yet, they commonly stated that college tutors tended, in contrast, to be more sceptical. Similarly, buy-in from students was more difficult to ensure, with student sign-up to Project Success being especially low in certain colleges.

Availability of information

The availability of information for project and subject leads, college tutors, students, and study supporters was crucial to ensure buy-in to the intervention and successful implementation. Information was provided through various sources,

targeted at different groups of those involved with the programme. There was a range of views regarding the effectiveness of information provided. Whilst project leads felt well informed, some students and study supporters reported the information they received was not comprehensive.

Informational workshop

Project leads felt the workshops delivered by BIT were important to ensure an effective introduction to the intervention. In general, interviewees found the workshops useful and informative, and thought they helped improve their understanding of why their engagement was essential for successful implementation. The workshops also gave an opportunity for tutors to learn about the research evidence supporting Project Success, which contributed to their buy-in.

'After that presentation we were very positive. We could see ... it was definitely worth giving it a go ... And there was a little bit of evidence to suggest it might help, so, you know, even if it helped a few of our students it was worth it' (subject lead, English teacher).

Cascading knowledge to college tutors

After their introduction to the intervention, project and subject leads were responsible for cascading information to other college staff. Project leads felt teachers had a pivotal role in raising awareness amongst students and facilitating student sign-up during lessons. Therefore, the cascading of knowledge from project leads to college tutors was considered an important aspect of successful implementation.

'I mean, teachers need to be the salesman of this kind of thing, because it's their students that are participating, so yeah, I think that's crucial' (project lead, Learning Innovation Manager).

College-wide awareness raising

Although project and subjects leads reported the value of making the intervention 'high profile', students raised the issue that the intervention did not gain much publicity at their college.

'I'd say it wasn't publicised quite like as a big thing. It was quite a small thing' (Student, re-sitting GCSE maths).

As such, they felt it should have been publicised more to raise awareness and therefore increase student sign-up.

Instructions from teachers

Alongside college-wide publicity, project leads felt that it was important to provide eligible students with clear and comprehensive information prior to signing-up.

'It goes back to just explaining what the project is, what kind of information they're getting and how, how frequently they're going to get it. And as long as that's done then you've got I think a better opportunity to get them to sign up' (project lead, director of English and maths).

As sign-ups usually occurred in maths or English lessons, subject tutors were responsible for providing students with information about Project Success and addressing student queries. However, it was felt by some students that the information provided lacked detail and clarity, resulting in confusion and a reluctance to sign-up. To counter this, one project lead suggested that the presentation of the intervention might have been more informative if it came directly from the developers rather than being delivered by college staff.

'They're hearing everything second hand from us. I think if Project Success had spoken to the students they might have got a better buy-in from our college' (project lead, head of English and maths).

Online information video

Although BIT did not deliver a presentation for students, they provided an online video designed to inform them about the intervention and the trial. In general, the video was well received by students as they thought it provided an informative introduction to Project Success. However, one student commented it would have been preferable to receive a written letter alongside the video presentation to remind them about the details of the intervention.

In addition, some students felt the video did not contain comprehensive information about group allocations, which led to uncertainty about who would be receiving the text messages. While this has to do with the trial rather than the intervention, it is revealing of broader students' concerns around who receives the messages.

Information provided to study supporters

As study supporters did not receive any information from the developers or college until the first text message, they were reliant on the information provided by students. Often, students did not inform their study supporter about the intervention prior to receiving the first text message. In most cases, once study supporters had the opportunity to ask the student about the text message, they understood the intervention and agreed to continue to take part. However, interviews with early-opt out study supporters revealed the lack of information may have caused a small minority of study supporters to withdraw from the intervention.

'If I'd understood a lot better, like in the beginning, I might have taken part' (study supporter, older brother).

Properties of the intervention

The findings from the process evaluation indicated that perceptions around different aspects of the intervention varied among individuals. Views around key features of the intervention contributed to initial and sustained buy-in amongst participants, and so influenced successful implementation. In general, the intervention was appealing to project and subject leads whereas students and study supporters were more critical around certain aspects.

Text message based

A majority of project and subject leads considered the mode of communication to be an appealing feature of the intervention. In general, they felt communicating via a mobile phone was an appropriate and effective way to communicate with the target student cohort.

'The kind of age groups that we work with, they're always on their phones' (project lead, senior maths tutor).

Furthermore, project and subject leads felt the instant nature of text messaging and the minimal effort required to access the content was an advantage over other methods of communication.

'The good thing about text messages is they arrive at your fingertips as opposed to having to access something' (project lead, head of skills for employment).

Across colleges, project and subject leads commonly reported previous experience of using text messages to improve attendance, which contributed to their confidence in the use of text messages to improve student outcomes.

'It is something I already think works because we do text messaging ourselves, and I'm a massive advocate of using text messages in the college to get students to attend, especially exams' (project lead, curriculum area manager for maths).

In contrast, study supporters were less positive about this feature of the intervention. In one case, a study supporter decided to opt out of the intervention because of their dislike towards using a mobile phone.

'I don't like using my phone too much, so I hate getting texts and stuff like that, so I opted out of it straightaway' (study supporter, older brother).

Furthermore, a few project leads, importantly, drew attention to the fact that not all students and study supporters would have access to mobile phones, which was a barrier for some students signing-up to the intervention.

'We often take for granted sort of mobile phone technology, but some students don't have them, some parents don't have them' (project lead, head of campus).

Low risk

Project and subject leads felt there was little risk involved in taking part as they considered the BIT texting programme a 'seamless addition' to current practice that would involve no extra 'stress or strain' for teachers. Their expectations about the minimal disruption to business as usual contributed to their initial buy-in.

Yet, project leads also reported that tutors had expressed concerns about the impact of the intervention on their workload. In particular, tutors were worried about the burden of the sign-up process as it coincided with the start of the academic year, which usually involves a higher workload compared to the rest of the college year. While concerns about the extra workload involved with project set-up at the start of the academic year were confirmed post-intervention, those who mentioned this also stated that this was not too onerous.

Project and subject leads also thought the intervention was low risk because they felt it would not cause detrimental outcomes for students. Likewise, students felt that there was 'nothing to lose' by taking part.

'There's very little risk involved if you're receiving a positive text message or your parents are ... if somebody's saying, "well done", it's unlikely to be detrimental' (project lead, head of English and maths).

Study supporter element

Two divergent discourses emerged around the study supporter element. A majority of project leads held the view that this element of the intervention was pivotal in improving student outcomes, describing the inclusion of study supporters as 'innovative' and 'valuable'. They felt study supporters would be able to provide support and motivation above and beyond that offered by teachers due to the close personal relationship with the student. In support of this view, some students indicated they believed this aspect of the intervention was important as their study supporter provided motivation for them to work harder.

In contrast, other interviewees expressed concern about this element of the intervention. A few project leads reported that at this age, students are independent and so the involvement of a study supporter may not be appropriate. In line with this view, some students explained they viewed the involvement of a study supporter as a 'deal breaker' to them signing-up.

'The students are quite independent, and they, they didn't really want to give parents their permission to help them with their GCSEs, if that makes sense, or a study supporter' (project lead, vice principal).

Alongside the involvement of study supporters, concerns were also raised about the selection of study supporters. Project leads noted some students had difficulties in selecting a study supporter, either because they were unsure of who to select or because they felt that they did not have a suitable person who would be able to support them.

'They don't have a trusted relationship with anyone, even sometimes their parents ... So joining up to a site where they're asked to provide somebody to give them support, that isn't something they can necessarily do' (subject lead, maths tutor).

Some project leads expressed concerns about the texts causing safe-guarding issues, particularly if there were previous concerns about the relationship between the student and study supporter.

'It actually becomes a negative kind of telling-off situation rather than a supportive, encouraging thing' (subject lead, college maths lead).

In addition, a few study supporters questioned the value of their involvement as they held the view that teachers are better placed to support revision.

'I didn't mind helping her, but, I mean, things like this I would have thought would have been dealt with at college' (study supporter, grandfather).

In light of these problematic aspects, project and subject leads argued that the involvement of study supporters should not be compulsory.

Developed by BIT

Project and subject leads tended to hold favourable opinions around BIT and their research. This contributed to their initial buy-in and positive expectations of the intervention.

'I was familiar with who the Behavioural Insights Team were prior to seeing it was advertised. I really like behavioural sciences and the work that the Behavioural Insights Team specifically have undertaken ... it was

a case of a bit of a no-brainer, knowing that there was a tool out there designed by a team of experts who could potentially influence behaviours' (project lead, acting head of campus).

Furthermore, projects leads reported very positive experiences of their working relationship with BIT, which contributed to sustained engagement with the intervention.

Student motivation

As the intervention was not compulsory but instead relied on students to opt-in, their willingness to sign-up influenced implementation. There were some students who were not motivated enough to sign-up. In line with findings from the BIT diagnostic survey reported earlier, subject and project lead interviews importantly drew attention to how the students who sign up are also likely to be the ones who are already motivated and engaged, highlighting limits in the capacity of the intervention to reach those who would need it the most.

Furthermore, students' perceived lack of motivation affected buy-in from study supporters, some of whom felt their involvement would bring little benefit to the outcomes of the student given their disengagement.

'If she would like [to] actually get on and knock down with English, yeah, I would help her. Why not? It's just she's not bothered so I don't see there's any of point of me helping her' (study supporter, friend).

Sign-up process

The sign-up process was important to ensure eligible and willing students would be involved in the intervention. However, a number of aspects of the sign-up process were viewed as barriers by project and subject leads and by students, which resulted in some students deciding not to take part.

Online access

Online access was necessary for the completion of the online sign-up form. Therefore, the number of students who were able to sign-up was dependent on them having access to adequate IT equipment. A popular approach used by colleges to facilitate student sign-ups was to use a computer room during a lesson, usually during the lessons students would be receiving texts for (English or maths). However, issues with using this approach were reported by some colleges, most commonly related to difficulties booking computer rooms.

'We had to book computer rooms for 250 students, which in itself is a bit of a headache' (project lead, head of English and maths).

In response to this barrier, some colleges encouraged students to use their own mobile phones to sign-up to Project Success, thus eradicating the need to use a computer room.

'But that was good about it that it was mobile-friendly, so they could do it on their phone... anything that's mobile-friendly just make it's so much easier for them to do, don't it?' (project lead, senior maths tutor).

Length of sign-up process

There were contradictory views expressed by students about the length of the sign-up process. Whilst some students said the sign-up process took the right amount of time to complete, others thought the sign-up survey took too much time. In one case it took a student one hour to complete and this was a deterrent to them signing-up.

"Cause a couple of my classmates were saying, "Oh, these questions are too long, forget it", like a handful of them, so you know. If there was a little bit less then you know why not?' (student, re-sitting GCSE English).

Although the length and content of the sign-up questionnaire is related to data collected for the evaluation rather than the intervention, it is an aspect worth noting for future delivery.

Concerns about sharing personal data

A perspective reported by students was concern about sharing their personal data during the sign-up process. Although most of this data was collected for the evaluation rather than the delivery, students still expressed wariness around sharing their personal mobile phone number. In most cases, this was driven by concern about how their data would be

used by developers. Project leads noted that often students influenced their peers, which resulted in wide-spread concerns in some colleges that impacted sign-up.

'There was this little, little posse of students that were saying, "Oh no. It's kind of, it's big, you know, the kind of big brother watching us and our data's gonna be used." ... there was this sort of conspiracy type thing' (project lead, English and maths).

For some students, the suspicion was dispelled after receiving reassurances from their teachers that it was a 'safe platform'. However, other students felt it would have been helpful to receive reassurance from the developer to eliminate their concerns.

Fidelity

As the intervention required minimal input from college, the intervention was, on the most part, delivered as intended. However, three aspects identified by participants that may have damaged fidelity relate to the delivery of texts messages, the editing process, and contamination across conditions.

Delivery of text messages

It was intended that all students and/or their study supporters would receive texts for the duration of the intervention period. However, as a few students had changed their mobile number and did not inform BIT, they stopped receiving texts. BIT provided students with a form to update their contact details and those of their study supporter. While some students took advantage of this opportunity, not all of them did. In another case, a student reported that they stopped receiving texts for no apparent reason. For students in the supporter-only or student/supporter arms, supporters could opt-out without informing them, which is a possible explanation for students no longer receiving texts.

Editing of text message content

It was intended that project and subject leads would plan the content of the text messages in advance of the college term. The majority of project leads reported planning the texts according to the scheduled plan of work for their college. However, project and subject leads commented that it was difficult to schedule texts for the whole term, especially because the schedule of work may change. They expressed a preference to prepare messages for each half-term.

Project and subject leads were given the opportunity to edit their planned texts three days prior to when they were due to be sent to students. In general, it was reported that the texts were easy to edit, with project leads stating the process took them between five to fifteen minutes, and as the process became more familiar, the editing process became quicker. The process was considered straightforward and not burdensome.

'I wanted to sort of add something in terms of motivation and they were fine about it and edited it and adapted it and it was really great' (subject lead, curriculum manager for English).

However, a few project and subject leads said they did not have enough time to edit the text messages to match the lesson content.

Contamination across conditions

There were instances when students who were involved in the trial were listed as another student's study supporter. These students, particularly those in the control group, reported using texts addressed to them as a study supporter for their own revision. A full breakdown of contamination can be found in Table 15.

Data collected as part of the sign-up survey shows that 275 students who signed up for Project Success provided the same telephone number for student and study supporter. This may have led to contamination between the student-only/student-and-supporter and the supporter-only arms, as these students received BIT text-messages intended for their supporters. An additional 241 students who signed up were also nominated by others as supporters, which may have led to contamination between the student-only/student-and-supporter and the supporter-only arms and to spill-over to the control group.

Usual practice

This study adopted a 'business as usual' model, whereby students randomised to the control group did not receive Project Success text messages. As part of the Memorandum of Understanding signed to take part in the trial, colleges agreed to minimise the use of their own text messages and limit the content of texts to procedural issues in order not to over-burden students and interfere with the trial.

Usual practice in intervention colleges, to which students in both treatment and control groups were exposed, was assessed through qualitative interviews conducted as part of the IPE; there was no survey or formal data collection on other 'business as usual' practices during the trial. Although interviews were not conducted with control group students, discussions with project leads and students in receipt of the intervention indicated that participating colleges delivered several other activities and strategies aiming to improve students' attendance and attainment in GCSE maths and English, alongside Project Success.

Colleges used a range of techniques to promote attendance, including text message reminders with exam dates, rewarding high attendance with bursaries and prize draws, and engaging parents where students' attendance was low. Other schemes were intended to support students' attainment in their re-sits, such as awards for the 'most improved student of the year' and an online platform providing students with feedback about their performance in maths and English sessions and identifying areas to revise further. In some cases, colleges arranged additional revision classes in lunchtime or during the holidays, and students were given action plans or supported by Learning Support Assistants or student buddies where an additional need was identified. Finally, other approaches were described which aimed to support students' broader wellbeing and achievement in turn, including mindfulness stress reduction activities, breakfast vouchers on exam days, and provision of pastoral care.

Perceived outcomes

The overall view amongst all those interviewed was that the intervention resulted in positive outcomes across a range of variables. However, the reported outcomes of the intervention varied between students, study supporters, and project and subject leads. As project and subject leads were not aware of which students were receiving the intervention, they were often reluctant to comment on the outcomes or attribute observed changes to Project Success.

Increased revision

Most students reported increased engagement in revision outside of college, thus providing some evidence in support of the programme logic model where the reception of messages by students and discussions with study supporters are expected to lead to increased students' engagement.

'It's made me work harder and do more work outside of lessons' (student, re-sitting GCSE English).

The extent to which the texts led to increased revision varied between students. Whilst some mentioned revising a couple of times more, other students cited notable increases in the amount of revision they were completing in comparison to previous years.

Although some project leads also noted increased engagement amongst students, they were not sure if the increases in motivation could be attributed to Project Success.

'I've seen students that are motivated to attend lessons ... But I'm not entirely sure whether they've been texted or not, actually, so it's hard to tell' (project lead, English lead).

Despite declared increases in revision at home, increases in engagement did not appear to extend to work completed during maths or English lessons.

The text messages sent to students functioned in two related ways to help them with learning and increase revision, outlined below.

Organisation and planning

The text messages reminded students they should be revising, by either directly telling them to revise or informing them about upcoming mocks or exams, which indirectly prompted them to revise.

'It just gave me a reminder that ... I need to go and revise for maths' (student, re-sitting GCSE maths).

'It actually reminded me that I've got maybe a mock or an exam and stuff like that, so I revise more' (student, re-sitting GCSE maths).

Similarly, text messages sent to study supporters also functioned as a reminder for the student to revise. In these instances, the reminder was delivered by the study supporter to the student in person, usually immediately after receiving the text or when the student returned home from college.

'When she gets the text, she just tells me, "Oh, go and revise." I think it's like a reminder for her, the text, mainly, 'cause I get it nearly every week, especially for her as well' (student, re-sitting GCSE maths).

The texts also assisted students' revision by helping them to plan and organise their revision schedule. In particular, the texts provided reminders about what topics students had already learnt in class and which topics they were going to be learning about. In one case, a student used the texts to get information about topics they had missed whilst off sick so they were able to catch-up at home.

'They've helped me organise myself ... They send me the messages and it helps me decide on what. 'Cause when you're just going in without an idea of what you're going to revise, that helps me, just with the text messages to help me look. And I've seen them and I've thought, right, I'll revise that one, that topic' (student, re-sitting GCSE maths).

Provision of resources

Additionally, the texts provided students with resources they could use to revise. In general, students felt the weblinks were helpful as they informed them about a range of websites they could use to aid their revision.

'Whenever I get the text messages they're like links to help, websites to help, so I just tend to look at them on my breaks because I'm not in lessons for them, so ...' (student, re-sitting GCSE English),

Although most students claimed to have used the weblinks in the texts, some students instead preferred to use the weblinks that were provided by their teachers during lessons.

'I haven't really gone on to the links. I've just gone on some of the websites that the teachers have said' (student, re-sitting GCSE maths).

Feeling supported and more confident

In the logic model for Project Success, students' confidence is expected to increase because of receiving BIT text messages as is their engagement with their studies and in conversations around these with study supporters. Students reported that they felt more positive or confident about exams they had completed and also for upcoming exams. In contrast, project leads and study supporters did not comment on increased confidence in students.

Increased quantity and quality of supportive communication between student and study supporters

The text messages were related to a reported increase in frequency of communication between students and study supporters. Study supporters attributed this increase to the texts acting as a 'conversation starter'.

'My dad usually gets a text message and then when I get back from college he's like, "You should start revising and have like a break and then start again on track." ... he's trying to motivate me a lot to get through it, which is helpful' (student, re-sitting GCSE English).

In addition, some study supporters mentioned that the texts led to an increase in their knowledge in either maths or English, which meant they felt more confident in their ability to talk to the student about their college work and revision after receiving the texts.

Finally, some study supporters reported that they helped students with their revision after receiving the texts.

'Well, if he knew I had an exam coming up, he'd tell me. He'd come and help me. We'd sit at the table and do like a paper together. And we'd just do it together' (student, re-sitting GCSE maths).

These findings lend support to the programme logic model, indicating that BIT text messages received by study supporters can serve as 'conversational prompts' and to improve study supporters' understanding of the content covered by students in college, thus facilitating discussion around studies and the provision of help.

However, the extent to which the texts led to an increase in quantity and quality of communication between students and study supporters appeared to be moderated by several factors:

- **prior relationship**—students who described a supportive relationship with their study supporter prior to Project Success did not report an increase in quantity and quality of communication;
- **frequency of communication**—students who had frequent contact with their study supporters, usually in college or at home, noticed increases in communication; however, students who had limited contact reported that there were no increases in communication; and
- **knowledge and confidence in maths or English**—a couple of students reported that their study supporters had limited understanding of the subject, which resulted in them being able to offer limited help.

Better attendance to maths or English classes and exams

The extent to which project leads reported increases in attendance varied at the college level. Some project leads noted increases in lesson and exam attendance for maths and English. Students reported that the texts functioned as a reminder and encouragement for them to attend lessons.

'It would just remind me that I like should be going to my lessons' (student, re-sitting GCSE maths).

However, project and subject leads often considered these increases small.

'The attendance in English and at maths throughout the year have improved slightly but, but not massively compared to last year' (project lead, director of English and maths).

Furthermore, in one college, project and subject leads reported an increase for maths attendance but not for English.

Nevertheless, most were reluctant to attribute increases to Project Success, especially as the colleges often had additional measures in place to increase attendance (for example, attendance cards, contacting parents, emails, and monetary rewards).

'It's hard to gauge whether it's just the texting or is it the texting plus all the other things that we're doing?' (project lead, director of English and maths).

Finally, some students did not think their attendance had increased, often because their attendance rate was already high and so there was little room for improvement.

Fewer college drop-outs

A related outcome to increased attendance was smaller college drop-out rates for pupils involved in Project Success. However, like attendance, project leads were reluctant to attribute the reduction in drop-out rates to Project Success.

Improved GCSE mock and exam results

A few project leads reported mock exam results were higher than expected. However, they were reluctant to attribute increases in results to Project Success. Ultimately, while hopeful of improvements, project leads and study supporters were cautious to comment on the impact of GCSE grades until results had been released.

'We won't know till we get the results 'cause they might say, "It helped me", but if it didn't improve the grade, we don't know if it did help them' (project lead, senior maths teacher).

Formative findings

Overall, the findings from the process evaluation indicated that Project Success was well received by students, study supporters, and teachers. Nevertheless, interviewees highlighted a number of ways in which the intervention and its delivery could be improved. These included amendments to the project set-up and student enrolment process, proposed changes to text messaging, including length, frequency and scheduled timings, as well as increased tailoring of messages. The study supporter element was discussed as an area for change as teachers felt that students did not always select a suitable person to engage with the texts. Similarly, study supporters reported not feeling equipped to support students' learning. Finally, it was acknowledged that students rarely discussed the texts with teachers or peers, suggesting that the intervention was low profile within the colleges.

Participation and sign-up process

Respondents raised several suggestions for improvements to the sign-up process. As described in relation to college-wide publicity, some felt that the intervention had not been promoted to students sufficiently prior to sign-up. This was particularly pertinent for large or multi-site colleges where provision of advertisement materials was felt to be limited:

'At the very beginning there were very few posters ... we each got four posters. And obviously because we work on different sites it meant it was racing on the ground in terms of making it ... visually apparent ... We did photocopy some of the posters ourselves to put up, because obviously it helped to have it on the doors ... I think that was really important' (project lead, director of English).

Project leads proposed multiple adaptations to intervention delivery to increase uptake, including making sign-up compulsory for all students re-sitting their maths or English GCSEs. The sign-up window was also considered to be key to student uptake. Some project leads reported that the sign-up process fell too early in the academic year, when students were trying to settle into their new learning environment. As such, it was suggested that the sign-up window could be extended to allow students to opt-in to receiving the intervention at any point in the year. In contrast, another project lead held that early sign-up was crucial to successful delivery and suggested embedding this into students' enrolment week, as motivation is high and potential study supporters may be present:

'If you want the best opportunity to see a student when they're motivated and looking forward to doing something productive with their future, it would be the enrolment week. So yeah, if we integrated that in our college systems, promoted a [study supporter], then I think that would work really well' (project lead, curriculum area manager for maths).

As outlined in relation to student sign-up, students expressed concerns around data protection, indicating that the sign-up form should clearly emphasise that data will not be used for purposes other than intervention delivery:

'The reaction from learners ... that they weren't prepared to give away certain details to an unknown third party ... They see all this stuff in the news about data protection and people giving away their information and they probably just think that that's happening every single instance' (project lead, curriculum area manager for maths).

Timing and frequency of text messages

Contradictory feedback was shared about text message timing and frequency. Students reported receiving messages in the evening at the latter end of the college week. In some cases, this was regarded as a timely reminder for weekend revision, whereas other students and study supporters felt less engaged with the messages by this point and would have preferred for them to arrive closer to the start of the week:

'To my recollection the texts always came in on a Friday evening just after a long ... you know fairly hard week of work. And maybe the timing of the texts just, you know is ... had enough sort of thing during the week and then you think, "Oh God I've got something else to do now!"' (study supporter, father).

Likewise, students and study supporters presented varying perspectives as to the preferred frequency of message delivery. In some cases, participants wanted to receive fewer messages as weekly messages were found to be excessive:

'Get them every Thursday, so it's every, it is a weekly thing ... To me, I'd say that's probably a little bit overkill. I think every fortnight would have been all right' (study supporter, older brother).

In contrast, some students and study supporters suggested adaptations to delivery in the form of two messages a week. Alternatively, a variable message schedule with either increased messages in the lead-up to exams or based upon student need was also suggested:

'They could send more texts if they'd like or, you know, more often, or, you know, it depends really on the person, how they're feeling. Like do they need it once a week, twice a week? ... You could put that in the sign-up process, you know, as one of those things.' (student, re-sitting GCSE maths).

Text message content

Text message content was frequently discussed as an area for revision. Whilst many thought that message content was appropriate, some felt that messages were too long, or that the tone of messages was better suited to younger students. Project leads stated that messages containing subject-specific information rather than motivational content were better suited to their cohort.

'I just think the type of motivational text that Project Success created didn't hit the right note with our type of students ... there were other more valid ones about the connection between attendance and achievement, which is obviously clear. But I think the, sort of, poetic metaphorical stance that some of the messages took was not necessarily for our type of cohort. So it'd be more practical and more fact-based and more reminders' (project lead, GCSE English teacher).

This was mirrored by students, who preferred to receive factual information or links to helpful resources.

'They say like you can go to your tutor or whatnot but ... if someone didn't have the courage to go to their tutor and ask for help or whatnot, they could put a link in for like for an exam or to a past paper or to some sort of question that they might be struggling with or something, I don't know. Or something a bit more specific' (student, re-sitting GCSE maths).

Project leads also reflected on tutor involvement in reviewing message content before these were sent to students and noted that they perhaps should have made more amendments to adapt messages for their cohort. BIT provided tutors information on how to amend text messages as part of the tutor workshop. As addressed in the section on Fidelity, while those who used this function reported this to be straightforward, project leads seldom felt they had enough time to adapt messages.

Project leads and study supporters finally emphasised a need for increased tailoring of messages. Text messages were often regarded as generic in their tone, and students seemed to be aware that all participating students in their cohort had received the same message. In one case, a project lead felt concerned that students might not engage or act upon the messages if they felt that texts were automated and not devised by their tutors:

'If students end up realising that text messages they're being sent are not written for them specifically or they're not even written by their tutor, a real person ... they'll see it as an automated text message' (project lead, curriculum area manager for maths).

Study supporter element

On the whole, the study supporter component was well received by participants. However, project leads stipulated that in order to be beneficial to students, parents/carers or other family members should be nominated to assume this role, rather than friends. Nevertheless, as highlighted previously, students in some cases showed resistance to involving their parent as a study supporter.

'I know a lot of them put a friend. Now, if I was whatever age, 16/17/18, and I got a text saying, "Why don't you encourage so and so with their maths revision?" I think I'd probably rather not. I'd rather just, you know, go round and chill ... I think that it should have been over 18s for the study supporter, and it should have been a bit more of a "strongly recommend" that it's your mum, a family member or carer' (project lead, head of English and maths).

The idea that study supporters should be provided with better guidance to best support students was raised in several cases. Where study supporters were a parent or carer, participants suggested it would have been useful to have received information about what the role would entail and how study supporters could help the student. Study supporters also highlighted a preference for information to be communicated directly by the developers or the college rather than the student:

'Um, I think maybe initially ... maybe just a little sheet with an introduction to the project ... but just so when the text messages appear it's kind of, you know, it's clear what is kind of expected ... at the moment it's a research project, but, you know, when it's a proper project if it keeps continuing, you know, what kind of the aim of the project is. So the parent understands that yeah, it is to support the child. That may be helpful' (study supporter, mother of student re-sitting GCSE maths).

Similarly, while links to relevant information were included where BIT felt study supporters might not be familiar with a topic, project leads suggested that messages directed at study supporters should contain additional resources to improve their understanding of a topic. As such, study supporters would be better equipped to open a conversation with the student about study concepts:

'Say we're doing Pythagoras this week and a parent says to the student ... "Oh, are you doing Pythagoras this week?" And if the student then says, "Yeah, do you know anything about it?", and if you say, "no", it seems a bit ridiculous mentioning that subject. I think almost like if the study buddy had extra text saying, "Look, we're doing this, this week, or we're gonna send a text next week about this, here's a link to where you can learn about it."' (project lead, senior maths teacher).

Limited discussion about text messages

Overall, a common point raised by project leads was that students rarely or never discussed taking part in the intervention and receiving texts. Project leads were often unaware of which students had signed up to participate and seldom sought feedback from students about the intervention. BIT did not encourage project leads and tutors to ask students about the intervention to minimise spill-over as they were meant to be blind to allocation. Nevertheless, interviewees also reported that students did not approach tutors as a result of encouragement within the messages to seek support:

'I suppose we haven't stopped every so often and gone, "Oh, Project Success, how's it going?" ... it was just something that we were asked to set up, which we did, and then we were told we didn't have to have anything more to do with it, so we haven't, it's just been ticking along. And our students haven't voluntarily said, "Oh wow, this, these texts are brilliant and they're really making me work", or there's been no positive or negative or anything' (project lead, director of English).

Cost

This section estimates the cost to colleges of taking part in the intervention, assuming it had been delivered without external funding. Our estimate includes:

- text messaging costs;
- costs for BIT to provide training;
- costs for BIT to develop website and online sign-up form;
- costs for BIT to devise and review texts;
- costs for BIT to provide support to colleges; and
- the costs of materials needed to advertise the project.

Our estimates assume the following:

- the number of colleges delivering Project Success per year is 31 (exactly as in the trial);
- the average number of students participating per college is 92.26 (exactly as in the trial);
- all students within a given college adopt the same treatment model;
- training would be repeated on a three year basis (based on staff turnover within maths and English departments);
- BIT developing the website (video, sign-up form) would be a one-off cost; and

- text messages would be tailored each year (messages are written based on lesson content and important reminders such as exam dates, which would need to be revised for each new cohort).

Costs for the two intervention arms sending texts to one group of people—(1) students only or (2) study supporters only—are identical. These are presented in Table 18. It shows the costs of delivering either of these arms over three years, and the total cost per pupil per year over three years.

Table 19 shows the costs of delivering the students and study supporters arm over three years, and the total cost per pupil per year over three years. This is more expensive to deliver than the other two as it costs more to send text messages to both students and study supporters. Aside from this extra cost, all other costs of the intervention were equivalent across each arm. Delivery of student only arm or study supporter only arm would result in a per pupil cost of £12.75 per year over three years. Delivery of students and study supporters arm would result in a per pupil cost of £14.67 per year over three years.

Table 18: Cost of delivery—students only arm or study supporters only arm (cost of each arm is identical)

Item	Type of cost	Cost	Total cost over 3 years	Total cost per pupil per year over 3 years
Training: cost of BIT staff time to deliver training workshop	Start-up cost per college	£677.00	£677.00	
Training: travel and subsistence costs for BIT	Start-up cost per college	£50.18	£50.18	
Cost of BIT developing website (video, sign-up form)	Start-up cost per college	£174.19	£174.19	
Printing / photocopying	Running cost per college	£11.23	£33.69	
Cost of BIT devising and reviewing text messages	Running cost per college	£666.68	£2,000.03	
Responding to queries and opt-out requests	Running cost per college	£21.52	£64.56	
Text messaging costs at standard messaging rate tariff	Running cost per college	£176.86	£530.58	
Total			£3,530.23	(£3,530.23/3/92.26) = £12.75

Table 19: Cost of delivery—students and study supporters arm

Item	Type of cost	Cost	Total cost over 3 years	Total cost per pupil per year over 3 years
Training: cost of BIT staff time to deliver training workshop	Start-up cost per college	£677.00	£677.00	
Training: travel and subsistence costs for BIT	Start-up cost per college	£50.18	£50.18	
Cost of BIT developing website (video, sign-up form)	Start-up cost per college	£174.19	£174.19	
Printing / photocopying	Running cost per college	£11.23	£33.69	
Cost of BIT devising and reviewing text messages	Running cost per college	£666.68	£2,000.04	
Responding to queries and opt-out requests	Running cost per college	£21.52	£64.56	
Text messaging costs at standard messaging rate tariff	Running cost per college	£353.72	£1061.16	
Total			£4,060.82	(£4,060.82/3/92.26) = £14.67

Costs over time

Most of the costs come from running costs, which would be the same each year. Nevertheless, the costs for delivering Project Success are expected to reduce over time since colleges would not need to attend training each year. Based on staff turnover within maths and English departments, it is assumed training would be repeated on a three-yearly basis. It is also assumed BIT would only develop the sign-up form and videos once.

Table 20 shows the cumulative cost per-pupil for student only and study support only arms increased by £9.50 from Year 1 to Year 2 and by the same amount from Year 2 to Year 3. The cumulative cost per pupil for students and study supporters arm increased by £11.42 from Year 1 to Year 2 and increased by the same amount from Year 2 to Year 3. The higher costs for the study supporters arm are due to the higher text messaging costs; all other costs are the same across the trial arms.

Table 20: Cumulative costs of students only arm, study supporters only arm, and students and study supporters arm per student over three years (assuming delivery over three years)

	Year 1	Year 2	Year 3
Students only arm	£19.27	£28.77	£38.26
Study supporters only arm	£19.27	£28.77	£38.26
Students and study supporters arm	£21.18	£32.60	£44.01

College staff time spent on Project Success

Table 21 displays the time spent on Project Success by college staff. Participating colleges provided a breakdown of staff time spent on Project Success and the average was calculated across all colleges that submitted activity data ($n = 16$, 52%). On average, college staff spent around six working days across the year (46.28 hours) on the delivery of Project Success. Staff reported spending most time on delivering the project during the normal college day (22.78 hours on average), which included initial student briefing and sign-up as well as an ongoing requirement to check content of text messages throughout the year.

A large proportion of the time spent on Project Success by colleges was used in training to deliver the intervention (18.19 hours on average). This was likely due to several staff members from each of the senior leadership team, maths, and English departments attending the initial training workshop facilitated by BIT and potentially spending time on cascading the training to colleagues who were not in attendance. Colleges also spent time in preparation for Project Success, including providing BIT with a schedule of work as well as time spent on liaison and communications about the project, internally and with the developers. A small amount of time was spent on delivering Project Success in addition to the normal college day, as well as on other unspecified activities (1.56 and 0.94 hours respectively).

Table 21: College staff time spent on Project Success (hours)

Activity area	Average (mean) staff time spent in hours per college
In training to deliver Project Success	18.19
Preparation for Project Success	1.75
Liaison and communication	1.06
Delivering Project Success during the normal college day	22.78
Delivering Project Success in addition to the normal college day	1.56
Other	0.94
Total staff time	46.28

Conclusion

Table 22: Key conclusions

Key conclusions
1. There is no evidence that the Project Success intervention had any impact on the GCSE English or maths re-sit pass rate.
2. There is no evidence that the Project Success intervention had any impact on the attendance of students re-sitting GCSE English or maths.
3. The intervention did not have a differential impact on the GCSE re-sit pass rate by gender, by eligibility for free school meals at 15, by the subject being re-examined, or by the number of re-sits being taken.
4. The use of mobile phone technology was perceived as a highly appropriate, effective, and low risk means of engaging with the target student cohort, though mobile phone use was less popular among study supporters.
5. There were significant limitations to the programme's ability to engage those who need it the most as it was the highly motivated students that were more engaged with their studies and with college generally who were more likely to sign up to the intervention.

Impact evaluation and IPE integration

Evidence to support the logic model

The logic model was not refined as a result of the trial. However, findings from qualitative interviews with college project and subject leads, students, and study supporters, conducted as part of the IPE, provide mixed evidence in support of the intervention logic model.

In line with the intermediate outcomes identified in the logic model, most students stated that receiving BIT text messages prompted increased engagement in revision outside of college. Yet, the extent of additional revision reported varied considerably among participants. According to students, text messages sent either to them or their study supporters encouraged revision by functioning as reminders and planning tools and by providing useful web resources. Motivational texts were perceived to be less useful. Interviews with students and college tutors, however, suggest that this increased engagement in revision outside of college did not translate into increased engagement in maths and English classes.

In the logic model, students' confidence is expected to improve as a result of increased engagement with their studies and in conversations around these with peers and study supporters. Students mentioned feeling more confident about exams they had completed and about upcoming exams, although project leads and study supporters did not comment on increased confidence in students. In support of the logic model, most students and study supporters reported an increase in the frequency of communication around maths and English studies. Study supporters attributed this increase to the texts acting as a 'conversation starter' and improving their understanding of the content covered by students in college, thus facilitating discussion and the provision of help. However, where the relation was already supportive, or, in contrast, characterised by infrequent contact or distance, or where study supporters had very limited understanding of the subject, text messages did not appear to improve communication and support.

The logic model for Project Success includes increased attendance or reduced drop outs as intermediate outcomes with prior evidence suggesting that attendance promotes improved GCSE results, whether re-sits or first-time attempts. While some project leads noted increases in lesson and exam attendance for maths and English as well as decreases in drop outs, these were often considered small and difficult to attribute to the intervention, especially as colleges usually employed additional measures to increase attendance.

Interpretation

This trial found no evidence that texting students, their study supporters, or students and their study supporters had any impact on the likelihood of passing a GCSE English or maths re-sit exam in 2018. Furthermore, subgroup analysis showed that the intervention had no effect on pupils who had been eligible for free school meals at the end of KS4 or during the preceding six years, and the effect did not differ for these pupils relative to their peers. Similarly, there was

no differential impact on the pass rate between males and females, by subject being re-sat, or by the number of exams being re-sat.

This contrasts with the evidence from a recent evaluation of a similar intervention (Miller et al., 2016), Texting Parents. This evaluation found statistically significant impacts for secondary school pupils in both English and maths. This may in part be explained by the secondary outcome, attendance, a key interim output of the intervention. Miller et al. (2016) also found a statistically significant reduction in absenteeism. As we have no evidence to suggest that this interim outcome has been affected by the intervention for this evaluation, this may explain the null findings on the primary outcome.

Importantly, the impact evaluation suggests that the intervention had no impact on college attendance in any of the intervention arms. This contrasts with findings from two recent studies of similar texting interventions (Miller et al., 2016 and Groot et al., 2017) which both found evidence that such interventions may improve college attendance, which is hypothesised to improve attainment. Kraft and Rogers (2015) argue that teacher-to-parent communication, such as texts, may improve college attendance. As there is evidence that this intervention had no impact on college attendance, this would suggest that changes in the primary outcome were unlikely. Though the content of text messages was not exclusively designed to encourage attendance, this may in part explain the null findings of the primary analysis. A further explanation of the limited impact on attendance may arise from attendance being measured over the entire academic year rather than over the period in which the trial was active. It is possible that that attendance was highest outside of the trial period at the start of the college year and fell over the course of the trial with time (Department for Education, 2019). If so, overall attendance within the trial could be overestimated affecting the analysis of the impact of the intervention on attendance.

Moreover, the general content of the text messages may not have maximised the opportunity to engage with, and motivate, the students or study supporters, whether it be to attend college or revise outside of college. Message content was generic and the sense of automation which students detected is likely to have led to a lack of engagement with the material and hence diminished the impact and led to a low effect size. The delivery of content was not deemed by some college staff as universally suitable to all students or study supporters, with some participants describing variation in individual preferences for delivery times or frequencies, which could affect their level of engagement with texts and consequent motivation. Tutors highlighted a need for more tailoring though they did not suggest a lack of training for doing so was an issue. Similarly, college-wide engagement with the texting intervention was low, potentially missing a further opportunity to maximise the impact of the regular delivery of motivational messages. These issues are a likely consequence of scaling up the intervention previously used in smaller trials, posing a challenge to tailored delivery and personalisation of messages within a larger number of colleges.

Although these results provide no evidence of an effect on the pass rate, it should be noted that the impact evaluation was reliant on a binary outcome of pass/fail at GCSE re-sit. This did not allow the impact evaluation to inspect for more finely graded improvements in student outcomes meaning a considerably large step-change in attainment was required in order to detect a difference in academic performance between study arms. It is possible that improvements in student attendance or overall academic performance might have been facilitated by the intervention; project leads did report improvements in mock results where raw marks on a continuous scale were available as opposed to categorical grade boundaries used in the impact evaluation. Project leads also reported changes in drop-out rates in some colleges, though they were reluctant to attribute this to the intervention without formal evaluation.

The results of this trial may not be fully generalisable however, and in different populations under different conditions a different result may be possible. This is because the student survey at sign-up suggests that those who opted into the trial were more likely to be highly motivated and generally more engaged with their college studies. Although these students were randomised to provide each study arm with similar levels of engaged students, this self-selection and formation of a more homogenous sample did not test for the effect of the intervention within a representative cross-section of prior grade 3-achieving students.

Limitations and lessons learned

The IPE of Project Success presented three key challenges. First, the recruitment of students and study supporters for in-depth interviews was particularly difficult, with few expressing the intention to take part and some of those who did, especially students, cancelling the interview or not showing up. While it was originally planned that data collection and analysis would include student-study supporter dyads, difficulties in securing participation meant that only three pairs

were interviewed. Second, collecting information from colleges both at the outset, for student ULNs, and at the end of trial, for cost and attendance data, was extremely difficult. This was primarily an issue with time constraints among college staff and project leads and the need to involve staff from different teams within the college where project leads had less influence. Staff turnover at participating colleges, including multiple changes to the project lead in some of the settings, complicated communication between the research team and colleges throughout the duration of the intervention and evaluation. Lastly, 'business-as-usual' approaches were not systematically documented as part of the IPE. Colleges were instructed to minimise texting to students unconnected to the trial and limit the content of texts to procedural issues yet interviews indicate that text messages were used in some colleges to encourage attendance, although the extent to which this happened is not known. It is also not known whether similar programmes aimed at improving attainment were being offered to students while they were active participants in this trial. In terms of the challenges faced by the impact evaluation, the interpretation of the findings from Project Success needs to account for the issue of contamination across study arms which may limit our ability to test for effect size of the intervention. There were a number of students in the control arm who received text messages, which may suggest that they were also nominated as study supporters, though contamination may have occurred for other reasons. This diminishes the power we have to test for an effect between the control group and all three intervention arms.

Although the trial does not suggest any effect on improving the GCSE re-sit pass rate via any of text messaging programmes used, or across any of the subgroups investigated, there are a number of modifications to the delivery of the intervention and the design of the trial which may induce a measurable effect in future. Regarding delivery of the intervention, the content, length, frequency, and delivery of the text messages to students or their supporters could be better tailored to the needs of the individuals. For instance, some participants regarded the arrival of a text message on a Thursday as a timely reminder to revise over the weekend, whereas others would have preferred to receive text messages at the start of the week to motivate them for the days ahead at college. This suggests that future trials ought to collect information at sign up on the preferred time of text message delivery. In terms of content, students and project leads preferred to receive factual information and links to study materials rather than motivational content. A minority of project leads did attempt to tailor messages, but these were generally conveying information around timetables and mock exams. Overall, it could be useful if more project leads tailored the content of text messages for their students to diminish the sense that text messages are uniform and automated, which may contribute to them not being acted upon. The practical, time-consuming implications of tailoring ought to be factored into the design of a future intervention.

The IPE suggests that the role of 'study supporter' ought to be made clearer in any future trial. It was felt that the supporter's role should be explained by BIT to the student themselves rather than college staff, again, to tailor the approach and support to the needs of the individual student. Further support could be provided to the student in choosing a study supporter given that 14.5% of students who opted out at recruitment did so because they had difficulty thinking of a study supporter. Additionally, the nomination of the study supporter could be more tightly controlled to generate impact, as, in many instances, the nominee was unable or unwilling to provide support. The inclusion of a study supporter arm was in itself problematic to recruitment as nearly one in five students who opted-out stated difficulties with nominating a supporter.

The roll-out of the intervention within colleges ought to have a higher profile. This trial largely adopted a light-touch approach where once the intervention was in place there was little discussion between tutors and students over the content of the messages. The lack of tutor-student discussion was a deliberate feature of the trial to attempt to minimise crossover. It is possible that a more general openness about regular delivery of messages supplemented by ongoing discussion of the content of the text messages within an intervention (that is, outside of a trial) may be able to capitalise on the motivational content so that the message becomes less likely to be ignored. It is possible that a more general openness about regular delivery of messages supplemented by ongoing discussion of the content of the text messages may capitalise on the motivational content so that the message becomes less likely to be ignored.

With respect to improving the impact evaluation of the trial, the study design deployed individual randomisation within colleges, which contributed to the moderate to high levels of contamination found across study arms, with some students receiving text messages where they were not supposed to or where students nominated themselves as study supporters. To mitigate, further evaluation designs should consider randomising at a college level to avoid such contamination issues, or exclude these students prior to randomisation. One major limitation of the design was the reliance on a binary outcome of pass/fail at GCSE re-sit. This did not allow the impact evaluation to inspect for more finely graded improvements in student outcomes meaning a considerably large step-change in attainment (at least one whole grade) was required in order to detect a difference in academic performance between study arms. RRR's were converted to Hedge's *g* effect sizes to define the months' progress that the treatment provided over the control group.

Furthermore, although our data indicates whether texts were sent or not, we cannot know if the individual received and read the content of the message. This is a limitation of the analysis assessing the impact of the intervention for different dosage thresholds. Difficulties with the collection of attendance data from colleges meant it was not possible to differentiate between attendance over the course of the entire year and attendance during the period over which the trial took place. This limitation may undermine the impact of the intervention on attendance if attendance varied between the pre-trial and trial study periods.

Lastly, with regards to any future work using mobile phones and texting technology, it will be important to consider at the study design stage how relevant the SMS medium will be to young people during the course of the trial. There has been a proliferation in messaging platforms since Project Success was devised, most of which use internet connections rather than traditional telephony. Future study designs ought to consider how young people use their phones in an era when SMS could be considered dated and potentially less engaging to young people.

Future research and publications

There are currently no plans for future research or publications. Given the mixed (and often weak) evidence of effectiveness of this type of intervention to date, a meta-analysis would be very timely to properly judge the weight of evidence in favour (or not) of texting as a means of improving student outcomes.

References

- Angrist, J. and Imbens, G. (1995) 'Identification and Estimation of Local Average Treatment Effects', *Econometrica*, 62 (2), pp. 467–475. <http://www.nber.org/papers/t0118>
- Angrist, J. D. and Pischke, J. (2009) *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton: Princeton University Press.
- Department for Education (2016) 'Level 1 and 2 Attainment in English and Maths by Students Aged 16–18: Academic Year 2014/15', main text: SFR15/2016. <https://www.gov.uk/government/statistics/level-1-and-2-englishand-maths-16-to-18-students-2014-to-2015>
- Department for Education (2017) 'GCSE and Equivalent Results: 2016 to 2017' (provisional), main text: SFR57/2017. <https://www.gov.uk/government/statistics/gcse-and-equivalentresults-2016-to-2017-provisional>
- Department for Education (2018) 'Retention Measure, Completion and Attainment Measure and Tech Level Minimum Standards Tables', SFR03/2018 (data set). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/689932/SFR03_2018_Retention_CA_Tech_level_minimum_Standards_tables.xlsx
- Department for Education (2018) Schools, Pupils and Their Characteristics. <https://www.jcq.org.uk/examination-results/gcses/2018/main-results-tables/gcse-full-course-results-summer-2018>
- Department for Education (2019) Pupil absence in schools in England: 2017 to 2018, Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/787463/Absence_3term_201718_Text.pdf
- Dong, N. and Maynard, R. A. (2013) 'PowerUp!: A Tool for Calculating Minimum Detectable Effect Sizes and Sample Size Requirements for Experimental and Quasi-Experimental Designs', *Journal of Research on Educational Effectiveness*, 6 (1), pp. 24–67, DOI: 10.1080/19345747.2012.673143
- Education Endowment Foundation (2018) 'Statistical Analysis Guidance for EEF Evaluations', London: Education Endowment Foundation. https://educationendowmentfoundation.org.uk/public/files/Grantee_guide_and_EEF_policies/Evaluation/Writing_a_Protocol_or_SAP/EEF_statistical_analysis_guidance_2018.pdf
- Frontier Economics (2017) 'Further Education Workforce Data For England: Analysis of the 2015-2016 Staff Individualised Record (SIR) Data', report prepared for the Education Training Foundation. <http://www.et-foundation.co.uk/wpcontent/uploads/2017/06/FE-workforce-data-2017-FINAL.pdf>
- Gerber, A. S. and Green, D. P. (2012) *Field Experiments: Design, Analysis and Interpretation* (1st edn), New York: W. W. Norton.
- Griggs, J. (2017) 'Amended Protocol for the Evaluation of the "Project SUCCESS"', London: Education Endowment Foundation. https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Protocols/Post_16_-_Project_SUCCESS_AMENDED.pdf
- Groot et al. (2017) 'I Get By with a Little Help from My Friends: Two Field Experiments on Social Support and Attendance in Further Education Colleges in the UK', London: Behavioural Insights Team. http://www.behaviouralinsights.co.uk/wp-content/uploads/2017/06/Study-Supporter-Working-Paper_2017.pdf
- Hedges, L. (2007) 'Effect Sizes in Cluster-Randomized Designs', *Journal of Educational and Behavioral Statistics*, 32 (4), pp. 341–370, DOI: 10.3102/1076998606298043
- Impetus – The Private Equity Foundation (2017) 'Life After School: Confronting the Crisis'. https://impetuspef.org.uk/assets/publications/Report/2017-03-14_Impetus-PEFConfronting-Crisis.pdf

Joint Council for Qualifications (2018) 'Provisional GCSE (Full Course) Results – June 2018 (All UK Candidates - Aged 17 and Over)'. <https://www.jcq.org.uk/examination-results/gcses/2018/main-results-tables/gcse-full-course-results-summer-2018>

Kraft, M. A. and Rogers, T. (2015) 'The Underutilized Potential of Teacher-to-Parent Communications: Evidence from a Field Experiment', *Economics of Education Review*, 47, pp. 49–63.

McNally, S. and Wyness, G. (2017) 'Education and Skills: The UK Policy Agenda', CEP Performance Paper CEPEA041, Centre for Economic Performance, London School of Economics.

Miller et al. (2016) 'Texting Parents: Evaluation Report and Executive Summary', London: Education Endowment Foundation.

https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/Texting_Parents.pdf

Smith, N. and Ashworth, K. (2018) 'Statistical Analysis Plan for "Texting Students and Study Supporters" ("Project SUCCESS")', London: Education Endowment Foundation.

[https://educationendowmentfoundation.org.uk/public/files/Projects/Post_16_-_Texting_Students_\(Project_SUCCESS\)_SAP.pdf](https://educationendowmentfoundation.org.uk/public/files/Projects/Post_16_-_Texting_Students_(Project_SUCCESS)_SAP.pdf)

Wolf, A. (2015) 'Heading for the Precipice: Can Further and Higher Education Funding Policies Be Sustained?', London: Policy Institute, King's College London.

Wooldridge, J. M. (1995) 'Score Diagnostics for Linear Models Estimated by Two Stage Least Squares', in Maddala, G. S., Phillips, C. B. and Srinivasan, T. N. (eds), *Advances in Econometrics and Quantitative Economics: Essays in Honor of professor C. R. Rao*, Oxford: Blackwell (pp. 66–87).

Appendix A: EEF cost rating


Appendix Table 1: Cost Rating

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

Appendix B: Security classification of trial findings

OUTCOME: GCSE re-sit English/ Maths, all three intervention arms

Please use this template to assign a separate security rating for each primary outcome.

Rating	Criteria for rating			Initial score	Adjust	Final score
	Design	MDES	Attrition			
5	Randomised design	<= 0.2	0-10%	5		
4	Design for comparison that considers some type of selection on unobservable characteristics (e.g. RDD, Diff-in-Diffs, Matched Diff-in-Diffs)	0.21 - 0.29	11-20%		 Adjustment for threats to internal validity [-1]	4
3	Design for comparison that considers selection on all relevant observable confounders (e.g. Matching or Regression Analysis with variables descriptive of the selection mechanism)	0.30 - 0.39	21-30%			
2	Design for comparison that considers selection only on some relevant confounders	0.40 - 0.49	31-40%			
1	Design for comparison that does not consider selection on any relevant confounders	0.50 - 0.59	41-50%			
0	No comparator	>=0.6	>50%			

Threats to validity	Threat to internal validity?	Comments
Threat 1: Confounding	Moderate	<p>This was a randomised trial and allocation was performed independently. It was powered to detect an MDES of 0.15 with very low attrition (1%). There was some imbalance between groups on certain characteristics, but this was likely due to chance. The evaluation team conducted appropriate sensitivity tests which did not change the conclusions of the primary analysis.</p> <p>There is a risk that students who self-selected for the sample were highly motivated and this may have impacted outcomes. Although this was balanced between groups, this still carries a risk of reducing the opportunity for the intervention to have an effect, as they may or may not have been the most appropriate students to participate.</p>
Threat 2: Concurrent Interventions	Moderate	Insufficient data has been provided on concurrent interventions in intervention settings and usual practice in control settings. The IPE reports intervention schools using various similar strategies (including texting), but this is not transparently reported or controlled analytically, as relevant data were not collected consistently.
Threat 3: Experimental effects	Moderate	Each intervention arm showed evidence of contamination, as did the control group. In two of the three intervention arms, 17% of participants did not receive the intended treatment. 7% of the control students received unintended texts. This contamination is not controlled for analytically and the reasons for it are unclear. The extent to which students nominated each other to be study supporters is also unclear (95.7% missing responses to relevant survey question).
Threat 4: Implementation fidelity	Moderate	A proportion of participants did not receive the number of texts intended, while some of the control group received intervention texts. There was no explanation provided as to why this was, however, the

		number of texts received (compliance) did not appear to moderate effects.
Threat 5: Missing Data	Low	Missing data was extremely low (1%) and so the complete case analysis should be unbiased.
Threat 6: Measurement of Outcomes	Low	GCSE Maths and English were the outcomes and were independent of the intervention.
Threat 7: Selective reporting	Low	No evidence of selective reporting or of unjustified changes from SAP.

- **Initial padlock score:** 5 Padlocks – MDES of .15 and very low attrition (<1%).
- **Reason for adjustment for threats to validity:** 4 Padlocks – Removed one padlock due to four moderate risks (direction of bias unclear/ inconsistent).
- **Final padlock score:** initial score adjusted for threats to validity = 4 Padlocks

Appendix C: Effect size estimation (secondary outcome)

Appendix table 2: Effect size estimation (secondary outcome)

Outcome	Unadjusted differences in means	Adjusted differences in means	Intervention group		Control group		Pooled variance	Population variance (if applicable)
			n (missing)	Variance of outcome	n (missing)	Variance of outcome		
Attendance: Students only	-0.21	-0.01	653 (296)	0.40	648 (297)	0.56	0.48	n/a
Attendance: Supporters only	0.12	0.01	653 (286)	0.82	648 (297)	0.56	0.70	n/a
Attendance: Students and supporters	-0.02	0.00	657 (289)	0.86	648 (297)	0.56	0.73	n/a

This work was produced using statistical data from ONS. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

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

To view this licence, visit <https://nationalarchives.gov.uk/doc/open-government-licence/version/3> or email: psi@nationalarchives.gsi.gov.uk


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

This document is available for download at <https://educationendowmentfoundation.org.uk>



The Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP

<https://educationendowmentfoundation.org.uk>

 @EducEndowFoundn

 [Facebook.com/EducEndowFoundn](https://www.facebook.com/EducEndowFoundn)