Evidence vs testimonial messaging approaches to adopt evidence-based programmes: a nimble RCT of the Embedding Formative Assessment project

Nimble Trial Report

April 2022

Kimberly Bohling, Pujen Shrestha, Eleanor Collerton, Tim Hardy, Alex Sutherland

The Education Endowment Foundation (EEF) is an independent grant-making charity dedicated to breaking the link between family income and educational achievement, ensuring that children from all backgrounds can fulfil their potential and make the most of their talents.

The EEF aims to raise the attainment of children facing disadvantage by:

- identifying promising educational innovations that address the needs of disadvantaged children in primary and secondary schools in England;
- evaluating these innovations to extend and secure the evidence on what works and can be made to work at scale; and
- encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus – Private Equity Foundation) and received a founding £125m grant from the Department for Education.

Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.

For more information about the EEF or this report please contact:

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
This project was designed and evaluated by a team from the Behavioural Insights Team: Kimberly Bohling, Pujen Shrestha, Jo Milward, Eleanor Collerton, Tim Hardy, and Alex Sutherland (Principal Investigator).
The project lead was Kimberly Bohling.

Contact details:
Behavioural Insights Team
4 Matthew Parker Street
London, SW1H 9NP

e: info@bi.team

1. Executive summary

The project

Embedding formative assessment (EFA) is a professional development programme delivered by SSAT which aims to improve pupil outcomes by embedding the use of formative assessment strategies across a school. The EFA scale-up evaluation is delivering between September 2021 and July 2023. This 'nimble' trial has been run alongside the scale-up evaluation. It aimed to test which of two marketing letters for the intervention (evidence vs testimonial) was more effective in encouraging the adoption of EFA, an evidence-based programme, by school leaders. Results from this trial would inform the recruitment strategy for scaling the programme.

This was a two-armed randomised controlled trial (RCT), a superiority trial comparing two treatments, rather than a treatment and control. Superiority trials assess whether one treatment performs better than the other, or if the two are equivalent. The RCT was randomised on a school level, and the outcome measures were interest in and purchase of the EFA programme. The trial was led by the Behavioural Insights Team (BIT) and ran between March and September 2021. Mailed marketing materials were sent to 2,000 headteachers at secondary schools in England. Half of the headteachers received the 'evidence-based' and half received the 'testimonial' version of the mailed marketing materials.

Table 1. Summary of findings

Research question	Finding
RQ1: Does testimonial or evidence-based marketing material perform better in encouraging school leaders to express interest in EFA?	● We find no statistically significant difference in response rates to testimony versus evidence-based marketing of the EFA programme in our primary model. ● The evidence-based marketing material resulted in 21 expressions of interest (EOIs) and the testimony mailers generated 11 EOIs. ● The pre-specified analytical model finds a difference of 0.9 percentage points in EOIs between the two types of messaging (95% CI: –0.1, 3.1). ● Overall, EOI rates were 1.6%.
RQ2: Does testimonial or evidence-based marketing material perform better in encouraging school leaders to purchase EFA?	● We find no difference in the rates of sales between evidence and testimony-based marketing materials. Two sales were recorded per arm. ● The pre-specified model finds a difference of 0.03 percentage points (95% CI: –0.17, 1.72).

Additional findings

Overall, only 1.6% of schools that received the mailed materials (n=32 in sample of 1,998) expressed interest in learning more about the EFA programme and only 4 schools went on to purchase the programme. However, it is important to note that the selected sample was 'cold' (i.e., they had not recently received EFA-specific marketing or expressed interest in the programme), so this particular sample was, perhaps, more difficult to engage than warmer contacts (e.g., schools that had more recently downloaded materials or attended an informational day about the programme).

The trial was affected by the 2020 partial school closures caused by the Covid-19 pandemic, so it is difficult to know whether the unusual context may have dampened response rates. A similar trial run as part of the National Tutoring Programme evaluation had a similar rate of expressions of interest (EOIs) (~1%) to email marketing in a similar time period, which suggests the response was typical for the time period and context (Harland et al., 2021).

On average, it cost £310 to generate each EOI (n=32) and £2,478 for each sale (n=4).

The letters were printed, assembled, and delivered as planned. We have no record of any letter being returned or otherwise marked as undelivered.

Interpretation & recommendations

There is currently little causal evidence about how to improve take-up of evidence-based programmes in schools. This trial adds evidence about the potential of using a routine part of the scaling process, marketing, as a way to bring more science into implementation science. We specifically sought to test the impact of communicating evidence to school leaders to understand if variants of this affected decision making. However, we caution that this trial tested a particular set of materials with a particular sample, and that when compared to marketing trials used in the private sector, the sample size is much smaller than would otherwise be used (Luca & Bazerman, 2020). The pilot study that inspired this research found US principals were more interested in a programme after being presented with testimony, and a similar trial run as part of the National Tutoring Programme evaluation found no difference between evidence and testimony in encouraging take-up of a particular tutoring provider (Harland et al., 2021).

In this study, we have demonstrated the feasibility of conducting a fairly low-cost 'nimble' trial to support the delivery teams' goal of scaling programmes and contribute to the evidence base around how to support scaling of education interventions. We recommend that EEF and other What Works Centres consider commissioning more of this type of trial. This would help us to understand which approaches are more effective in promoting the adoption of evidence-based programmes by schools.

2. Introduction

Background evidence

Although there has been a significant effort to produce better evidence about what works in education – which the EEF has been integral to – there is still very little known about how to encourage schools to adopt evidence-based interventions. This means that many well-researched interventions 'sit on the shelf' rather than being put to use to improve the lives of pupils. Therefore, insights generated from this research could help the EEF, researchers and practitioners based in the UK, and elsewhere, better design communications that promote the adoption of evidence-based practices, which might otherwise not achieve optimal uptake and impact.

Our research builds on an unpublished pilot survey conducted by Dr Todd Rogers, Professor of Public Policy at the Harvard Kennedy School. He wanted to explore why innovative ideas with research backing fail to be widely adopted and theorised that some leaders find anecdotes more motivating than evidence. He conducted an exploratory pilot survey of 200 US school principals in a large urban public school district about their interest in learning about a low-cost, well-evidenced absenteeism intervention – more principals indicated that they would be interested in learning about the intervention when they were told that a single principal recommended it, as opposed to being presented with evidence from three large-scale RCTs.

There is limited evidence in the UK context as to why anecdotal or testimonial recommendations are responded to more positively than when evidence is presented. One suggestion taps into ideas surrounding the messenger effect, whereby greater attention may be paid to an advert or information if the person relaying that information is seen to be similar to them, relatable or a credible source of information (Wood et al., 2005). Within an education setting, a testimonial from another educator regarding a programme may resonate more than evidence-based information.

Evidence-based information in relation to education programmes tends to be attended to and adopted more when the evidence presented aligns with the receivers' pre-existing beliefs (Coburn et al., 2009). There has been fairly limited research focusing specifically on head teachers' adoption of evidence-based approaches; however, one study found that a head teacher's response to proposed evidence-based initiatives is moderated by their existing knowledge around the topic area (Spillane et al., 2002).

The trial tests two variations of marketing materials for embedding formative assessment (EFA), a whole-school teacher development programme led by SSAT, which has strong evidence of effectiveness in improving pupil outcomes (Speckesser et al., 2018). An effectiveness trial found that in schools randomly assigned to implement the programme, pupils made the equivalent of an additional two months of progress on their Attainment 8 GCSE scores when compared to pupils in schools who implemented their business-as-usual instruction. The messaging was designed to be sent as a printed material via post, as we hypothesised that this messaging medium will be more effective than email for several reasons:

- Email messages are easily skipped, deleted, or filtered out. Sending the materials via post with recorded delivery ensures the message stands out and may be more likely to be opened and read.
- Email marketing is very common; whereas mail-based marketing now has a novelty factor.
- Printed materials have the potential to become 'social artefacts' – that is, they can be easily shared with others and become physical reminders. Research in the energy domain found that providing households with printed copies of their energy-use report resulted in them reducing their energy use (Schultz et al., 2007). Todd Rogers incorporated this into research focusing on reducing pupil absenteeism, by sending parents postcards, which displayed their child's absence count. These postcards were felt to have a longer shelf life and be more salient to parents in comparison to receiving a text message or email, which could be easily ignored. In one trial, these postcard absence updates reduced student absenteeism by 15% in comparison to a control group which only received a reminder about the importance of attendance (Rogers et al., 2017). Given that decisions about school programme purchases are often made collectively among senior leadership team (SLT) members, we felt the social artefact element of providing printed materials was particularly important.

The rationale for the study was to test whether certain marketing messages can encourage the adoption of an evidence-based programme by school leaders. We tested two message variations – 'evidence' and 'testimonial'. Specifically, the aim of this field trial was to assess which type of message is more effective in encouraging school leaders to adopt the EFA programme.

SSAT are currently running a multi-year scale-up of their EFA programme and aiming to increase the number of schools involved. This created an opportunity to explore further whether the type of marketing schools receive has an impact on the number of EOIs or sign ups by schools. Due to the multi-year nature of the scale-up, findings from this study can hopefully be incorporated by SSAT into their marketing plan for their next recruitment cycle.

Intervention

Table 2. Intervention description

Intervention name	EFA Mailer Trial: Mail-delivered testimonial and evidence-based marketing material
Why (theory/ rationale)	The rationale of this study was to explore whether testimony or evidence-based marketing material would result in differential expressions of interest (EOIs) and sales of embedding formative assessment (EFA). This would enable us to make recommendations about how to improve design of school recruitment materials, as well as make a contribution to the wider knowledge about the adoption of evidence-based practices at scale.
Who (recipients)	Head teachers of secondary schools received the mailers.
What (materials)	The two treatment groups received a letter and a supplemental programme description for EFA that was either constructed using solely testimonial (from a previous case study) or evidence-based support (from the previous EEF-funded evaluation). See Appendix A for intervention materials and the section below for further detail on intervention design.
What (procedures)	Schools received the intervention by mail using recorded delivery. If interested in learning more about the programme, schools were given three ways they could express interest: 1) download additional informational materials after providing contact details; 2) register online for an informational Open Day; 3) request a consultation. The letters contained web addresses that corresponded to each of the three EOI activities above. The pages were specifically set up for the trial and were not used or linked elsewhere. On each of the web pages, recipients were asked to provide their school's Unique Reference Number (URN) alongside their EOI in order to track outcomes more easily, and the school's URN was included on the letter to ensure the recipient had the number easily available. This sort of tracking is part of standard marketing practices for SSAT.
Who (provider)	BIT designed the intervention materials in collaboration with SSAT. The intervention was printed, assembled and disseminated via a printing/mailing service.
How (delivery mode)	The intervention was mailed to the identified schools and addressed to the headteacher using Signed For delivery.
Where (setting)	The intervention was delivered at the identified schools.
When & how much (dosage)	The marketing material was delivered once. It was timed to be delivered in March when schools may be planning for the next school year.
Tailoring (adaptation)	N/A

Elements of design

In order to rigorously test the effect of one type of marketing approach versus the other (see Figure 1), we designed the materials to be of the same length and formatting. This was to ensure that differences in response were not due to length or a more visually appealing design. We also distinguished the testimony/evidence features with call-out boxes and red font to maximise attention on this information. Some of the other key design features included:

- *Personalisation:* People respond better to stimuli that is specific to them and contains personalised information, such as directly addressing the letter to their name (Carmody & Lewis, 2006).
- *Simplification:* People are more likely to act on a message if it is easy to understand (Lasky-Fink et al., 2020). The letters were kept short with an easy-to-read message. The format also included features to make the key messages salient to the reader (boxes and bullet points).
- *Clear call to action:* Both letters included a clear call to action in the form of a bullet point list of next steps, making it easy for the recipient to know what they should do next. Responses to each call to action were also easy to measure.
- *Reduced friction:* The letters included short links in bold of the websites that the recipient would need to type out, making it easier for them to locate the relevant information. The letters also included the school's URN at the bottom of the page to make it easy to locate when expressing interest on the website.

One element to note in the design is that we needed to rely upon evidence and testimony that had already been collected. Each letter consisted of three pieces of information. We ensured both letters' first data point was in regard to impact on Attainment 8, which is an important metric of attainment that will be of interest to school decision-makers. However, we did not have comparable data points for the remaining two pieces of information presented, but this is precisely due to the nature of the two interventions we are presenting – evidence-based and testimonial data points can be inherently different. A case study/testimonial can speak in detail to an individual school's experience and will generally be very positive about the impact of the programme on their particular school, whereas an RCT will be generating average findings across many schools. We do not see the differences in findings presented as a limitation of the design, but a feature of how evidence and testimony vary in the types of findings they can produce.

Implementation agreements with SSAT

In order to assess the impact of the letters, BIT and SSAT collaboratively agreed on some marketing arrangements, which attempted to balance the needs of the trial with SSAT's marketing goals. First, it was agreed with SSAT that they would not directly market EFA further to the schools in the sample before the primary data collection window closed (28 May 2021). This was agreed, so we could largely attribute any EOIs in this period to the letter. This did impose some burden on SSAT's marketing efforts, but as the sample was not in their current sales pipeline, this was an acceptable arrangement. They were allowed to resume normal marketing efforts after 28 May.

EFA is not the only programme that SSAT offers to schools and a risk to the trial was that competing SSAT programmes had the potential to crowd out interest. It was agreed with SSAT that they could market other programmes to the schools in the sample, as long as the marketing was not targeted to headteachers. Any further restrictions would have placed a significant burden on their standard marketing practices, particularly for other programmes. It was possible that other marketing may have dampened interest in EFA, but as SSAT were blind to condition, we do not expect any differences in additional marketing between the two groups.

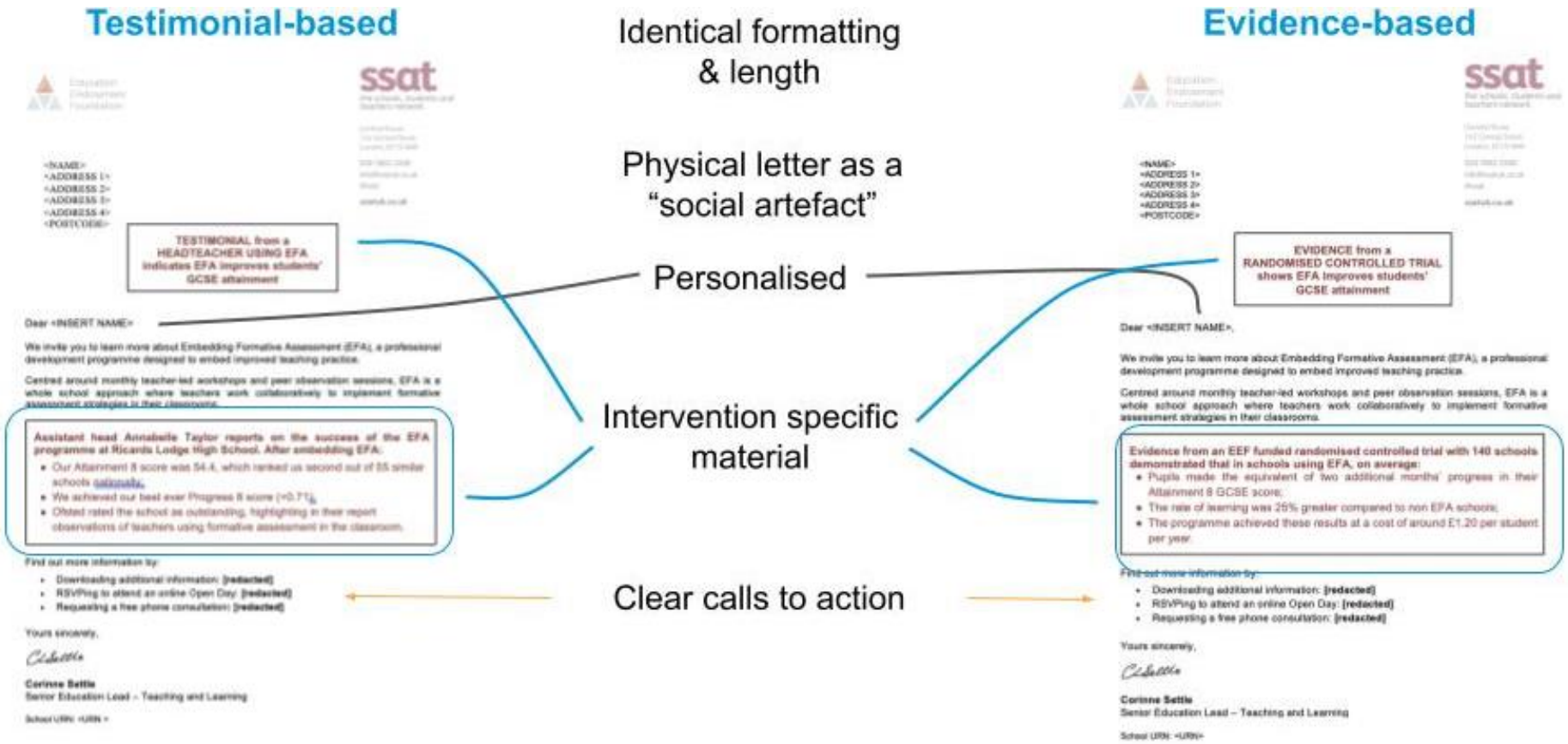
The trial and materials were designed December 2020–February 2021 in collaboration with SSAT. The letters were posted in March 2021. Further details on the project team, timeline, procedure on ethical review and data protection can be found in Appendix B, C, E and F, respectively. The pre-registration of this trial can be found on OSF at: https://osf.io/wvmrz. The trial protocol for this trial can also be found on the EEF website at: https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/embedding-formative-assessment-re-grant/

Research questions

RQ1 (primary): Does testimonial or evidence-based marketing material perform better in encouraging school leaders to express interest in EFA?

RQ2 (secondary): Does testimonial or evidence-based marketing material perform better in encouraging school leaders to purchase EFA?

Figure 1. Summary of intervention materials



3. Methods

Trial design

Details of the trial design are presented in Table 3 and the flow diagram for the trial is presented in Figure 2.

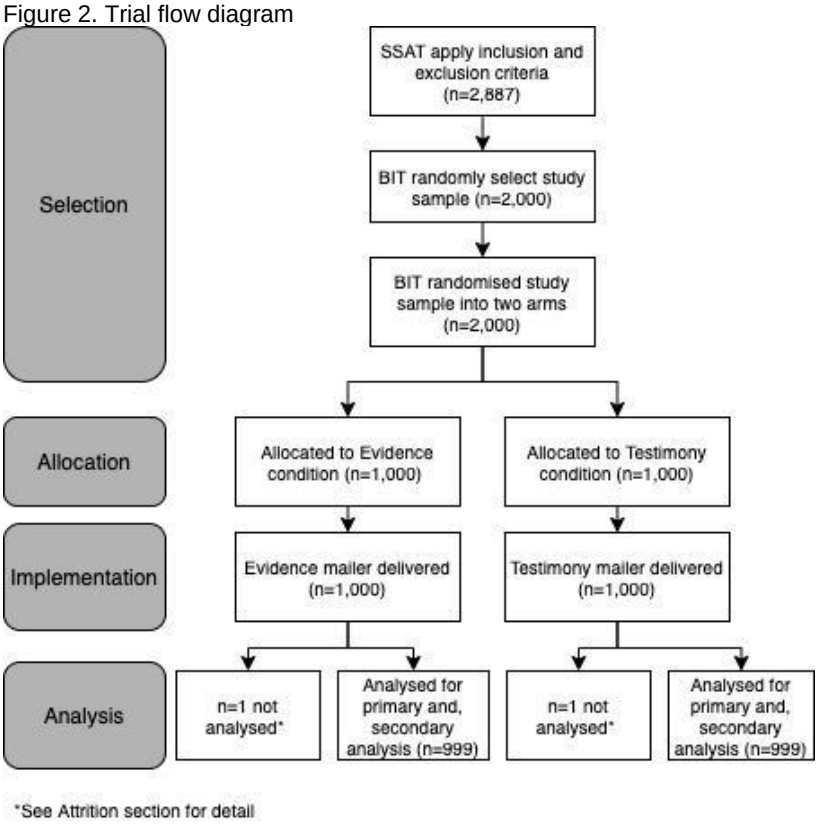
Table 3. Trial design summary

Trial design, number of arms		Two-arm, randomised control trial
Participants	Inclusion criteria	Secondary schools in England
	Exclusion criteria	Schools were excluded from the trial if they had previously used or expressed interest in EFA. Specifically, the exclusion criteria included: 1. Schools that had taken part in the EEF EFA research 2. EFA Mentor and Ambassador schools 3. Schools that are or have undertaken the EFA support package (toolkit + support from an EFA mentor) 4. Schools that have purchased the EFA toolkit (without support package) 5. Schools that expressed an interest (EOI) in EFA or requested a EFA consultation 6. Schools that have had staff members booked to attend or had attended EFA open days 7. Schools that have had a staff member download content from SSAT website related to EFA in an identifiable manner. We also excluded schools that had no listed address or headteacher in the Get Information About Schools Edubase dataset, as we would not be able to deliver or personalise the material. SSAT provided a list of 2,887 schools that met this inclusion/exclusion criteria. The randomisation process is described in full below.
	Target number	2,000 schools
Unit of randomisation		School
Stratification variables		N/A
Primary outcome	Variable	Expression of interest (EOIs) submitted by 28 May 2021 (slightly less than 3 months after the mailing was posted).
	Measure (instrument, scale, source)	EOI was recorded by SSAT. This is a binary indicator generated from whether the school expressed interest through one of the three options described above (download, Open Day registration, consultation request).
	Direct measure or proxy?	Direct
	Baseline measure	N/A
	Time of collection	16th March–28th May 2021 (rolling data collection)
Secondary outcome	Variable(s)	Purchase of EFA by 1 September 2021 (approximately 7 months after the mailing was posted).
	Measure (instrument, scale, source)	Purchase of EFA was measured by SSAT. This is a binary indicator.
	Direct measure or proxy?	Direct
	Baseline measure	N/A
	Time of collection	16th March 2021 to 1st September 2021 (rolling data collection)
	Adjusting for multiple comparisons?	No – only considers primary outcome

There was no untreated control condition in this trial, as it would have required SSAT to not market to a large number of schools for three months. This would have posed quite a significant negative impact on SSAT's marketing efforts and potentially their ability to reach scaling targets.

No changes were made to the trial design as set out in the pre-registered trial protocol.

Figure 2. Trial flow diagram



Participant selection

A 'cold' sample (i.e., schools that had not recently received EFA-specific marketing or expressed interest in the programme) was intentionally constructed in order to assess the impact of the letters. It is possible with a 'warmer' sample (e.g., schools that had recently downloaded information or attended an information session) that the letter may have simply served as a reminder and the letter's content would have had very little impact on behaviour. BIT and SSAT worked together to establish the school inclusion and exclusion criteria, which are provided in Table 3. The list of eligible schools was provided by SSAT (n=2,887). BIT were responsible for randomly selecting the 2,000 schools to be included in the trial (see Sample size for more detail).

Outcome measures

The study was designed with the proximal measure (EOIs) as the primary outcome, as there are potentially a number of steps and different pathways between expressing interest in the programme and eventually purchasing the programme. We felt it was important to establish a causal link between the intervention and the most direct outcome measure first, before assessing impact on our distal outcome of sales.

EOIs and sales were recorded by SSAT, using an internal customer relationship management platform. EOI was a binary indicator generated from whether the school expressed interest through (1) downloading additional informational materials, (2) registering online to attend an informational Open Day, and/or (3) requesting a consultation. The different modes of expressing interest were specifically designed to allow for varying levels of interest/engagement with a simple download of informational materials offering the lowest level of engagement and requesting a consultation as the highest level of engagement. EOIs were measured at two time points: 28 May (primary outcome) and 1 September (exploratory outcome). Sales was a binary indicator generated from whether the school purchased the EFA programme or not by 1 September 2021. Timelines for outcomes were decided in collaboration with SSAT and the EEF, based on:

(i) what SSAT believed would be reasonable for schools to respond to:
(ii) SSAT's data on how long it takes for schools to convert from marketing to EOI, and from EOI to purchase;
(iii) timelines for the project itself; and

(iv) the 'rhythm' of the school year (i.e., when schools are typically thinking about planning for the next school year).

Both measures were recorded as described in this trial's protocol.

The mailings were addressed to the headteacher to increase the likelihood that they were opened (as opposed to having no specified recipient and potentially being discarded unopened). However, all outcome data was recorded at the school level. This was to take into account that decisions about teacher professional development may be made by other leaders within a school, so EOIs and sales may actually come through school leaders other than the headteacher. Also as described in the background section, letters can be viewed as a 'social artefact' that can be passed around to others or brought to meetings to share, and we expected that this sort of sharing was likely within a school setting (although it is possible the Covid context may have discouraged this sort of sharing of physical documents).

Sample size

The estimate of 2,000 schools was based on the number of schools needed in order to achieve a minimum detectable effect size (MDES), in line with expectations for this sort of light-touch intervention. The agreed sample size also took into account the limited sample (2,887 eligible schools) and that the trial imposed limitations on SSAT's marketing efforts with the study sample (see Implementation for details on agreed marketing limitations). Table 4 presents the assumptions and MDES for a range of response rates (EOIs) to the lowest performing arm. Codes for power calculations are presented in Appendix G.

Justification for the assumptions in the Table 4 are as follows:

- **Alpha and power:** These are standard assumptions provided by the EEF (see 2019 trial protocol template).
- **One or two-sided test?** A two-sided test was performed to err on the side of caution. There is little existing evidence relating to the effect of interventions of this type of marketing material, so we cannot assume the direction of any effect that we might observe.
- **Number of schools:** The sample size was partially determined by not wanting to limit the market for SSAT too much for EFA, as they agreed to changes in their marketing practices for the duration of the trial (see Implementation agreements' section for further detail). Further, as described in the exclusion criteria, we excluded a number of schools based on previous engagement with SSAT/EFA, which limited our eligible sample.
- **Baseline:** SSAT does not routinely engage in cold-call mass mail-outs, so we do not know what a typical response rate is. We proposed a sample of 2,000 schools with an equal allocation between the two arms. The MDES for a range of response rates (EOIs) to the lowest performing arm is presented in Table 5.

Table 4. Minimum detectable effect size (MDES)

		OVERALL
MDES (Cohen's h)		**0.13**
Alpha		0.05
Power		0.8
Number of schools	Testimonial intervention	1,000
	Evidence-based intervention	1,000
	Total	2,000
Software used for calculations		Rstudio 1.2.5001 & R 3.6.1

In Table 5 we present what the MDES would be, substantively depending on the response rate in the lowest performing arm. That is, if the lowest performing arm had an EOI rate of 1%, then the trial would be powered to detect a difference of 1.6 percentage points (the higher performing arm has a response rate of at least 2.6%).

Table 5. Substantive minimum detectable effect size (MDES)

Baseline (EOI rate in lowest performing arm)	**1%**	**5%**	**10%**
MDES*	1.6pp	3.1pp	4.1pp

**pp = percentage points*

Randomisation

SSAT provided a list of 2,887 schools that met the inclusion/exclusion criteria. BIT randomly selected 2,000 schools from the list to be included in the trial (see Table 9 of balance checks in the Findings section for further detail on the schools selected).

In addition to the random sampling from the initial study population, schools were randomly assigned into the two treatment arms. Random allocation was a 50/50 allocation and was conducted in February 2021 using RStudio with R version 4.0.3 (see Appendix G for randomisation code).

BIT designed and undertook the randomisation process and sent two sets of letters (with publicly available recipient name and address already merged into the correct letter) directly to our partner printing agency. This guaranteed that SSAT would be blind to allocation and ensured that if a school submitted an EOI, any subsequent sales interactions would not be influenced by SSAT's knowledge of the treatment arm. Additionally, schools were blind to allocation because they did not know they were part of a trial. Our mailer partner was responsible for the printing and delivery of the letter.

Statistical analysis

The statistical analysis approach is set out in Table 6. All analyses are an intention-to-treat and were carried out using Rstudio 1.2.5001 & R 3.6.1. We included six covariates in our analytical model that we hypothesised might be associated with our outcome, and we expected to improve precision. Further technical detail on the analytical approach can be found in Appendix D.

We have opted to convert the log-odds into percentage points change (instead of odds ratios) in order to make findings directly interpretable and more accessible. See Appendix D for details on conversion.

Table 6. Impact analysis summary

RQ	**Sample**	**Dependent variable**	**Independent variable**	**Control variable(s)**	**Analytical method**	**Interpretation**
RQ1 (primary)	Secondary schools that do not meet our exclusion criteria (see Table 2)	EOI in EFA after 3 months (binary)	Treatment assignment (evidence=1; testimonial=0)	Proportion of FSM eligible pupils, school size, opportunity area, region, whether the school is part of a multi-academy trust, and Ofsted rating[a]	Logistic regression	The evidence-based marketing intervention affects EOI rate by X percentage points in comparison to the testimonial marketing intervention.
RQ2 (secondary)	Secondary schools that do not meet our exclusion criteria	Purchase of EFA after 7 months (binary)	Treatment assignment (evidence=1; testimonial=0)	Proportion of FSM eligible pupils, school size, opportunity area, region, whether the school is part of a multi-academy trust, and Ofsted rating	Logistic regression	The evidence- based intervention affects purchase rate by X percentage points in comparison to the testimonial intervention.

Note: [a]All covariate data except Ofsted rating was obtained from Get Information about Schools (https://www.get-information-schools.service.gov.uk/). Ofsted data was obtained from ***https://www.compare-school-performance.service.gov.uk/***

4. Findings

Summary

- We find no statistically significant difference in response rates to testimony versus evidence-based marketing of the EFA programme in our primary model.
- The evidence-based marketing material resulted in 21 EOIs and the testimony mailers generated 11 EOIs. When analysed, this translates to a difference of 0.9 percentage points (95% CI –0.1, 3.1), However, this difference is not statistically significant on a 95% confidence level (p-value = 0.100).
- We also find no difference (statistically or practically) in sales rates between the two conditions. Two sales were recorded per arm.
- The letters were printed, assembled and delivered as planned. We have no record of any letter being returned or otherwise marked as undelivered.

Outcomes and analysis

The results of the primary analysis are presented in Table 7 and Figure 3. Overall, the rate of EOIs was 1.6% approximately three months after the mailings were posted. The raw rate of EOI for the testimonial-based marketing was 1.1% (11 EOIs) and for the evidence-based marketing it was 2.1% (21 EOIs). There was also one EOI that came from a multi-academy trust with two schools (each randomised to a different arm), which could not be assigned to either condition. That EOI and the two schools have been excluded from the analytical sample. The pre-specified full model finds a difference of 0.9 percentage points, but that difference is not statistically significant on a 95% confidence level (p-value = 0.100).[1] When we run the model removing all covariates other than condition, the result changed marginally, showing a slightly stronger effect for the evidence-based condition, but the result still does not reach significance at the 95% confidence level (p-value = 0.079).

The regression table for both models and process for converting log odds to percentage points can be found in Appendix D.

Table 7. Primary analysis results

Primary analysis: Expressions of interest (EOIs) (3 months) (n = 1,998)			
	Unadjusted mean (95% CI)	Log odds[a] (95% CI)	Treatment effect converted to percentage points (95% CI)
Evidence *n = 999*	2.1% (1.2%, 3.0%)	0.6 (–0.1, 1.4)	0.9 (–0.1, 3.1)
Testimonial *n = 999*	1.1% (0.5%, 1.7%)	–	–

Notes: p-value for this result is p=0.100. [a]Odds ratios can be calculated by exponentiating the log-odds coefficient.

We intentionally gave schools a variety of options to express interest at varying levels of engagement, ranging from fairly low (download materials) to fairly high (requesting a consultation). Table 8 tabulates the highest level of engagement for each school in the full sample.[2]

[1] The pre-specified model includes the following covariates: Proportion of FSM eligible pupils, school size, opportunity area, region, whether the school is part of a multi-academy trust, and Ofsted rating

[2] Given overall low numbers of EOIs, we are intentionally not presenting the types of EOI by arm.

Table 8. Rate of types of EOI for the full sample

Type of EOI	Proportion of total EOIs
Download materials	25 (78%)
Open day	2 (6%)
Request consultation	5 (16%)
Total	32 (100%)

The results of the secondary analysis are presented in Table 9. The raw proportion of sales for both testimonial-based and evidence-based marketing was 0.2% (2 purchases in each arm). The pre-specified full model finds an effect of 0.03 percentage points that is not statistically significant (p-value = 0.915). *We find no difference in sales rates between the two marketing approaches.*

Table 9 also presents the results of the exploratory analysis. Over seven months, there were 43 EOIs across the two arms (18 in the testimonial arm and 25 in the evidence arm) in the analysis sample – i.e., 11 more EOIs between the three-month point at which the primary outcome was collected and this seven-month point. We find no difference in the rate at which schools expressed interest in EFA over this longer period (p-value = 0.376).

Table 9. Secondary and exploratory analysis results

Secondary analysis: Sales (n = 1,998)			
	Unadjusted mean (95% CI)	Log odds (95% CI)	Treatment effect converted to percentage points (95% CI)
Evidence *n=999*	0.2% (0.0, 0.5%)	0.12 (–2.04, 2.28)	0.03 (–0.17, 1.72)
Testimonial *n=999*	0.2% (0.0, 0.5%)		
Note: p-value for result is p=0.915.			
Exploratory analysis: Expressions of interest (EOIs) (7 months) (n = 1,994)			
	Unadjusted mean (95% CI)	Log odds (95% CI)	Treatment effect converted to percentage points (95% CI)
Evidence *n=998*	2.5% (1.5%, 3.5%)	0.28 (–.34, 0.90)	0.57 (–0.51, 2.52)
Testimonial *n=996*	1.8% (1.0%, 2.6%)		
Note: p-value for result is p=0.376.			

Protocol deviation

A missing data strategy was not specified in the trial protocol. We included six covariates in our analytical model that we hypothesised might be associated with our outcome and we expected to improve precision (see Appendix D for full specification). Our approach to missingness of covariates was that if any covariate was missing for more than 5% of the observations, we would drop the covariate from the analysis. For variables with missingness <5% of cases, we replaced the missing value with an extra category flagging the missingness. This approach is consistent with the EEF's 2018

statistical analysis guidance.[3] We did not have any covariate missing for more than 5% of observations, so we were able to run the model specified in the protocol without dropping any covariates.

Attrition

For the primary and secondary analysis, we excluded one school from each arm (see Table 10). One EOI was recorded at the multi-academy trust level rather than at the school level. For this particular trust, there were two schools from the trust in the sample and each had been randomised into a different arm. As such, we have excluded both schools from the analytical sample. This results in a final sample size for primary and secondary analysis of 1,998 schools.

For the exploratory analysis (EOIs at seven months), we exclude a further four schools in one trust (one in testimony, three in evidence) and one EOI for the same reason as above. This results in a final sample size for the exploratory analysis of 1,994 schools.

There is no further attrition. We were able to observe all EOIs, and any school that did not express interest was defined to have a zero outcome.

No letters were returned or recorded as undelivered, so we have no reason to believe any schools did not receive the letter and were not provided the means to express interest. It is possible some schools never opened the letter. However, lack of delivery or not reading the materials would not be considered attrition and would be ignored in these intention-to-treat analyses.

Table 10. Attrition from the trial (primary outcome)

		Testimonial	Evidence	Total
Number of schools	**Randomised**	1,000	1,000	2,000
	Analysed	999	999	1,998
School attrition (from randomisation to analysis)	**Number**	1	1	2
	Percentage	0.1%	0.1%	0.1%

Participant characteristics and balance checks

When checking balance between treatment and control groups, we coarsened two covariates that were continuous in the regression analysis (percentage of FSM pupils and number of pupils) into categorical variables. We coarsened percentage of FSM pupils into 10 categories (0–10%, 11–20%, …, 91–100%) and number of pupils into 11 categories (0–200, 201–400, …, 1801–2000, over 2000). For each covariate, we also included a category for missing values, which allowed us to examine whether treatment and groups were balanced on observable values and missingness together. We observed balance across both arms for all categories. Due to the size of the table, we present the detail on all categories in Appendix D.

Implementation

Conditions for success

It is necessary that schools are open for the success of this intervention in order to ensure that there are no barriers to delivery being experienced. The mailers were sent two weeks after schools had reopened after a nationwide closure of

[3] Accessed at: https://educationendowmentfoundation.org.uk/public/files/Evaluation/Writing_a_Protocol_or_SAP/EEF_statistical_analysis_guidance_2018.pdf

several months due to COVID-19. However, it is worth noting schools were operating under unusual conditions during Covid, which may have impacted the trial.

Fidelity

The intervention was delivered by our mailer partner and we did not observe any instances of deviation from our guidelines. There were no recorded incidents of the letters being returned or otherwise recorded undeliverable. In both treatment groups, a BIT staff member postal mailing address was included so we were able to confirm letters were delivered correctly and in the expected timeline. We were unable to observe whether letters were actually opened or read.

Perceptions of content

We intentionally did not conduct interviews with school SLT members as part of this trial, as we expected it would be very difficult to recruit a sufficient number given the many competing priorities school leaders had after schools re-opened. We do have plans to interview SLT members as part of the broader scale-up evaluation and will be asking for perceptions of the mailers to enhance our understanding of how leaders engaged with the two types of information.[4] These findings will be published within the scale-up evaluation reporting.

Cost

The cost to print, assemble and post the 2,000 mailers was £9,910.80, or approximately £4.95 per mailer. On average, it cost £310 to generate each EOI (n=32) and £2,478 for each sale (n=4).

[4] The EFA scale-up study plan can be accessed at: https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/embedding-formative-assessment-re-grant

5. Conclusion

Table 11. Summary of findings

Research question	Finding
RQ1: Does testimonial or evidence-based marketing material perform better in encouraging school leaders to express interest in EFA?	● We find no statistically significant difference in response rates to testimony versus evidence-based marketing of the EFA programme in our primary model. ● The evidence-based marketing material resulted in 21 expressions of interest (EOIs) and the testimony mailers generated 11 EOIs. ● The pre-specified model finds a difference of 0.9 percentage points in EOIs between the two types of messaging (95% CI: –0.1, 3.1). ● Overall, EOI rates were 1.6%.
RQ2: Does testimonial or evidence-based marketing material perform better in encouraging school leaders to purchase EFA?	● We find no difference in the rate of sales between evidence and testimony-based marketing materials. Two sales were recorded per arm. ● The pre-specified model finds a difference of 0.03 percentage points (95% CI: –0.17, 1.72).

Interpretation

Given a lack of evidence about how to improve take-up of evidence-based approaches, this evaluation was designed to test two variations on marketing materials to learn if programme evidence or testimony was more effective in improving uptake by school leaders.

Overall, we find 1.6% of schools expressed interest in learning more about the EFA programme. We find no statistically significant difference in response rates to testimony versus evidence-based marketing of the EFA program in our primary model (11 and 21 EOIs, respectively). The pre-specified full model finds an effect of 0.9 percentage points (p-value = 0.100).

The trial specifically prioritised measuring impact on EOIs over sales, as we wanted to establish a connection to the most direct outcome first, especially as we expect that the pathway from EOI to sales could vary substantially between schools. Four schools ultimately purchased the programme, and *we do not find any difference in rate of purchase between the two arms* (2 sales per arm).

This trial was launched shortly after schools reopened after a long period of Covid-related closure, and it is difficult to know how much this unusual context affected overall response rates. Another trial launched in the same time period observed a similar overall EOI rate (~1%) and sign-up rate (0.4%) to this trial (n=1,929) (Harland et al., 2021). This suggests that the response rates were, perhaps, typical for a time with a number of atypical challenges. Senior leaders were busy with added responsibilities around pupil Covid testing and remotely supporting pupils required to quarantine, meaning it is likely their attention may have been elsewhere whilst the trial was ongoing. Unfortunately, we do not have response rate data from pre-Covid that would be a suitable comparison.

Although not a significant difference, the direction of the results from this trial indicate that evidence may have been better in encouraging interest from school leaders. Interestingly, the results from an unpublished pilot study that inspired this trial ran in the opposite direction, with more US school leaders interested in the programme after being presented with a testimony. A similar trial, commissioned after this one and run as part of the National Tutoring Programme evaluation, found no difference between evidence and testimony in encouraging take-up of a particular tutoring provider (11 and 14 EOIs, respectively) (Harland et al., 2021).

One potential explanation for the difference in the direction of the findings between the current study with the US pilot could be because of the difference in established norms about where school leaders look for guidance on programme adoption between England and the US. A recent report by the Sutton Trust found that 79% of surveyed school leaders reported using research evidence to help inform their decision making, when thinking about which programmes or approaches to use within their schools. Of the surveyed school leaders, 69% reported specifically using the EEF toolkit when making decisions about which programmes to adopt (Sutton Trust, 2021). These numbers are consistent across both primary and secondary schools, highlighting that many school leaders are aware of the merits of using evidence-

based approaches and are familiar with the EEF. In the US, a survey of a nationally representative sample of school districts found that the most commonly reported source of information about strategies to adopt in support of school improvement was recommendations from educators in other districts (Wei & Johnson, 2020). The least frequently reported source of information was from centres that are specifically set up to disseminate and rate evidence in the education sector (i.e., institutions similar to the EEF, such as the US What Works Clearinghouse).

Limitations

First, this trial tested a particular set of materials with a particular sample and found that, when compared to marketing trials used in the private sector, the sample size is much smaller than would otherwise be used (Luca & Bazerman, 2020). Given the observed response rates, a larger sample size is likely necessary for future research in this space.

Second, the sample used was atypically 'cold' for a marketing effort (i.e., not recently marketed to or actively interested in the EFA programme), but was intentionally selected in order to provide the best estimate of the impact of the mailers. The approach used also differed from SSAT's business-as-usual marketing strategy, which is typically more digital than print. It is also atypical to contact a school once and not send any additional marketing for three months, but it would have been difficult to estimate the impact of the mailer content if headteachers were recontacted shortly after receiving the mailer. Therefore, it is difficult to compare any impact of the current study with what SSAT would usually expect.

Third, this trial tested two sets of messages, which were delivered by post. The trial was not designed to draw conclusions about the effectiveness of mail marketing over other strategies and, as stated above, the response rates are drawn from a cold sample, so are not representative of more typical marketing efforts with schools that may be in the sales pipeline. Recommendations about marketing modality sit outside the scope of the research questions tested in this trial, so none have been made.

Finally, as highlighted above in the interpretation, a large number of UK school leaders are familiar with the EEF. Although both letters contained the EEF logo in the letterhead, the evidence letter specifically cited that the evidence came from an EEF study. It is possible that a slightly higher response rate to the evidence-based letter was more due to the trust in evidence coming from the EEF than in the specific results presented.

Recommendations

From the perspective of the EEF and other What Works Centres, we have demonstrated a fairly low cost and relatively easy trial to implement that both supported SSAT's goal of scaling the EFA programme (32 EOIs and 4 sales from an otherwise very cold sample) and also expands the evidence base around how to support scaling of successful education interventions. Given the dearth of causal evidence about how to support successful scaling in education, *our primary recommendation is to continue this line of experimentation with different programmes, modes of delivery, and content.* Further experimentation is crucial to supporting evidence-based programmes being adopted and achieving their aims of improving pupil outcomes. Further qualitative research may also be beneficial in better understanding the barriers and facilitators to the take-up of evidence-based programmes to inform future experiments. We also recommend that any funding in support of scaling programmes includes a requirement that there is generation of causal evidence about how to effectively take programmes to scale, or on the process of scaling itself.

Future research and publications

As addressed in the two preceding sections, this is a single study that aimed to start building evidence on a topic that has little causal evidence. Given the conflicting direction of the results between this study and the US pilot, we think there is value in further testing of evidence versus testimony in encouraging interest and take-up of evidence-based programmes, and we intend to explore this qualitatively as part of the broader scale-up evaluation. Further testing would likely benefit from a larger sample (a minimum sample of ~2500 schools would be needed to detect an effect size similar to that found in this trial) and could explore effects with a warmer sample and/or using another modality (email). Additional testing would be helpful in understanding the impact in a context that more closely resembles 'business as usual' and would allow for a comparative cost analysis. The trial design could also be repeated and include an extra arm to help explore the potential 'EEF effect' further.

As addressed in the Recommendations section, there is little to no causal evidence in this space and any number of variations on further hypotheses can be tested. Both this trial and the Harland et al. (2021) trial found response rates of ~1% to fairly light-touch approaches to programme marketing and it may be that a more hands-on approach is necessary. Future research could further explore general barriers and facilitators for the adoption of evidence-based programmes and further interventions designed to address them.

References

Carmody, D.P., and Lewis, M., 2006. Brain activation when hearing one's own and others' names. *Brain Research*, *1116*(1), pp.153–158.

Coburn, C.E., Touré, J., and Yamashita, M., 2009. Evidence, interpretation, and persuasion: Instructional decision making at the district central office. *Teachers College Record*, *111*(4), pp.1115–1161.

Harland, J., Bradley, E., Staunton, R., Lord, P., Sims, D., and Styles, B., 2021. National Tutoring Programme – Tuition Partners – Reach and engagement randomised controlled trial: Testing different types of recruitment emails. Available from: https://d2tic4wvo1iusb.cloudfront.net/documents/pages/projects/TP_RE_RCT_-_Report_Final.pdf?v=1630925173

Lasky-Fink, J., Robinson, C.D., Chang, H.N.L., and Rogers, T., 2020. Using behavioral insights to improve school administrative communications: The case of truancy notifications. *Educational Researcher*, p.0013189X211000749.

Liberman, A., 2005. How much more likely? The implications of odds ratios for probabilities. *American Journal of Evaluation*, 2*6(2),* pp. 253–266.

Luca, M., and Bazerman, M., 2020. *The power of experiments: Decision-making in a data-driven world*, The MIT Press, Cambridge, MA.

Rogers, T., Duncan, T., Wolford, T., Ternovski, J., Subramanyam, S., and Reitano, A., 2017. *A randomized experiment using absenteeism information to 'nudge' attendance. REL 2017-252*. Regional Educational Laboratory Mid-Atlantic.

Schultz, P.W., Nolan, J.M., Cialdini, R.B., Goldstein, N.J., and Griskevicius, V., 2007. The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, *18*(5), pp.429–434.

Speckesser, S., Runge, J., Foliano, F., Bursnall, M., Hudson-Sharp, N., Rolfe, H., and Anders, J., 2018. Embedding formative assessment: Evaluation report and executive summary. Available from: https://educationendowmentfoundation.org.uk/public/files/EFA_evaluation_report.pdf

Spillane, J.P., Reiser, B.J., and Reimer, T., 2002. Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, *72*(3), pp.387–431.

Sutton Trust, 2021. School funding and pupil premium 2021. Available from: https://www.suttontrust.com/our-research/pupil-premium-2021/.

Wei, T., and Johnson, E. 2020. How states and districts support evidence use in school improvement. Available from: https://ies.ed.gov/ncee/pubs/2020004/pdf/2020004.pdf.

Wood, N.T., Solomon, M.R., and English, B.G., 2005. Personalisation of online avatars: Is the messenger as important as the message? *International Journal of Internet Marketing and Advertising*, *2*(1–2), pp.143–161.

Appendices

Appendix A: Intervention materials

Evidence-based intervention materials (page 1 of 2)

Education Endowment Foundation

ssat
the schools, students and teachers network

Central House
142 Central Street
London, EC1V 8AR

020 7802 2300
info@ssatuk.co.uk
@ssat

ssatuk.co.uk

<NAME>
<ADDRESS 1>
<ADDRESS 2>
<ADDRESS 3>
<ADDRESS 4>
<POSTCODE>

EVIDENCE from a RANDOMISED CONTROLLED TRIAL shows EFA improves students' GCSE attainment

Dear <INSERT NAME>,

We invite you to learn more about Embedding Formative Assessment (EFA), a professional development programme designed to embed improved teaching practice.

Centred around monthly teacher-led workshops and peer observation sessions, EFA is a whole school approach where teachers work collaboratively to implement formative assessment strategies in their classrooms.

Evidence from an EEF funded randomised controlled trial with 140 schools demonstrated that in schools using EFA, on average:

- Pupils made the equivalent of two additional months' progress in their Attainment 8 GCSE score;
- The rate of learning was 25% greater compared to non EFA schools;
- The programme achieved these results at a cost of around £1.20 per student per year.

Find out more information by:

- Downloading additional information: **[redacted]**
- RSVPing to attend an online Open Day: **[redacted]**
- Requesting a free phone consultation: **[redacted]**

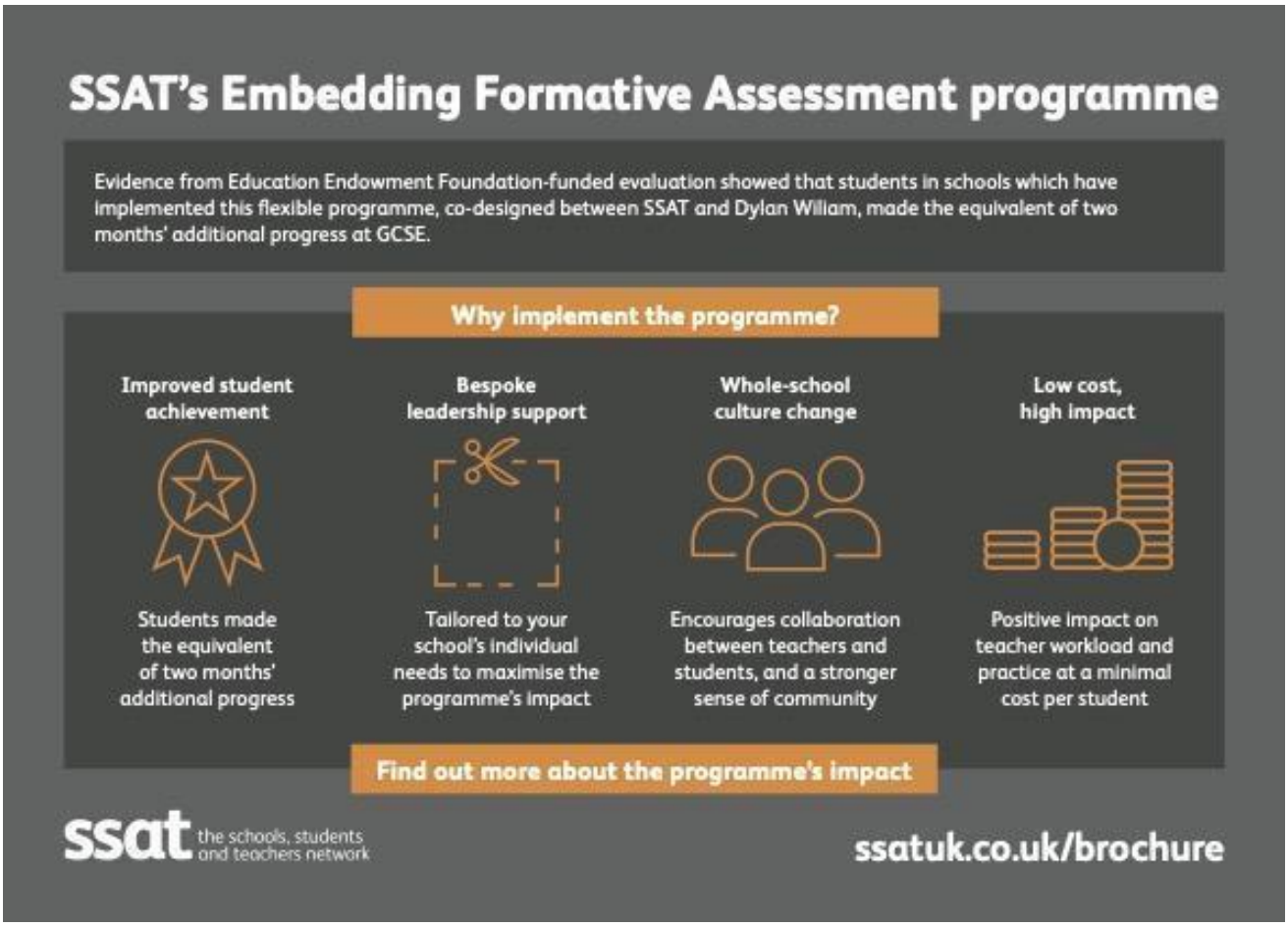
Yours sincerely,

CSettle

Corinne Settle
Senior Education Lead – Teaching and Learning

School URN: <URN>

Evidence-based intervention materials (page 2 of 2)



Testimonial intervention materials (page 1 of 2)

Education Endowment Foundation

ssat
the schools, students and teachers network

Central House
142 Central Street
London, EC1V 8AR

020 7802 2300
info@ssatuk.co.uk
@ssat

ssatuk.co.uk

<NAME>
<ADDRESS 1>
<ADDRESS 2>
<ADDRESS 3>
<ADDRESS 4>
<POSTCODE>

TESTIMONIAL from a HEADTEACHER USING EFA indicates EFA improves students' GCSE attainment

Dear <INSERT NAME>

We invite you to learn more about Embedding Formative Assessment (EFA), a professional development programme designed to embed improved teaching practice.

Centred around monthly teacher-led workshops and peer observation sessions, EFA is a whole school approach where teachers work collaboratively to implement formative assessment strategies in their classrooms.

Assistant head Annabelle Taylor reports on the success of the EFA programme at Ricards Lodge High School. After embedding EFA:

- Our Attainment 8 score was 54.4, which ranked us second out of 55 similar schools nationally;
- We achieved our best ever Progress 8 score (+0.71);
- Ofsted rated the school as outstanding, highlighting in their report observations of teachers using formative assessment in the classroom.

Find out more information by:

- Downloading additional information: **[redacted]**
- RSVPing to attend an online Open Day: **[redacted]**
- Requesting a free phone consultation: **[redacted]**

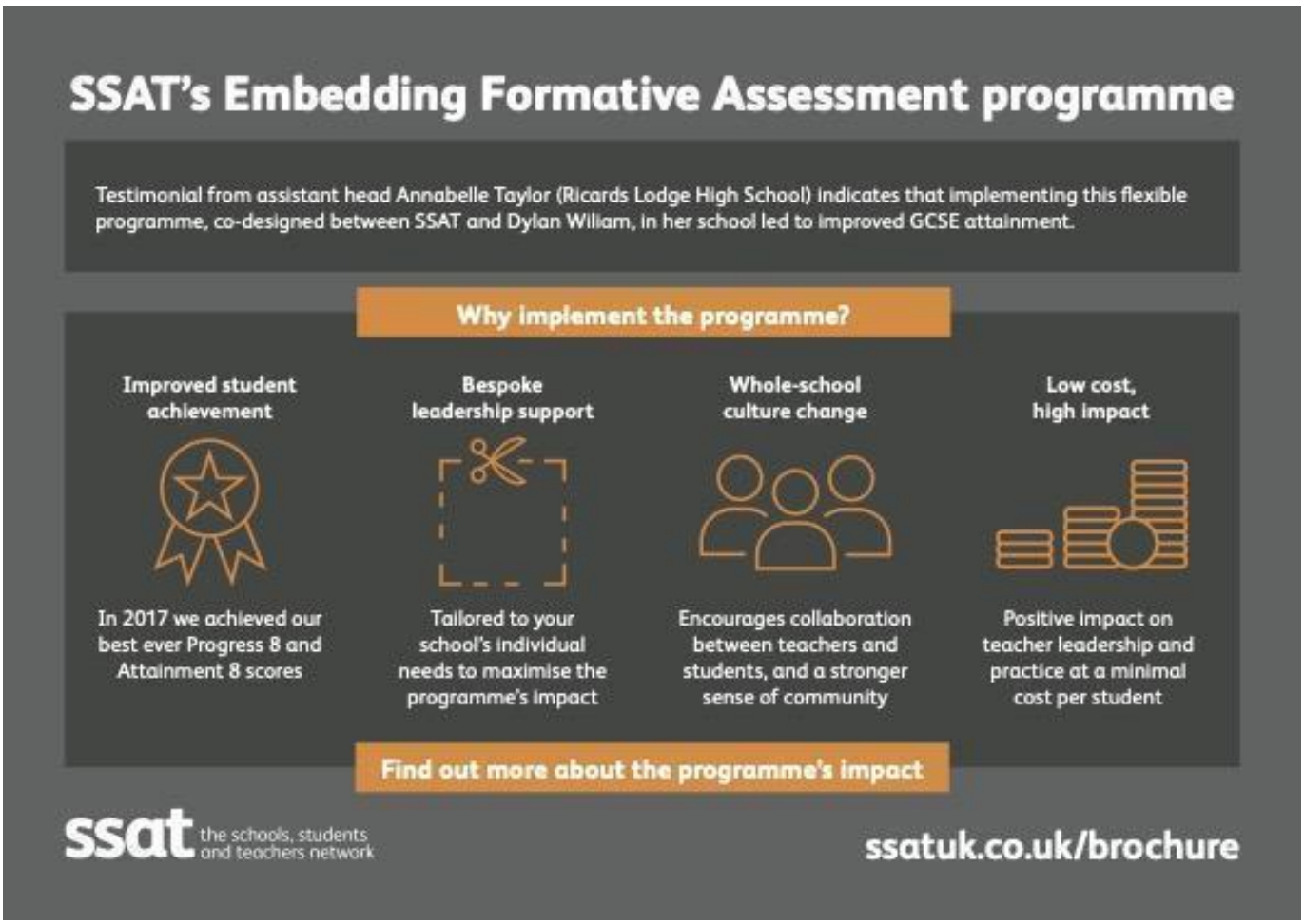
Yours sincerely,

C Settle

Corinne Settle
Senior Education Lead – Teaching and Learning

School URN: <URN >

Testimonial intervention materials (page 2 of 2)



Appendix B: Project team

The project was a collaboration between teams at the Behavioural Insights Team (BIT) and SSAT.

From BIT, the project was managed by Kimberly Bohling, with Alex Sutherland as Principal Investigator. Jo Milward led on letter design and implementation, Pujen Shrestha was the lead analyst, and Eleanor Collerton supported with reporting. Quality assurance was provided by Tim Hardy, Lev Tankelevitch, Dilhan Perera and Clément Bisserbe.

The team at SSAT also advised on letter design and content and managed data collection and sharing: Jennifer Farrell, Corinne Settle, Mike Jones, and Ellen Renton Pearce.

We also thank Professor Todd Rogers and Emma Starr, both of Harvard Kennedy School, for advising on the trial and letter design.

Appendix C: Project timeline

Dates	Activity	Staff responsible/ leading
Dec 2020	Project kick-off meeting	EEF
W/C 15th February 2021	Randomisation	BIT (Research Team)
26th Feb 2021	Finalising intervention materials	BIT (Project Team)
W/C 15th March 2021	Launch trial; mailers sent out	BIT and Mailer Partner
May 2021	Data collection (EOI)	SSAT and BIT (Research Team)
September 2021	Data collection (Purchase and EOI)	SSAT and BIT (Research Team)

Appendix D: Impact analysis technical detail

Primary outcome model specification

Outcome

The primary outcome is an EOI in the EFA programme being submitted by a school within approximately three months after the mail-out being received by the school. This is a binary indicator.

Analysis

Primary analysis is intention-to-treat (ITT). Analysis uses the following logistic regression model:

$$outcome_i \sim bernoulli(p_i)\ ;\ logit(p_i) = \alpha + \beta_T treatmentt_i + \beta_C covariates_i$$

where:

- $outcome_i$ is the outcome for EOI, binary (1 if the school expresses interest, 0 if not);
- $treatment_i$ is a binary indicator for the treatment assignment for school i (1 if the school is assigned to evidenced-based marketing treatment; 0 if not); and
- $covariates_i$ is a set of all covariates of interest for school i:
  - proportion of FSM eligible pupils,
  - region (as a dummy variable, 'London' as reference group),
  - whether the school is part of a multi-academy trust (1 if the school is part of a MAT; 0 if not),
  - school size (number of pupils),
  - Ofsted rating (as a dummy variable, 'outstanding' as reference group), and
  - opportunity area (1 if the school is located in an opportunity area; 0 if not).

The analysis was first conducted by the primary analyst on the trial. The primary analyst also produced a blinded dataset, which was analysed independently by a second analyst who was not working on the trial. Both analysts followed the trial protocol and produced the same results.

Table 12 provides the results of the primary and exploratory analyses, which are described in the body of the report. Some categories (e.g., East of England region in column 1) have very large coefficients and standard errors because there is little variation in the outcome for those categories.

Table 12. Primary and exploratory analyses (EOIs)

	Dependent variable: **Expression of interest (EOI)** Logistic regression coefficients (standard errors in parentheses)		
	(1) Primary – EOIs after 3 months, main specification	(2) Primary – EOIs after 3 months, no covariates	(3) Exploratory outcome – EOIs after 7 months
Evidence-based marketing	**0.624**	**0.657**	**0.280**
	(0.380)	(0.375)	(0.316)
Percentage of FSM-eligible pupils	0.026		0.025
	(0.016)		(0.015)
PercentageFSM (mean imputed) x PercentageFSM missing	–0.886		–0.668
	(514.072)		(47.207)
Region (reference category is London)			
East Midlands	–0.247		–0.734
	(0.853)		(0.824)
East of England	–16.262		–0.618
	(1,107.988)		(0.722)

North East	–0.556		–0.963
	(1.103)		(1.079)
North West	–0.450		–0.402
	(0.740)		(0.604)
South East	0.246		0.223
	(0.636)		(0.539)
South West	0.542		0.116
	(0.667)		(0.619)
West Midlands	0.315		–0.094
	(0.643)		(0.604)
Yorkshire and the Humber	0.453		0.012
	(0.683)		(0.620)
Part of a multi-academy trust	–0.229		–0.340
	(0.381)		(0.329)
Number of pupils	0.0002		0.001*
	(0.0005)		(0.0004)
Number of pupils (mean imputed) x Number of pupils missing	0.015		0.011
	(9.100)		(0.836)
Ofsted rating (reference category is 'outstanding')			
Good	–0.340		0.014
	(0.460)		(0.418)
Requires improvement	–0.911		–0.265
	(0.728)		(0.606)
Inadequate	–0.128		0.460
	(0.747)		(0.666)
Missing	–0.191		0.093
	(1.215)		(1.212)
In an Opportunity Area	–16.475		0.278
	(2,020.573)		(0.782)
Constant	–4.607**	–4.498**	–4.853**
	(0.920)	(0.303)	(0.818)
Observations	1,998	1,998	1,994
Note:	* p<0.05; ** p<0.01		

Secondary outcome model specification

Secondary outcome analysis followed the same model specification used for the primary outcome.

The secondary outcome is purchase of the EFA programme within approximately seven months of the school receiving the mail-out. This is a binary indicator. Table 13 provides the results of the secondary analysis, which is described in the body of the report. Some categories (e.g., regions) have very large coefficients and standard errors because there is little variation in the outcome for those categories.

Table 13. Secondary analysis (sales): logistic regression coefficients (standard errors in parentheses)

	Dependent variable: Sale Logistic regression coefficients (standard errors in parentheses)
	(1) Main specification
Evidence-based marketing	**0.118**
	(1.103)
PercentageFSM	0.051
	(0.035)
PercentageFSM (mean imputed) x PercentageFSM missing	0.032
	(6,008.541)
Region (reference category is London)	
East Midlands	–34.513
	(12,493.380)
East of England	–16.507
	(4,445.267)
North East	–18.947
	(18,411.010)
North West	–36.003
	(10,442.860)
South East	1.045
	(1.796)
South West	1.356
	(1.812)
West Midlands	–36.062
	(11,612.560)
Yorkshire and the Humber	–38.007
	(12,440.290)
Part of a multi-academy trust	–1.124
	(1.301)
Number of pupils	–0.002
	(0.002)
Number of pupils (mean imputed) x Number of pupils missing	–0.019
	(107.809)
Ofsted rating (reference category is ‘outstanding’)	
Good	–0.626
	(1.385)
Requires improvement	–34.686
	(8,928.601)
Inadequate	–19.358
	(14,066.880)
Missing	–20.528
	(17,429.680)
In an Opportunity Area	21.351
	(4,445.267)

Constant	–4.383
	(2.480)
Observations	1,998
Note:	* p<0.05; ** p<0.01

Table 14. Participating school characteristics

		Testimony (n=999)	Evidence (n=999)
Ofsted Rating	Outstanding	192 (19.2%)	195 (19.5%)
	Good	558 (55.9%)	516 (51.7%)
	Requires improvement	146 (14.6%)	167 (16.7%)
	Inadequate	61 (6.1%)	80 (8.0%)
	Missing	42 (4.2%)	41 (4.1%)
Multi-Academy Trust (MAT)	Non-MAT	392 (39.2%)	399 (39.9%)
	MAT	607 (60.8%)	600 (60.1%)
Opportunity Area (OA)	Non-OA	967 (96.8%)	961 (96.2%)
	OA	32 (3.2%)	38 (3.8%)
Free School Meals (FSM)	0–10%	283 (28.3%)	273 (27.3%)
	11–20%	371 (37.1%)	342 (34.2%)
	21–30%	189 (18.9%)	198 (19.8%)
	31–40%	83 (8.3%)	98 (9.8%)
	41–50%	28 (2.8%)	39 (3.9%)
	51–60%	3 (0.3%)	4 (0.4%)
	61–70%	0 (0%)	2 (0.2%)
	71–80%	0 (0%)	0 (0%)
	81–90%	0 (0%)	0 (0%)
	91–100%	0 (0%)	1 (0.1%)
	Missing	42 (4.2%)	42 (4.2%)
Number of pupils	<200	23 (2.3%)	18 (1.8%)
	201–400	54 (5.4%)	67 (6.7%)
	401–600	103 (10.3%)	97 (9.7%)
	601–800	136 (13.6%)	158 (15.8%)
	801–1000	171 (17.1%)	199 (19.9%)
	1001–1200	193 (19.3%)	147 (14.7%)
	1201–1400	133 (13.3%)	119 (11.9%)
	1401–1600	68 (6.8%)	84 (8.4%)
	1601–1800	43 (4.3%)	38 (3.8%)
	1801–2000	19 (1.9%)	19 (1.9%)
	>2000	16 (1.6%)	12 (1.2%)
	Missing	40 (4.0%)	41 (4.1%)

Converting log-odds to percentage point change

For clear interpretation we have presented the performance of treatment arms in percentages. This is possible as log odds are directly interpretable as a percentage[5]. We have included a worked example for context.

1. Calculate the proportion of the EOIs occurring in the 'testimonial' group.

2. Calculate the logit of the proportion of EOIs in the 'testimonial' group using the logistic function log(p / (1 – p)).

3. Calculate the 'evidence' group estimated proportion by taking the inverse-logit of the treatment effect coefficient + the logit of the proportion of EOIs in the 'testimonial' group.

4. Calculate the percentage point effect by subtracting the true ‘testimonial’ proportion from the estimated ‘evidence’ proportion.

Using the steps above, the primary analysis can be worked as:

1. The proportion of schools in the ‘testimonial’ group with an EOI after 3 months was 0.011 (i.e., 1.1%).
2. Applying the logistic function to this gives log(0.011/(1–0.011)) = –4.498.
3. The estimated treatment coefficient from column 1 of Table 12 is 0.624, which when added to –4.498 gives –3.874. The ‘evidence’ group estimated proportion is invlogit(–4.498 + 0.624) = exp(–3.874)/(1 + exp(–3.874)) = 0.020 (i.e., 2.0%).
4. The estimated percentage point effect is thus 0.020 – 0.011 = 0.009 (or 0.9pp).

Example code:

```
### Example for converting odds ratios to percentages

library(tidyverse)

library(LaplacesDemon)

# Get the proportion of the outcome occurring in the 'treatment 0' group

testimonal_rate <- filter(data, treatment == 0) %>%  summarise(proportion_outcome = mean(eoi_f == "Yes")) %>%

as.numeric()

testimonal_rate_logit <- logit(testimonal_rate)

# Get the logit treatment effect coefficient

evidence_logit <- primary_model$coefficients["evidence based"]

# Calculate the 'treatment 1' group estimated outcome proportions evidence_est_rate <- invlogit(testimonal_rate_logit
+ evidence_logit)

# Calculate the percentage point effect

(percent_effect <- (evidence_est_rate - testimonal_rate) * 100)
```

[5] See Liberman, A.M., 2005. for a longer discussion.

Appendix E: Ethical review

All BIT trials need to have been through BIT's internal research ethics process. This trial was assessed as being Low Risk with regards to the dimensions of our ethical review process. We were using standard research methods commonly applied in education and participants were all adults with no known or expected vulnerabilities. The trial was being delivered in the context of the EFA's scale-up evaluation using materials that are inline with standard SSAT marketing. There is minimal use of personal data, which is limited to the use of head teachers names in addressing the letters. All analysis was conducted on school-level (non-personal) data. BIT has extensive experience both in conducting research in the education sector, as well as testing variations in communications. This trial did not alert participating schools to the fact that they were taking part in a trial. It is not unusual for marketing teams to test different messaging approaches and track response rates to determine which is more effective. All respondents who expressed interest in the programme were presented with SSAT's data privacy notice which says that the data can be used for the purposes of evaluating service delivery and marketing. All analysis was conducted at the school-level; not at the individual level.

Before launching the trial, BIT registered it at the Open Science Framework (osf.io). The trial registry will be updated with outcomes once this report is published by the EEF.

Appendix F: Data protection

- BIT acted as SSAT's data processor for the purposes of this evaluation. SSAT provided BIT with school-level marketing data (not personal data) for the purposes of randomisation and analysis. BIT combined this data with publicly available personal data to appropriately address the letters.

- Any person who submitted an expression of interest was presented with SSAT's standard privacy notice as concerns collection and processing of personal data. Their legitimate interests include evaluating service delivery and marketing to potential customers, which is in alignment with the purposes of this research. However, SSAT did not provide any personal data to BIT.

- BIT engaged a sub-processor (Rapidity Printing) for the purposes of printing, assembling, and posting the letters. This was part of the data sharing agreement between BIT and SSAT.

- All personal data collected as part of the study has been treated with the strictest confidence by BIT and processed only in accordance with the requirements of the GDPR and the Data Protection Act 2018. BIT has not used any personal data in any report arising from this project. BIT collected and processed personal data solely for the purposes of proper delivery of the intervention.

- We needed to employ personal data (the headteacher's name of all selected schools) for the implementation of our intervention.

- BIT did not process any special data in this project.

- No data from this project will be transferred to the EEF archive.

Appendix G: Randomisation and analysis code

Power calculations code:

```
## 1% baserate
## we know the baseline rate and n:
baselineRate <- 0.01
n <- 1000

## first get h
pwr.2p.test(power=0.8,n=n)

## keep hold of it
pwr.2p.test(power=0.8,n=n)$h -> h ## h = 0.1252923

## now we calculate p
phi.b <- 2*asin(sqrt(baselineRate)) ## 0.2003348
phi.t <- phi.b + h ## 0.3256272

## treatment rate
(treatmentRate <- sin(0.5*phi.t)^2) ## 0.02627486

## MDES
treatmentRate - baselineRate ## 0.01627486


## 5% baserate
## we know the baseline rate and n:
baselineRate <- 0.05
n <- 1000

## first get h
pwr.2p.test(power=0.8,n=n)

## keep hold of it
pwr.2p.test(power=0.8,n=n)$h -> h ## h = 0.1252923

## now we calculate p
phi.b <- 2*asin(sqrt(baselineRate)) ## 0.4510268
phi.t <- phi.b + h ## 0.5763191

## treatment rate
(treatmentRate <- sin(0.5*phi.t)^2) ## 0.08076291

## MDES
treatmentRate - baselineRate ## 0.03076291

## 10% baserate
## we know the baseline rate and n:
baselineRate <- 0.1
n <- 1000

## first get h
pwr.2p.test(power=0.8,n=n)
```

```
## keep hold of it
pwr.2p.test(power=0.8,n=n)$h -> h ## h = 0.1252923

## now we calculate p
phi.b <- 2*asin(sqrt(baselineRate)) ## 0.6435011
phi.t <- phi.b + h ## 0.7687934

## treatment rate
(treatmentRate <- sin(0.5*phi.t)^2) ## 0.140625

## MDES
treatmentRate - baselineRate ## 0.04062496
```

Randomisation code:

```
---
title: "EFA Mailer Randomisation"
author: "Pujen Shrestha"
---

rm(list = ls())

library(data.table)
library(lfe)
library(tidyverse)
library(TREX)
library(scales)
library(lemon)
library(randomizr)

# Set up SSAT selected school----------------------------------------------------------------

## set seed
set.seed(01032021)

## import dataset, SSAT selected schools
data_selected_schools <- read.csv("/Volumes/GoogleDrive/.shortcut-targets-by-
id/1gP_9rMQyGFpKsDgjB9lKyO6GsB_VXRMw/P0EE4 - EFA Mailer Trial/Pujen - R/efaSchool.csv")

# clean data into booleans
data_selected_schools <- data_selected_schools %>%
  transmute(URN = URN,
            criterion6 = if_else(EFA.Open.Day.booking..Criterion.6. == "Criterion 6", 1, 0),
            criterion7 = if_else(Download..Criterion.7. == "Criterion 7", 1, 0),
            criterion8 = if_else(Thunderhead..Bullet.2. == "Criterion Bullet 2", 1, 0)
  )

nrow(data_selected_schools)

# Inclusion/exclusion SSAT selected schools ----------------------------------------

## Exclusion criteria
# 1. Taken part in the EFF EFA research
# 2. EFA Mentor and Ambassador schools
```

```
# 3. Schools that are or have undertaken the EFA support package
# 4. Schools that have purchased the EFA toolkit
# 5. Schools that expressed an interest (EOI) in EFA or requested a EFA consultation
# 6. Schools that have had staff members booked to or attended EFA open days
# 7. Schools that have had a staff member download content from SSAT website related to EFA

# 1-5 filter conducted by SSAT, we only need to filter out 6 and 7

data_selected_schools <- data_selected_schools %>%
  filter(criterion6 == 0 & criterion7 == 0)

nrow(data_selected_schools)

# import edubase dataset ----------------------------------------------------

## import edubase dataset
data_covars <- read.csv("/Volumes/GoogleDrive/.shortcut-targets-by-
id/1gP_9rMQyGFpKsDgjB9lKyO6GsB_VXRMw/P0EE4 - EFA Mailer Trial/Pujen - R/edubasealldata20210107.csv")

# data_covars$URN

# select variables for randomisation and mailmerger
data_covars <- data_covars %>%
  select(URN, EstablishmentName, PhaseOfEducation..code., PhaseOfEducation..name.,
         Street, Locality, Address3, Town, County..name., Postcode, HeadFirstName,
         HeadLastName) %>%
  mutate(headTeacherName = paste(HeadFirstName, HeadLastName, " "))

## merge selected schools and covars
data_merged <- left_join(data_selected_schools, data_covars) %>%
  arrange(URN)
nrow(data_merged)
sum(is.na(data_merged$EstablishmentName))

# randomisation -------------------------------------------------------------
# two arms, 1000 schools per arm

## set the seed first
set.seed(20210203L)

# randomisation into two different datasets for easy of mailer merger

data <- data_merged

# School Unique Reference Number (URN)
urn <- data$URN

assigned_a <- sample(urn, 1000, replace = FALSE)

unassigned <- urn[!(urn %in% assigned_a)]

assigned_b <- sample(unassigned, 1000, replace = FALSE)

sum(assigned_a %in% assigned_b)
```

```
assigned_treatment <- rbind(data.frame(URN = assigned_a, treatment = rep("A", times = 1000)),
                            data.frame(URN = assigned_b, treatment = rep("B", times = 1000)))

nrow(assigned_treatment)

assigned_treatment <- left_join(assigned_treatment, data_merged)

# write dataset for mail merger ------------------------------------------

path_out = '/Volumes/GoogleDrive/.shortcut-targets-by-id/1gP_9rMQyGFpKsDgjB9lKyO6GsB_VXRMw/P0EE4 - EFA Mailer Trial/Pujen - R/'

#Commented out to protect writeover

# # treatment A - testimonial data
# EFA_A_treatment <- left_join(data.frame(URN = assigned_a), data_merged) %>%
#   select(URN, EstablishmentName, headTeacherName, Street, Address3, Town, Postcode, treatment)

EFA_A_treatment <- left_join(data.frame(URN = assigned_a), assigned_treatment) %>%
  select(URN, EstablishmentName, headTeacherName, Street, Address3, Town, Postcode, treatment)
#
# fileName = paste(path_out, 'EFA_A_treatment.csv',sep = '')
# write_csv(assigned_a_mail, fileName)

# # treatment B - evidence based data
# EFA_B_treatment <- left_join(data.frame(URN = assigned_b), data_merged) %>%
#   select(URN, EstablishmentName, headTeacherName, Street, Address3, Town, Postcode, treatment)

EFA_B_treatment <- left_join(data.frame(URN = assigned_b), assigned_treatment) %>%
  select(URN, EstablishmentName, headTeacherName, Street, Address3, Town, Postcode, treatment)
#
# fileName = paste(path_out, 'EFA_B_treatment.csv',sep = '')
# write_csv(assigned_a_mail, fileName)


Analysis code:
rm(list=ls())
library(summarytools)
library(tidyverse)
library(knitr)
library(lmtest)
library(sandwich)
library(stargazer)
library(LaplacesDemon)
library(table1)
library(TREX)

data <- read_csv("G:/My Drive/BIT project data/P0EE4 - EFA Mailer Trial/Pujen - R/Datasets/efa_cleaned_TH.csv")
names(data)
table(data$treatment)
data$treatment <- relevel(as.factor(data$treatment), ref="testimonial")
table(data$treatment)

#1. "Cleaning"
sum(is.na(data$LA))
```

```
sum(is.na(data$PercentageFSM)) #84
sum(is.na(data$GOR))
sum(is.na(data$ofsted)) #83
sum(is.na(data$district))
sum(is.na(data$MAT))
sum(is.na(data$n_pupil)) #81
sum(is.na(data$oa))

#Replace missing values with mean (fortunately none are integers)
data$PercentageFSM_raw <- data$PercentageFSM
data$PercentageFSM_missing <- (is.na(data$PercentageFSM_raw))
mean(data$PercentageFSM, na.rm = TRUE)
data$PercentageFSM <- ifelse(is.na(data$PercentageFSM),
                        mean(data$PercentageFSM, na.rm = TRUE),
                        data$PercentageFSM)

data$n_pupil_raw <- data$n_pupil
data$n_pupil_missing <- (is.na(data$n_pupil_raw))
mean(data$n_pupil, na.rm = TRUE)
data$n_pupil <- ifelse(is.na(data$n_pupil), mean(data$n_pupil, na.rm = TRUE), data$n_pupil)

data$ofsted <- ifelse(is.na(data$ofsted), "Missing", data$ofsted)

#2. Balance checks
data <- data %>%
  mutate(FSM_cat = case_when(PercentageFSM_raw >= 0 & PercentageFSM_raw <= 10 ~ "0-10%",
                        PercentageFSM_raw > 10 & PercentageFSM_raw <= 20 ~ "11-20%",
                        PercentageFSM_raw > 20 & PercentageFSM_raw <= 30 ~ "21-30%",
                        PercentageFSM_raw > 30 & PercentageFSM_raw <= 40 ~ "31-40%",
                        PercentageFSM_raw > 40 & PercentageFSM_raw <= 50 ~ "41-50%",
                        PercentageFSM_raw > 50 & PercentageFSM_raw <= 60 ~ "51-60%",
                        PercentageFSM_raw > 60 & PercentageFSM_raw <= 70 ~ "61-70%",
                        PercentageFSM_raw > 70 & PercentageFSM_raw <= 80 ~ "71-80%",
                        PercentageFSM_raw > 80 & PercentageFSM_raw <= 90 ~ "81-90%",
                        PercentageFSM_raw > 90 & PercentageFSM_raw <= 100 ~ "91-100%"))
data$FSM_cat <- ifelse(is.na(data$FSM_cat), "Missing", data$FSM_cat)

data <- data %>%
  mutate(n_pupil_cat = case_when(n_pupil_raw >= 0 & n_pupil_raw <= 200 ~ "<200",
                        n_pupil_raw > 200 & n_pupil_raw <= 400 ~ "201-400",
                        n_pupil_raw > 400 & n_pupil_raw <= 600 ~ "401-600",
                        n_pupil_raw > 600 & n_pupil_raw <= 800 ~ "601-800",
                        n_pupil_raw > 800 & n_pupil_raw <= 1000 ~ "801-1000",
                        n_pupil_raw > 1000 & n_pupil_raw <= 1200 ~ "1001-1200",
                        n_pupil_raw > 1200 & n_pupil_raw <= 1400 ~ "1201-1400",
                        n_pupil_raw > 1400 & n_pupil_raw <= 1600 ~ "1401-1600",
                        n_pupil_raw > 1600 & n_pupil_raw <= 1800 ~ "1601-1800",
                        n_pupil_raw > 1800 & n_pupil_raw <= 2000 ~ "1801-2000",
                        n_pupil_raw > 2000 ~ ">2000"))
data$n_pupil_cat <- ifelse(is.na(data$n_pupil_cat), "Missing", data$n_pupil_cat)

#Chi-square tests
table1(~ factor(ofsted) | factor(treatment), data = data)
table1(~ factor(MAT) | factor(treatment), data = data)
table1(~ factor(oa) | factor(treatment), data = data)
```

```
table1(~ factor(FSM_cat) | factor(treatment), data = data)
table1(~ factor(n_pupil_cat) | factor(treatment), data = data)

chisq.test(table(data$treatment, data$ofsted))
chisq.test(table(data$treatment, data$MAT))
chisq.test(table(data$treatment, data$oa))
chisq.test(table(data$treatment, data$FSM_cat))
chisq.test(table(data$treatment, data$n_pupil_cat))
chisq.test(table(data$treatment, data$GOR))

#Reference categories
data$GOR <- relevel(as.factor(data$GOR), ref="London")
data$ofsted <- relevel(as.factor(data$ofsted), ref="Outstanding")

#3. Regressions ----------------------------------------------------------
#a. Primary outcome - EOI within 3 months of launch
##Main spec
table(data$eoi, data$treatment) #1.1% vs. 2.1%
p <- 11/999
100*(p - (1.96*sqrt(p*(1-p)/1000)))
100*(p + (1.96*sqrt(p*(1-p)/1000)))
p <- 21/999
100*(p - (1.96*sqrt(p*(1-p)/1000)))
100*(p + (1.96*sqrt(p*(1-p)/1000)))
rm(p)

primary_model <- glm(I(eoi=="Yes") ~ treatment + PercentageFSM + PercentageFSM:PercentageFSM_missing +
GOR + MAT +
n_pupil + n_pupil:n_pupil_missing + ofsted + oa, family = "binomial", data = data)
summary(primary_model)

primary_model_stargazer <- primary_model

##Naive reg (no covars)
secondary_model <- glm(I(eoi=="Yes") ~ treatment, family = "binomial", data = data)
summary(secondary_model)

#Interpret results of main spec
control_rate <- sum(data$eoi=="Yes" & data$treatment=="testimonial")/sum(data$treatment=="testimonial")
control_rate_logit <- logit(control_rate)

##Get the logit treatment effect coefficients
primary_model$coefficients["treatmentevidence based"]
evidence_logit <- primary_model$coefficients["treatmentevidence based"]

##Calculate the evidence estimated
evidence_est_rate <- invlogit(control_rate_logit + evidence_logit)

##Calculate the percentage effect
percent_effect <- (evidence_est_rate - control_rate) * 100

##Make BitBar
##Calculate CI for barplot primary regression
evidence_logit_upper <- primary_model$coefficients["treatmentevidence based"] +
(1.96*summary(primary_model)$coefficients[2, 2])
```

```
evidence_logit_lower <- primary_model$coefficients["treatmentevidence based"] -
(1.96*summary(primary_model)$coefficients[2, 2])
as.numeric(evidence_logit_upper)

evidence_est_rate_upper <- invlogit(control_rate_logit + evidence_logit_upper)
evidence_est_rate_lower <- invlogit(control_rate_logit + evidence_logit_lower)

percent_effect_upper <- (evidence_est_rate_upper - control_rate) * 100
percent_effect_lower <- (evidence_est_rate_lower - control_rate) * 100

BITBarplot(
  c(round(control_rate*100, digits = 1), round(evidence_est_rate*100, digits = 1)),                    ## bar heights
  CIMin = c(round(control_rate*100, digits = 1), evidence_est_rate_lower*100),
  CIMax = c(round(control_rate*100, digits = 1), evidence_est_rate_upper*100),
  barXLabels = c("Testimonial", "Evidence"),
  ylab="EOI rate (%)",          ## Y axis label
  yline = 1,                    ## move the label out by 1 line so it doesn't overlap
  extramar = 1,
  p=c(1,1),              ## p-values
  subtitle = "",
  controlGroup = TRUE,
  main="",
  col.main = BITBlack,
  barCol = BITGray,
  line.main = -1,
  N = 1998
)

#b. Secondary outcome - sale
##Main spec
table(data$sale, data$treatment)
p <- 2/999
rm(p)

primary_model <- glm(I(sale=="Yes") ~ treatment + PercentageFSM + PercentageFSM:PercentageFSM_missing +
GOR + MAT +
                     n_pupil + n_pupil:n_pupil_missing + ofsted + oa, family = "binomial", data = data)
summary(primary_model)
stargazer(primary_model, type = "html", title = "Sales regressions",
        out = "G:/My Drive/BIT project data/P0EE4 - EFA Mailer Trial/Pujen - R/Code/sales.html")

#Interpret results of main spec
control_rate <- sum(data$sale=="Yes" & data$treatment=="testimonial")/sum(data$treatment=="testimonial")
control_rate_logit <- logit(control_rate)

##Get the logit treatment effect coefficients
primary_model$coefficients["treatmentevidence based"]
evidence_logit <- primary_model$coefficients["treatmentevidence based"]

##Calculate the evidence estimated
evidence_est_rate <- invlogit(control_rate_logit + evidence_logit)

##Calculate the percentage effect
percent_effect <- (evidence_est_rate - control_rate) * 100
as.numeric(percent_effect)
```

```
##Calculate CI for barplot primary regression
evidence_logit_upper <- primary_model$coefficients["treatmentevidence based"] +
(1.96*summary(primary_model)$coefficients[2, 2])
evidence_logit_lower <- primary_model$coefficients["treatmentevidence based"] -
(1.96*summary(primary_model)$coefficients[2, 2])
as.numeric(evidence_logit_upper)
as.numeric(evidence_logit_lower)

evidence_est_rate_upper <- invlogit(control_rate_logit + evidence_logit_upper)
evidence_est_rate_lower <- invlogit(control_rate_logit + evidence_logit_lower)

percent_effect_upper <- (evidence_est_rate_upper - control_rate) * 100
percent_effect_lower <- (evidence_est_rate_lower - control_rate) * 100
as.numeric(percent_effect_upper)
as.numeric(percent_effect_lower)

#c. Exploratory outcome - EOI within 7 months of launch
##Main spec
table(data$eoi_7mo, data$treatment)
p <- 18/998
100*(p - (1.96*sqrt(p*(1-p)/1000)))
100*(p + (1.96*sqrt(p*(1-p)/1000)))
p <- 25/996
100*(p - (1.96*sqrt(p*(1-p)/1000)))
100*(p + (1.96*sqrt(p*(1-p)/1000)))
rm(p)

exploratory_model <- glm(I(eoi_7mo=="Yes") ~ treatment + PercentageFSM +
PercentageFSM:PercentageFSM_missing + GOR + MAT +
                  n_pupil + n_pupil:n_pupil_missing + ofsted + oa, family = "binomial", data = data)
        #note: eoi_7mo is missing for the 4 schools which need to be excluded
summary(exploratory_model)
stargazer(primary_model_stargazer, secondary_model, exploratory_model,
      type = "html", title = "EOIs regressions",
      out = "G:/My Drive/BIT project data/P0EE4 - EFA Mailer Trial/Pujen - R/Code/EOIs.html")

#Interpret results of main spec
control_rate <- sum(data$eoi_7mo=="Yes" & data$treatment=="testimonial" &
!is.na(data$eoi_7mo))/sum(data$treatment=="testimonial" & !is.na(data$eoi_7mo))
control_rate_logit <- logit(control_rate)

##Get the logit treatment effect coefficients
exploratory_model$coefficients["treatmentevidence based"]
evidence_logit <- exploratory_model$coefficients["treatmentevidence based"]

##Calculate the evidence estimated
evidence_est_rate <- invlogit(control_rate_logit + evidence_logit)

##Calculate the percentage effect
percent_effect <- (evidence_est_rate - control_rate) * 100
as.numeric(percent_effect)

##Calculate CI for barplot primary regression
```

```
evidence_logit_upper <- exploratory_model$coefficients["treatmentevidence based"] +
(1.96*summary(exploratory_model)$coefficients[2, 2])
evidence_logit_lower <- exploratory_model$coefficients["treatmentevidence based"] -
(1.96*summary(exploratory_model)$coefficients[2, 2])
as.numeric(evidence_logit_upper)
as.numeric(evidence_logit_lower)

evidence_est_rate_upper <- invlogit(control_rate_logit + evidence_logit_upper)
evidence_est_rate_lower <- invlogit(control_rate_logit + evidence_logit_lower)

percent_effect_upper <- (evidence_est_rate_upper - control_rate) * 100
percent_effect_lower <- (evidence_est_rate_lower - control_rate) * 100
as.numeric(percent_effect_upper)
as.numeric(percent_effect_lower)
```

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