Education Endowment Foundation

Embedding Formative Assessment scale-up evaluation

Interim report

May 2023

Patrick Taylor, Rizwaan Malik, Pujen Shrestha, Neus Torres

THE BEHAVIOURAL INSIGHTS TEAM

The Education Endowment Foundation is an independent charity dedicated to breaking the link between family income and education achievement. We support schools, nurseries and colleges to improve teaching and learning for 2 – 19-year-olds through better use of evidence.

We do this by:

- **Summarising evidence.** Reviewing the best available evidence on teaching and learning and presenting in an accessible way.
- **Finding new evidence.** Funding independent evaluations of programmes and approaches that aim to raise the attainment of children and young people from socio-economically disadvantaged backgrounds.
- **Putting evidence to use.** Supporting education practitioners, as well as policymakers and other organisations, to use evidence in ways that improve teaching and learning.

We were set-up in 2011 by the Sutton Trust partnership with Impetus with a founding £125m grant from the Department for Education. In 2022, we were re-endowed with an additional £137m, allowing us to continue our work until at least 2032.

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

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

info@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents

About the evaluator

The project was independently evaluated by a team from the Behavioural Insights Team: Patrick Taylor; Rizwaan Malik; Pujen Shrestha; and Neus Torres.

The principal investigator was Patrick Taylor.

Contact details:

Behavioural Insights Team

4 Matthew Parker Street,

London,

SW1H 9NP

Email: patrick.taylor@bi.team

Executive summary

The project

This evaluation is a collaboration between the Schools, Students, and Teachers (SSAT) Network, and the Behavioural Insights Team (BIT). It focuses on the scale-up of a professional development programme for teachers called Embedding Formative Assessment (EFA).

The subject of the evaluation is the scaling up of EFA. EFA is a professional development programme that aims to improve pupil outcomes by embedding the use of formative assessment strategies across a school. In the programme, schools receive initial training, 2 years of ongoing support from a mentor, and detailed resource packs to run structured monthly workshops, known as Teacher Learning Communities (TLCs). The programme focuses on enhancing understanding of formative assessment strategies, reflecting on formative assessment practice in the classroom, shared problem solving, and planning for future practice. Between TLCs, teachers conduct structured peer observations focusing on the use of formative assessment strategies.

The aim of the research project is to assess and better understand the process, outcomes, and impact of scaling up an educational intervention. The aim is to provide useful information to SSAT, but also to the Education Endowment Foundation (EEF) for when it is considering supporting education interventions to scale in the future. The study also contributes to the limited body of evidence on the barriers to and facilitators of scaling education interventions in general. While individual interventions are often evaluated for impact, little is known about how to scale-up those interventions effectively.

This report gives interim findings from the first phase of the research. A mixed-methods approach was used, which involved collecting and analysing data that came from schools and from SSAT. At the school level, six case studies were conducted in this first phase of the research; combining observations and interviews with school staff. Two surveys were also conducted with all schools that were part of this first academic year of new scaling (14 schools in total). At the SSAT level, a combination of observations, interviews, document reviews, and administrative data reviews were conducted. Some administrative data from SSAT was also integrated with the case study findings. The second phase of the research will cover the same research topics, but with more focus on the sustainability of the programme and on a different approach to scaling that will be adopted in the 2022/23 academic year (the use of a subsidy from the Department for Education).

The project began in December 2019 and is expected to finish in February 2024.

Table 1: Summary of findings

Research topics	Findings
Strategy	To support scaling, SSAT made two main changes to their approach to programme management. First, they introduced a new CRM system in the 2021/22 academic year, helping to digitise and automate a series of time-consuming manual processes. This helped reduce the burden on key SSAT staff, and also allowed them to better evaluate their marketing methods (e.g. measuring open/click rates). Second, the size of the team responsible for the EFA programme was expanded to provide additional capacity during the scale-up phase
	No significant changes were made to the core content of the programme (TLC sessions) to support scaling. The most significant change to the programme during the research period was a shift to online delivery due to Covid. Some schools had to complete their launch events online, and some meetings between school leads and mentors that usually happened in-person were conducted virtually
	The strategy included a target to reach an additional 50 schools in the 2021/22 academic year, but only 14 schools started the programme during this period. The scaling strategy was not therefore successful in achieving its core aim. Further details on possible reasons for this can be found in 'Reach and recruitment' below
Structures, systems, and processes	SSAT collected a range of qualitative and quantitative M&E data on the programme—some of these measures were collected through online forms, which made analysis

	more straightforward and reduced the burden on SSAT staff, which in turn supported the scaling of the intervention
	SSAT's current M&E approach relating to the monitoring of programme outcomes was deemed to be a proportionate and efficient approach to impact management (given the existing impact evidence from an effectiveness trial), and no further improvements were suggested in this area. Some potential improvements to programme M&E were identified across school reach and recruitment, programme implementation, and programme quality. It is suggested that SSAT: i. ensure complete data is collected where possible; ii. collect more structured quantitative feedback from stakeholders; iii. conduct periodic analysis of the data collected; and iv. develop formal M&E plans relating to each area
	The research found that SSAT collected some limited data on mentors (their programme support staff), but that no substantial analysis was conducted on the data currently collected. No systems were identified to ensure that the data collected was acted upon. The main potential improvement identified here was for SSAT to begin collecting and analysing data on mentor reach and recruitment, and on mentor training and support. To support all these suggested improvements, BIT has worked with SSAT to develop a comprehensive M&E plan and a new set of feedback surveys to support this
Reach and recruitment	In total, 23 schools were recruited in the 2021/22 academic year, of which 14 schools actually began the programme—this was considerably lower than the target of 50 new schools that SSAT were aiming for. Around 35% of the schools recruited had above-average levels of FSM and the majority (86%) were rated 'Good' or 'Outstanding' by Ofsted. Around 64% of the schools recruited were located in London or south-east, with 21% located in an 'Opportunity Area'[1]
	A total of 375 schools entered the EFA sales pipeline between January 2019 and June 2022, with 14 schools from this pool beginning the programme in the 2021/22 academic year—implying a sales conversion rate of 3.7%
	A range of factors were identified that help explain this conversion rate, including concerns from schools about increasing workload for their staff (particularly in the wake of Covid), concerns about the flexibility of the programme, and a potential lack of momentum in the sales approach
Fidelity, contextual factors, and sustainability	SSAT staff suggested that the intervention contains three essential features: i. teachers participating in TLCs and reflecting on their practice; ii. teachers observing one another between TLCs; and iii. producing personal action plans to support them as they try new techniques in their classroom. In general, school leads, TLC leaders, and classroom teachers considered the TLCs and peer observations to be central to the programme (and stuck faithfully to these components), but there was some evidence to suggest a more inconsistent approach to personal action planning
	A wide range of adaptations to the programme (made by school leads, TLC leaders, and classroom teachers) were identified—including small changes to the language used within the programme, changes to the delivery format, changes to the size of the TLC groups, as well as more substantive changes to the programme content. Two factors particularly facilitated the effective adaptation of the intervention: i. regular opportunities for TLC leaders and school staff to provide feedback on the programme to the school's leadership; and ii. the quality and regularity of the relationship between the school lead and the EFA mentor
	Intervention fidelity was encouraged by SSAT via the EFA mentors but was largely the responsibility of the school lead and TLC leaders. Some schools developed innovative systems to reward and incentivise adherence, while others incorporated EFA into their performance management processes. Some of the schools with the highest fidelity took additional measures to build staff buy-in to the programme, which in turn encouraged adherence

---

[1] See https://www.gov.uk/government/publications/social-mobility-and-opportunity-areas.

Fidelity to the intervention was monitored by SSAT (through the EFA mentor), the school lead and TLC leaders. The EFA mentor monitored intervention fidelity through a series of informal (often virtual) calls with the school lead, coupled with an in-person visit to the school at the end of the first year of implementation. The level of monitoring conducted by school leads and TLC leaders varied across schools. Schools with the highest fidelity to the programme tended to have school leads and TLC leaders that actively monitored staff attendance at TLC sessions and their completion of peer observations and personal action plans. Some schools took a less active approach to monitoring fidelity given the additional pressures that many schools had been under in the wake of Covid, which often resulted in lower overall fidelity to the intervention

A wide range of factors were identified that affected the adoption and implementation of the programme—including the alignment of the programme to other CPD priorities, the ongoing impact of Covid, the school's leadership, the school's culture, and specific characteristics of a school's staff and pupils

BIT, Behavioural Insights Team; CPD, Continuing Professional Development; CRM, Customer Relationship Management; EFA, Embedding Formative Assessment; FSM, Free School Meals; M&E, monitoring and evaluation; Ofsted, Officed for Standards in Education, Children's Services and Skills; SSAT, Schools, Students, and Teachers Network; TLCs, Teacher Learning Communities.

Interpretation and implications

There was evidence that SSAT had developed and implemented a range of measures to support the scaling up of the EFA programme. However, SSAT were unable to recruit their target number of schools for the 2021/22 academic year. There was some evidence to suggest that this was at least partly driven by the lasting impact of Covid on the UK's education system, with some schools more reluctant to begin new programmes during a period of stretched resources and instability. However, a range of other factors may also help explain the low sales conversion rate, such as schools' previous interactions with EFA, or their concerns about the flexibility and potential burden of the programme (including the fact that it is a whole-school intervention). To reach their scaling targets, SSAT should consider these factors, and then review and implement the relevant recommendations set out in Appendix B where possible.

Evidence was also collected on the factors affecting the adoption and implementation of the programme, and a school's level of fidelity. For example, there was evidence to suggest that some of the schools with highest levels of fidelity to the programme implemented a range of innovative and school-specific measures to encourage staff to engage positively with the programme, and coupled this with regular, consequential monitoring of fidelity. Some of these schools were also able to successfully implement the programme through a period of Covid-related instability. SSAT should explore these factors and consider ways that the programme can incorporate these additional measures that were suggested to support schools with implementation.

Based on the findings contained in this report, 47-specific recommendations were made to SSAT covering the full range of research topics. These can be found in Appendix B.

Five high-level general recommendations were also identified for organisations wishing to scale educational interventions:

- If there is evidence of sufficient demand to scale the intervention, ensure scaling strategies include a comprehensive plan for recruitment.
- Collect and analyse data that helps to explain the types of settings reached, and the possible reasons why some settings choose to not sign up.
- Automate and streamline internal processes where possible.
- Identify and encourage the facilitators of effective adaptation of the intervention.
- Distribute responsibility for encouraging and monitoring fidelity to the intervention.

Introduction

Background and intervention

Despite considerable research in recent years, identifying whole-school educational interventions that are cost-effective in improving pupil outcomes has proved challenging. At the same time, there has been growing demand from the education sector for programmes and interventions that have a proven evidence base. The Embedding Formative Assessment (EFA) is one such whole-school intervention that has evidence for improving pupil outcomes (EEF, 2018).

EFA is a professional development programme that aims to improve pupil outcomes by embedding the use of formative assessment strategies across a school. The programme was developed jointly by the Schools, Students, and Teachers (SSAT) Network and Dylan William. Black and William (2009) define formative assessment as practices in which:

> …evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited. (Black and William, 2009)

The programme was developed by SSAT to address three perceived barriers to EFA: a lack of understanding of the value of formative assessment; a lack of time dedicated to embedding it; and the complexity involved in changing teachers' practices. Taking account of these challenges, EFA combines regular meetings, ongoing feedback, and clear guidance in an attempt to make formative assessment part of routine practice. The delivery of EFA in a school is supported by the roles outlined in Table 2, split between SSAT and the participating school.

Table 2: EFA roles and responsibilities

Organisation	Role	Responsibilities
SSAT	Senior education lead	Oversight of EFA recruitment, on-boarding, and implementation across all participating schools. Training and line management of EFA mentors
	Programme manager	Day-to-day programme management
	EFA mentors[2]	Support school leads to set up and implement EFA through a programme induction meeting, ongoing ad hoc remote support, and a meeting at the end of year 1 of implementation
School	EFA school lead	Manages the delivery of EFA within the school. Appoints and line manages the TLC leads
	TLC leaders	Facilitate cross-departmental TLC workshops with teachers and senior leaders. There are multiple TLC leaders within each school, each leading a group of roughly 10–12 teachers. The number of TLC leaders within a school depends on the size of the school. TLC leaders are also class teachers
	Class teachers	Participate in the EFA programme by embedding the formative assessment in their own practice and by supporting their peers to do the same through observations, feedback, and practice sharing

EFA, Embedding Formative Assessment; SSAT, Schools, Students, and Teachers Network; TLC, Teacher Learning Community.

---

[2] Two types of mentors are employed within the EFA programme. School-based mentors are employed by a school and support a limited number of schools to implement the EFA programme. School-based mentors typically have experience implementing the programme in a school (e.g. as a school lead). The second type of mentors are consultants who do not typically work full-time in schools. They typically have considerable experience in the education sector (e.g. as senior leaders in schools), including experience implementing Continuing Professional Development (CPD) programmes.

To support delivery, schools receive detailed resource packs to run structured monthly workshops, known as Teacher Learning Communities (TLCs), which focus on enhancing understanding of formative assessment strategies, reflecting on formative assessment practice in the classroom, shared problem solving, and planning for future practice. Between TLCs, teachers conduct structured peer observations focusing on the use of formative assessment strategies.

More information on the EFA programme can be found in the report of the effectiveness trial that preceded this evaluation (Speckesser *et al.,* 2018) and on SSAT's website[3].

Study rationale and aims

Background

EFA is a well-developed and well-tested intervention. An Education Endowment Foundation (EEF) effectiveness trial of the programme with 140 schools found that students in the intervention schools made an additional 2 months' progress in their Attainment 8 General Certificate of Secondary Education (GCSE) score versus students in comparison schools (Speckesser *et al.*, 2018). As a result of this success, SSAT is trying to scale the programme up, with funding from the EEF.

Other programmes of formative assessment have also been shown to have positive effects when implemented well. One study estimated a very large positive effect (up to half of a GCSE grade per student), and a meta-analysis of studies on this approach suggests that about an additional 3 months' progress can be achieved if formative assessment practices are supported by good professional development (EEF, 2015). However, two studies in the most comprehensive meta-analysis on the topic date have estimated negative effects that are significant at the 5% level: one from a formative assessment intervention for middle-school science teachers; and one from an all-subjects formative assessment programme for Year 7 pupils in an English secondary school (Newman *et al.*, 2021). Some research has also shown that effective formative assessment practices can be difficult to implement (Gorard *et al.*, 2014). The process evaluation for the EFA effectiveness trial also highlighted substantial variation in implementation at the school level; however, SSAT felt most of the observed adaptations were acceptable and part of the necessary flexibility of the programme. Furthermore, only a small number of successful educational programmes have been scaled-up, with the process of scaling being formally evaluated. So, while there is strong evidence to suggest that EFA—and formative assessment more broadly—can be an effective way of increasing pupil attainment, implementing it well at scale is not a given.

Given this context, the Behavioural Insights Team (BIT) has been appointed by the EEF as an independent evaluator to assess and better understand the process, outcomes, and impact of scaling up. The aim is to provide useful information to SSAT, and also to the EEF for when it is considering supporting education interventions to scale in the future. The study also aims to contribute to the body of evidence on the barriers to and facilitators of scaling educational interventions in general.

Scaling targets

SSAT received funding and support from the EEF to enable the scaling of EFA. SSAT and the EEF agreed scaling targets separate from this evaluation, but which are being monitored throughout the evaluation. Where the evaluation refers to the 'scaling of EFA', this is a reference to the scaling targets agreed with the EEF.

The initial targets involved SSAT scaling the delivery of EFA from 20 secondary schools to an additional 125 secondary schools between September 2021 and July 2023, with the following specific targets.

- 50 new schools starting in the 2021/22 academic year ('year 1' of the research); and
- 75 new schools starting in the 2022/23 academic year ('year 2' of the research).

During the 2021/22 academic year, SSAT was awarded some new funding by the EEF through the Department for Education's (DfEs) Accelerator Fund. At this point, the scaling targets were revised. SSAT's new target is to have launched EFA in 150 new schools by July 2022 in three DfE target regions (the North, East Midlands and Humber, and West Midlands). To support the achievement of this new target, the DfE is now subsidising the fee paid by schools for

---

[3] The full report can be accessed at: https://bit.ly/3lOqpAi.

the programme. Normally, the programme costs schools £5,299 (for 2 years of support and a pack of resources), and this is what the 14 new schools in 2021/22 have paid. With the DfE subsidy, new schools will pay £1,589 (30% of the standard fee).

This interim report covers SSAT's initial scaling plans and the (unsubsidised) schools that were recruited as part of this. For the next phase of the research, the study plan will be revised to account for the new scaling plans.

Research questions

The research questions cover the following seven interrelated topics are:

- Topic 1: Strategy;
- Topic 2: Fidelity;
- Topic 3: Structures, systems, and processes;
- Topic 4: Reach and recruitment;
- Topic 5: Contextual factors;
- Topic 6: Sustainability; and
- Topic 7: Cost.

Findings relating to Topics 1 to 6 are included in this report. The research questions for each of these topics are presented below, and a more detailed breakdown of all the research questions can be found in Appendix A.

Topic 1: Strategy

- What is SSAT's strategy for scaling up the EFA programme?
- How does SSAT's strategy for scaling evolve over time?
- What factors influence changes to the scaling strategy?
- What role does the EEF play in helping SSAT to achieve readiness for scaling up?

Topic 2: Fidelity

- What are the essential features of the intervention, and what adaptations are appropriate (and required to support scaling)?
- How does the approach taken to scaling, support or hinder fidelity?
- How is intervention fidelity managed?
- What are the barriers to, and enablers of, the effective adaptation of EFA?

Topic 3: Structures, systems, and processes

- What challenges are there organisationally when making a sizeable change in the scale of implementation of the EFA programme, and how are these overcome?
- How well do SSAT's monitoring and evaluation (M&E) systems support data-based decision-making and how can they be improved?

Topic 4: Reach and recruitment

- What is SSAT's sales process/pathway for the EFA programme?
- How many and what types of schools are SSAT reaching and successfully recruiting?
- How do schools respond to the sales approach?

Topic 5: Contextual factors

- What school characteristics affect the adoption and implementation of EFA and how (e.g. culture, school-type, leadership, subject(s) taught, characteristics of individual teachers, and mentors)?

- What are the facilitators and barriers—in the context outside of schools—to the scale-up of the EFA programme (e.g. education policy, funding, networks between schools)?

Topic 6: Sustainability

- What indicates that EFA has been embedded in school practice?
- What are the facilitators of, and barriers to, embedding the EFA programme in a school?
- How viable is it for schools to sustain the use of EFA on an ongoing basis after the end of the scale-up? What are the factors that affect this?
- Is EFA being institutionalised at levels other than the school? What are facilitators of, and barriers to, 'vertical' scaling up,[4] and how can the barriers be addressed?

The only research question set out in the study plan that is not being addressed in this report relates to the cost of implementing the programme (Topic 7: Cost). The year 1 cost data had been collected at the time of writing this report, but the analysis is for 2 years and so for Topic 7: Cost, the data will be reported at the end of year 2.

---

[4] 'Vertical scaling' refers to the 'institutionalisation' of a programme or practice. At the highest level, vertical scaling could come in the form of new central government policy that supports the intervention. It could also come at different levels however, for example within local authorities or multi-academy trusts, and does not have to involve formal policy change (WHO, 2010).

Research design and methods

Summary of design and phases

This evaluation aims to improve our understanding of many aspects of the scaling up of a complex intervention and covers a lot of research questions. To help make this manageable, the research has been broken down into four phases, with each phase focused on a subset of topics (see Table 3). These phases were defined in collaboration with SSAT to ensure that the feedback and interim findings were given to SSAT in as timely a manner as possible.

Table 3: Research design broken down by phase

Phase	1 and 2. Mapping strategy, resources, and processes	3. Implementation in year 1	4. Embeddedness in year 2	5. Final reporting
Common research topics across all phases	Strategy; Fidelity; Structures, systems, and processes; Reach and recruitment; Contextual factors			
Focus of phase	M&E infrastructure, organisational processes, defining fidelity, MI collection and uses	M&E of year 1 in school programme implementation	M&E of year 2 in school programme embeddedness and sustainability	Strategy and support
Data collection methods	Administrative data, document review, SSAT observations and interviews, SLT interviews	Administrative data, school surveys, school observations and interviews, SSAT observations and interviews	Administrative data, school surveys, school observations and interviews, SSAT observations and interviews	Analysis feedback workshop with SSAT
Lines of inquiry with SSAT strategic leads	Strategy for scaling, defining fidelity, M&E infrastructure, and school recruitment process	Strategy for scaling, fidelity management, barriers to and enablers of fidelity, organisational challenges of scaling, reach and recruitment, contextual factors outside schools, and cost	Changes to strategy, fidelity management, barriers to and enablers of fidelity, organisational challenges of scaling, reach and recruitment, contextual factors outside schools, sustainability, and cost	Interpretation of findings
Lines of inquiry with EFA mentors (school-facing staff)	Defining fidelity, fidelity management	Fidelity management, barriers to and enablers of fidelity, and contextual factors inside schools	Fidelity management, barriers to and enablers of fidelity, and contextual factors inside schools	Interpretation of findings
Lines of inquiry with school staff	Barriers to and facilitators of adoption	Fidelity management, barriers to and enablers of fidelity, contextual factors inside schools, contextual factors outside schools, response to sales approach, and cost	Fidelity management, barriers to and enablers of fidelity, contextual factors inside schools, contextual factors outside schools, response to sales approach, and cost	NA

EFA, Embedding Formative Assessment; M&E, monitoring and evaluation; MI, Management Information; NA, not applicable; SLT, senior leadership team; SSAT, Schools, Students, and Teachers Network.

The detailed questions identified above are best answered by a combination of research methods, so a mixed-methods approach has been taken. This approach involves collecting and analysing data that comes from schools and from SSAT. At the school level, ten case studies are being conducted that combine observations, interviews, and surveys. Between them, these research activities cover four research topics: fidelity; reach and recruitment; contextual factors; and sustainability. At the SSAT level, a combination of observations, interviews, document reviews, and administrative data reviews cover all research topics. Some administrative data from SSAT is also being integrated with the case study findings. In addition to this, a cost evaluation is being conducted using structured interviews and administrative data to answer the research question for Topic 7: Cost, however this will not be reported until the end of year 2.

Data collection methods

School case studies

Methods

A collection of ten comparative case studies of participating schools (out of a potential 150 schools if recruitment targets are met) are being conducted to gain an in-depth understanding of how the programme is being implemented. Each case study has involved observations of key programme activities, interviews with staff involved in the delivery of EFA within a school, and analysis of SSAT's programme monitoring data for the school.

The observations addressed three of the research topics: fidelity; contextual factors; and sustainability. Three sessions were observed in each school: one TLC workshop; one peer feedback session; and one lesson observation. Semi-structured observation guides were used to capture field notes. These were developed with some input from SSAT to ensure that the details of the programme and its scaling were accurately reflected. All observations were non-participatory (i.e. the researchers did not participate in the activity being observed), direct and undisguised (i.e. the activity participants did know that researchers were present, and who were observing and taking notes).

Table 4: Case study methods for year 1 of the programme

Activity / participant group	Method	Number of data pieces	Number of people involved
TLC workshop	Observation	1	15[5]
Formative assessment in practice in the classroom	Observation	1	30
EFA school lead	Interview	1	1
TLC leaders	Paired interview	1	2
Class teachers	Paired interview	1	2
SSAT EFA mentor	Interview	1	1
	Total for each Wave 1 case study	**6**	**51**

EFA, Embedding Formative Assessment; SSAT, Schools, Students, and Teachers Network; TLC, Teacher Learning Community.

The interviews addressed four of the research topics: fidelity; reach and recruitment; contextual factors; and sustainability. In each case, interviews were conducted with the EFA school lead, two TLC leaders, and two class teachers. For the latter two categories the interviews were conducted in pairs to increase the diversity of the sample and

---

[5] This is 50% more than estimated in the protocol. SSAT advises that TLC groups contain approximately 10 to 12 teachers, but also accepts that groups may need to be larger for logistical reasons. Fifteen is at the upper end of SSAT's expectations of group size. TLC group size does not appear as an important factor in any of the analysis below.

to encourage some peer-to-peer reflection. Semi-structured guides were used for all interviews. These were also developed with some input from SSAT.

The data collection in these schools is taking place across the 2 years of the programme in order to gain insight into both early and later implementation. Schools were recruited by SSAT to the EFA programme in two cohorts ('waves'): Wave 1 schools were recruited to start in September 2021; and Wave 2 schools were recruited to start in September 2022. To take advantage of this, six case studies are being conducted with Wave 1 schools over both years of implementation.[6] A further four lighter touch case studies will be conducted with Wave 2 schools in their first year of implementation. For this interim report, data has only been collected for the six initial case studies from Wave 1.

Table 4 above provides an overview of the research activities for each case study school, the participant groups involved, and the number of qualitative pieces of data that were collected in year 1 (where a piece of data is either a transcript from an interview or a set of field notes from an observation).

Sampling

The aim of the sampling strategy for the case studies was to capture the range and diversity of experiences that supported a comprehensive response to the research questions. To achieve this, a stratified purposive sample was created, where schools were selected based on key characteristics, stratified by their level of fidelity to the intervention. To achieve this, two schools with low fidelity and four schools with high fidelity to the programme were selected. For this purpose, fidelity was defined by SSAT, using management data and the subjective judgement of programme management staff. Within these two groups (high and low fidelity), the sampling aimed for variation in the Office for Standards in Education, Children's Services and Skills (Ofsted) rating and school size (where greater than 1,000 pupils on roll was classified as 'large' and less than 1,000 pupils was classified as 'small'). Table 5 summarises the sample achieved for case studies in year 1. It shows that good variation was achieved in terms of fidelity and school size, but that all recruited schools received a 'Good' Ofsted rating at their last inspection. This lack of variation in Ofsted rating is reflective of the total population of schools starting the programme in the 2021/22 academic year and is commented on further in the section on 'Structures, systems, and processes' below.

Table 5: Summary of case study school sample for year 1 of the programme

School	Fidelity	Ofsted rating	Size
School 1	High	Good	Small
School 2	High	Good	Small
School 3	High	Good	Large
School 4	Low	Good	Large
School 5	Low	Good	Small

Ofsted, Office for Standards in Education, Children's Services and Skills.

School surveys

Methods

Two online surveys were conducted with all Wave 1 schools, based on the two key EFA staff groups in each school: TLC leaders; and class teachers. Both surveys covered fidelity, contextual factors, and sustainability. There was some overlap in survey questions between the two groups—because TLC leaders are also teachers who implement EFA with their students—but the TLC leaders survey contained additional questions that only relate to their additional role as facilitators of EFA.

---

[6] In the event a Wave 1 school drops out of the programme or the evaluation before the second year, we will replace that school with a Wave 1 school with similar sampling characteristics.

Surveys were administered to these groups at two time points in year 1. A baseline survey was issued during the first term (in October) to capture early impressions, and a follow-up survey, covering the same topics, was issued towards the end of the academic year (in April) to see if responses changed over the course of a year.

Sampling

A census approach was used for all survey sampling, where all members of each sampling population were invited to complete the relevant surveys. For the first year of sampling, the sampling population comprised all Wave 1 schools. Only data from Wave 1 schools will therefore be presented in this interim report. Table 6 shows that the response rate for TLC leaders was quite high for a non-compulsory survey of this nature. The response rate from class teachers was lower, which is to be expected given their lower level of engagement with the programme as compared to TLC leaders. Most of the survey data used in the analysis for this report comes from the endline survey and having the majority of TLC leaders responding to this survey gives us good confidence that we have captured a wide range of experiences from this cohort. While the 142 class teacher respondents to this survey are unlikely to be representative of the overall population (at 14% of the total), they still provide useful insights when contextualised with the qualitative data.

Table 6: Sample size for surveys

Participant group	Estimated pop. size in Wave 1	Baseline survey responses in Wave 1 (% of estimated pop.)	Endline survey responses in Wave 1 (% of estimated pop.)
TLC leaders	98	38 (39%)	57 (58%)
Class teachers	980	120 (12%)	142 (14%)

Pop., population; TLC, Teacher Learning Community.

SSAT observations and interviews

Methods

Key activities delivered by SSAT in the recruitment and set-up phase for new schools were observed, and interviews were conducted with a range of SSAT staff involved in the scaling process. The observations addressed four of the research topics: fidelity; reach and recruitment; contextual factors; and sustainability. Two types of sessions were observed at this level: one EFA open day and one EFA launch event. Open days were hosted by ambassador schools (schools that have embedded EFA and have volunteered to host open days) and provided an opportunity for staff from interested schools to learn more about EFA. The launch event was for EFA school leads, TLC leaders, and school governors, to introduce the programme at a high level and ensure that the senior school staff understood their roles. Semi-structured observation guides were used to capture field notes. All observations were non-participatory (i.e. the researchers did not participate in the activity being observed), direct and undisguised (i.e., the activity participants did know that researchers were present, and who were observing and taking notes).

Semi-structured interviews were conducted with key members of the SSAT team. The first round of interviews took place before the first year of scaling started, and were held with the senior education lead, head of brand, head of business development, project manager, and two EFA mentors. The purpose of these interviews was to understand how the organisation was structured and any changes that had been made or were planned in order to support scaling. They also covered the components of the scaling strategy that each person was responsible for and any challenges they foresaw within their strategic areas.

Table 7 provides an overview of the research activities with SSAT staff, the participant groups involved, and the number of qualitative pieces of data collected in year 1.

Table 7: Research activities with SSAT staff

Activity / participant group	Method	Number of data pieces	Number of people involved
EFA open day	Observation	1	NA
EFA launch event	Observation	1	NA
Senior education lead	Interview	1	1
Project manager	Interview	1	1
Head of brand	Interview	1	1
Head of business development	Interview	1	1
EFA mentors (x 2)	Interview	4–6	6
	Total	**6**	**10**

EFA, Embedding Formative Assessment; NA, not applicable; SSAT, Schools, Students, and Teachers Network.

Sampling

Observations of the open day and the launch event were selected for convenience, based on the timing of the research, and were independent of the case studies (i.e. neither school involved in these observations was part of the case study sample).

Interviews were conducted with all SSAT staff who had lead responsibility for an element of the scaling strategy.

Sales interviews

To better understand school leaders' broader perceptions of the programme and sales approach, a series of semi-structured interviews were conducted with senior leadership team (SLT) members who were in the sales pipeline for EFA. These interviews focused on responses to the EFA sales approach, as well as perceived barriers and facilitators to signing up to the programme.

Sampling

Six interviews were conducted with SLT members that were in the sales pipeline. It had been planned that these interviews would include SLT that were not interested in taking up the programme, as well as those that were unsure and those who were in signing up. Given challenges to recruitment driven by ongoing Covid disruption, it was not possible to interview SLTs that were not interested in the programme.

Administrative data review

SSAT's M&E processes were analysed to address three research topics: fidelity; structures, systems, and processes; and reach and recruitment.

The first stage of the M&E review involved an evaluation of the data that was collected and analysed by SSAT, and the systems that were used for this purpose. It asked whether these data and systems effectively supported the implementation and scaling of EFA.

The second stage of the M&E review involved analysis of the actual data that was collected on fidelity and reach and recruitment at the end of the first year of scaling. This analysis was conducted for all Wave 1 schools, as well as at the case study level in order to integrate it with the qualitative findings.

Deviations from the study plan

There were three deviations from the study plan.

- Challenges observing the mentor training. It had initially been planned to observe an EFA mentor training session at the start of the scaling period. However, there were no training programmes delivered in this period. Instead, a mentor training programme was observed in April 2022, but for the purposes of a separate project by the EEF with a different set of research questions. Where there is overlap with this study's research questions however, findings from this observation have been integrated into this report.
- Reduced sample for sales interviews. Ten interviews with SLT staff were planned to gather evidence relating to EFA's sales process. The original plan for recruitment had been for BIT researchers to attend an EFA open day and recruit SLT staff in-person. With all open days shifting online, this method of recruitment was no longer viable, and recruitment had to be conducted cold by email and telephone. This made it particularly challenging to recruit staff from schools that had been considering the programme at one stage but had ultimately chosen to not sign up.
- Reduced sample for TLC observations. In one case study school it was not possible to observe a TLC. This was because the school lead was unable to host visitors to the school and then clashes with exam season. All other research activities were completed for this school.

Timeline of research activities

Table 8: Project timeline

Phase	Timing	Key activities
1. Set up and kick-off	December 2019 – March 2020	- Kick-off meetings with SSAT and the EEF - Finalise data sharing documents and arrangements - EFA document review - Scaling strategy TOC workshop **TOC summary report (March 2020)**
2. Organisational processes, defining fidelity, MI data review	April 2020 – August 2021	**Study plan finalisation (April 2021)** **Pre-mortem workshop and summary report (April 2021)** - Interviews with key SSAT staff to map organisational processes and define programme fidelity (current and at scale) - Review MI extracts **First feedback (July 2021)**
3. Year 1 school implementation, reach and recruitment	September 2021 – August 2022	- Pre- and post-surveys with school staff - Case studies (x six schools) - Analyse MI - Interviews with key SSAT staff about strategy, organisational capacity, and processes **Second feedback (December 2021)** **Third feedback (July 2022)** **Interim report (September 2022)**
4. Year 2 school embeddedness and sustainability	September 2022 – August 2023	- Pre- and post-surveys with school staff - Case studies (x ten schools) - Analyse MI - Interviews with key SSAT staff about strategy, organisational capacity, and processes **Fourth feedback (July 2023)**
5. Final reporting	September 2023 – February 2024	**Final findings and post-mortem workshop (October 2023)** **Final report (draft November 2023, final February 2024)**

EEF, Education Endowment Foundation; EFA, Embedding Formative Assessment; MI, Management Information; SSAT, Schools, Students, and Teachers Network; TOC, theory of change.

The key research activities and deliverables for each phase are outlined in Table 8. At the time of writing this interim report, all activities in Phases 1, 2, and 3 have been completed.

Analytic approach

For the interviews with SSAT staff and reach and recruitment interviews with SLT, Braun and Clarke's (2006) six-step approach was used, which involved coding the transcripts and identifying emerging themes. Themes then underwent a further round of classifying and were sorted into high-level themes and sub-themes.

For the case study data, a framework approach was used with within-case analysis being conducted before between-case analysis (Ritchie *et al.*, 2013). This first involved identifying emerging themes through familiarisation with the data. Then, an analytical framework was created using a series of matrices each relating to an emergent theme. The columns in each matrix represented the key sub-themes drawn from the findings and the rows represented individual participants interviewed or activities observed.

The interview and observation data was summarised in the appropriate cell, which means that all data relevant to a particular theme was noted and easily accessible. This enabled a systematic approach to analysis that was grounded in participants' and schools' accounts. The next step of analysis involved working through the charted data to draw out the range of schools' experiences and participants' views, while identifying similarities, differences, and links between them. Thematic analysis (undertaken by looking down the theme-based columns in the framework) identified concepts and themes and the case-based analysis (undertaken by comparing and contrasting rows in the framework) allowed for links within cases to be established and cases to be compared and contrasted with each other.

For all qualitative analysis, a balance was maintained between deduction (using existing knowledge and the research questions to guide the analysis) and induction (allowing concepts and ways of interpreting experience to emerge from the data). Verbatim participant quotations and case examples are used to provide evidence and exemplify the theme(s) discussed in the paragraph before the quotation. Quotations have been selected by considering multiple factors including how well they exemplify the theme(s) discussed.

As qualitative data can only be generalised in terms of range and diversity and not in terms of prevalence, the analytical outputs focus on the nature of experiences, avoiding numerical summaries or language such as 'most' and 'majority'.

SSAT also played an informal role in the analysis, by offering their reflections in response to feedback and more formal findings that have been presented to them over the course of the evaluation.

Approach to feedback and reporting

To support the formative aims of this study, four types of feedback and reporting will be delivered.

- Pre-mortem workshop. This was conducted before the first year of scaling started. The goal of this session was to help SSAT assess the potential risks and threats to the scale-up, to support more effective and comprehensive planning. It brought the SSAT scaling leadership team together for a 2-hour session to jointly imagine ways in which the scale-up might fail, and then work backwards to imagine the causes of these failures. The output of this session was a set of newly identified risks for SSAT to take away and plan mitigation strategies for.

- Findings and formative feedback. Four feedback slide decks are being produced over the course of the evaluation. The aim of these slide decks is to provide timely findings to SSAT that they may wish to act upon during the scaling process. Three feedback slide decks have been provided to SSAT at the time of writing this report.

- Formal reports. Two formal reports are being produced: one interim report at the end of the first year of scaling (this report); and one final report after the second year of scaling. These reports cover a combination of information that SSAT will have already received in the feedback slide decks, as well as some new findings. SSAT is encouraged to review and act upon these new findings.

- Post-mortem workshop. This session will provide SSAT with headline draft findings from the final report. There are three aims of this: i. to give important results to SSAT as quickly as possible so that they can continue to develop their strategy; ii. to allow SSAT to offer feedback on the findings to add insight and depth to our final

stages of analysis; and iii. to support SSAT to assess the validity of the risks and mitigation strategies identified in the pre-mortem workshop to aid future planning.

Findings

The research completed in year 1 of the evaluation generated evidence relating to six of the research topics: Strategy; Fidelity; Structures, systems, and processes; Reach and recruitment; Contextual factors; and Sustainability. The findings relating to each topic are presented in the following sections. Note, the findings relating to fidelity, contextual factors, and sustainability have been presented in the same section, given substantial overlap between some of the findings.

Strategy

Scaling framework

Scaling up is a complex process that occurs within a dynamic system, involving a multitude of interactions between the different components of the scale-up process. In partnership with SSAT, a framework was developed to build a shared understanding of their scale-up approach, and to support BITs' analysis throughout the evaluation. The framework draws primarily on two existing scale-up frameworks by the World Health Organization and ExpandNet (2009)[7] and the Institute for Healthcare Improvement (Barker et al. 2015), and is depicted in Figure 1.
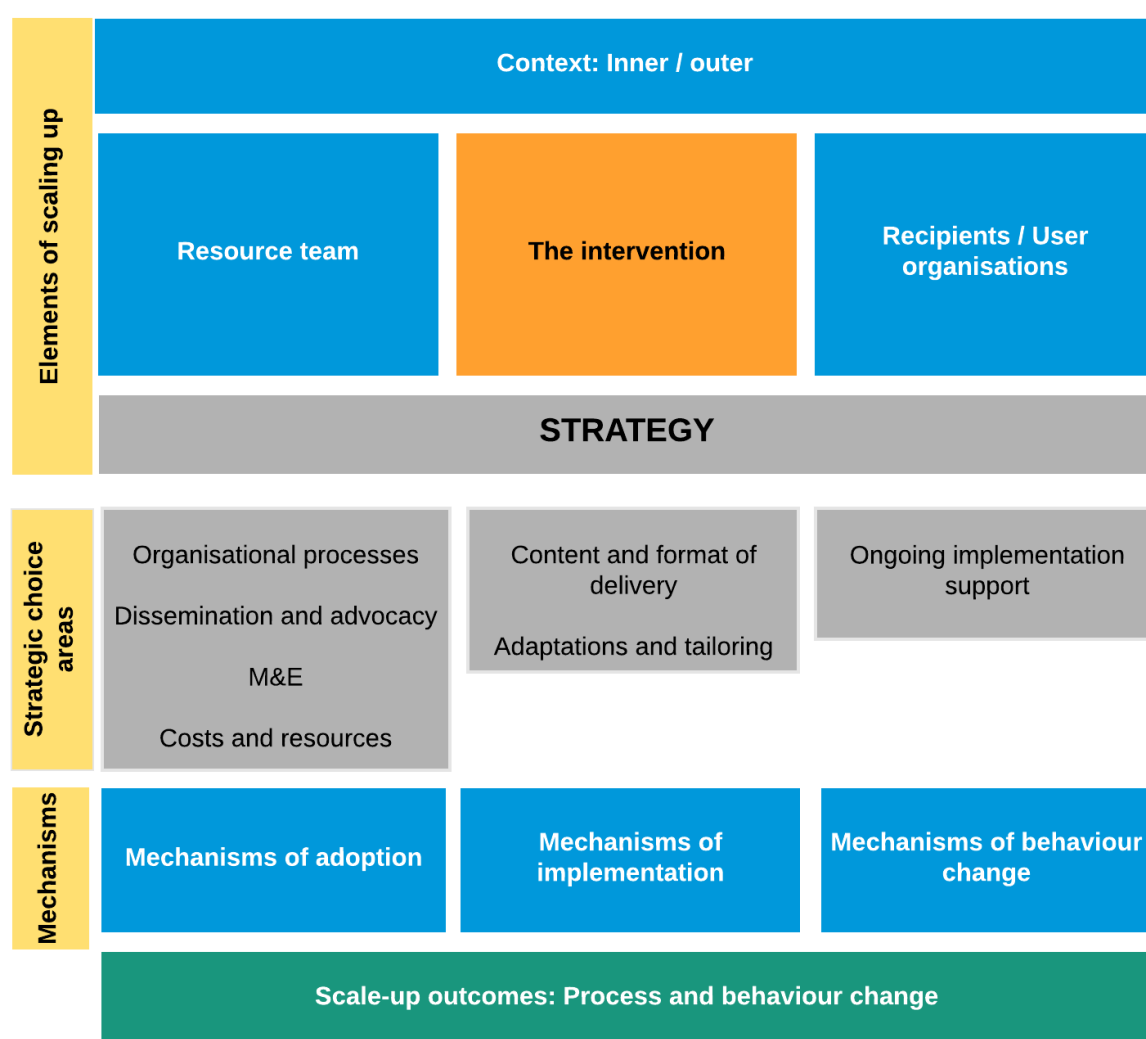

Figure 1: Scale-up framework. M&E, monitoring and evaluation.

The framework includes five elements of the scale-up: i. the resource team (SSAT); ii. the intervention (the EFA programme); iii. the user organisation (schools); iv. the context of operation; and v. the strategy.

---

[7] See https://expandnet.net/PDFs/WHO_ExpandNet_Practical_Guide_published.pdf.

The strategy is the set of actions and plans that binds the other elements together to achieve scale-up success. The strategic choice areas outline the various plans and actions that allow SSAT to roll out the programme on a larger scale. It is the sum of the plans that enable SSAT (the resource organisation) to transfer EFA (the intervention) to schools (the recipients) and that allows them to implement the programme. The following sections outline SSAT's strategic choices across these areas.

Strategic choices

Organisational processes

Before starting the scaling studied in this research, SSAT conducted a review of their Customer Relationship Management (CRM) processes and implemented a range of measures that made these processes more efficient. Automation, in particular, was incorporated into the process to reduce the burden on key members of staff. The system now automatically notifies the relevant teams when there are tasks for them to do. For example, if the programme lead signs up a school, they can click a button to automatically notify the EFA team who manage invoicing. Interviews with SSAT staff suggested that these automated processes were saving them time having to email other members of staff.

SSAT also updated the processes for their mentor support and management. The mentor records are now completed fully online, and the programme lead has access to a tracker spreadsheet that makes it easier to see whether any mentors have not completed the records. Mentors and the programme lead noted that this system was working much better than the previous solution, which involved extensive back-and-forth emails between the mentors and programme lead to check whether records had been completed.

SSAT also decided to expand the team responsible for EFA. There is now a dedicated programme lead who works solely on EFA, and the wider team involves a project officer and a project manager. Small tweaks were also made to the way the team functions, for example the programme lead now has an automated calendar booking system that allows anyone to schedule time in their diary. This was perceived to have reduced the number of unnecessary emails to schedule meetings.

Dissemination and advocacy

SSAT implemented a range of sales activities prior to a lead being generated. The EFA sales funnel consisted of marketing through various channels, including: emails, social media, conferences, features in Teach Secondary magazine and SecEd, post, and the education lead speaking directly to schools.

A small number of EFA ambassador schools also hosted open days for school leaders who were interested in the programme. This was considered to be a particularly powerful way of disseminating information about EFA as it was thought that schools were more likely to listen to and take advice from other schools. Central to SSAT's scaling strategy was using the ambassador schools network to expand the geographical reach of the programme. However, due to Covid, some of these open days were pushed to later in the year, and some of them had to be cancelled entirely. To replace these, SSAT began offering some online webinars, which involved the programme lead explaining the programme followed by an opportunity for schools to ask any questions.

In general, the dissemination and advocacy plans for the scale-up of the programme were targeted at headteachers and members of a school's SLT. SSAT's education lead also engaged with some local authorities about broader dissemination of the programme.

At the start of the scale-up evaluation, there was evidence to suggest that SSAT monitored the types of schools engaging with EFA. However, they were unable to use this information to target individual schools.

M&E

SSAT routinely collected data on recruitment and school characteristics, such as the proportion of pupils with English as an additional language (EAL), the level of pupils on Free School Meals (FSM), the school's Ofsted rating, and the school's current educational outcomes. SSAT also collected a range of data relating to the in-school implementation of the programme. This included contact details for the key school contacts, the results of the fidelity survey administered by mentors, the conversation records between the school lead and the mentor, and surveys of teachers and pupils.

The new CRM system introduced (see earlier sections) also allowed SSAT to evaluate their marketing methods in new ways. The new system allowed the team to identify the click rates and open rates for emails, which helped them to begin tracking the effectiveness of various marketing messages.

Content and format of delivery

The content of the TLCs has remained the same throughout the research period. However, the programme lead has begun the preparatory work to conduct a thorough review of the materials in the 2022/23 academic year.

One new addition to the delivery of the programme in the 2021/22 academic year was the introduction of regular mentor support sessions. These online webinars provided a forum for mentors to come together, share their experiences, and discuss any issues that they were facing. The EFA programme lead noted that these sessions were proving valuable for mentors, and this view was also shared by mentors themselves.

Adaptations and tailoring

The most significant adaptation to the programme overall was the shift to an online delivery due to Covid. Many launch events had to be completed online, and meetings between mentors and school leads had to also be conducted virtually where they may have been done in-person previously.

Ongoing implementation support

In the 2021/22 academic year, SSAT also took steps to try to formalise their response to typical questions from schools. It had previously been noted that a lot of the queries they received from participating schools were similar to each other (e.g. a large number of schools report issues with embedding peer observations). In response to this, the programme lead led efforts to produce documents that provide suggested solutions to common problems, with this resource available to mentors.

Scale-up outcomes

Central to SSAT's strategy for scaling was their intended outcomes for the scale-up. Based on discussions with SSAT and other materials seen at the start of scaling, their scale-up outcomes can be categorised into four distinct areas, as seen in Table 9.

Table 9: Scale-up outcomes

Category	Outcomes
Organisational outcomes	Recruit and train sufficient mentors to support schools
	Two models of delivery developed—i.e. 1-year and 2-year versions
Reach	Fifty new secondary schools starting the programme in the 2021/22 academic year
Embeddedness	Schools implement the programme within the broad parameters for fidelity of the programme
	Teachers across all subject areas are using formative assessment
Cost	EFA is scalable at a reasonable cost

EFA, Embedding Formative Assessment.

These targets were what SSAT were aiming for at the beginning of the scale-up process. SSAT were hoping to reach an additional 50 Wave 1 schools in the 2021/22 academic year. However, SSAT were unable to meet this target, with 14 schools starting the programme in the 2021/22 academic year. There are several possible reasons for this, including the ongoing impact of Covid for schools and their inability to take on whole-school interventions—this and other reasons are discussed in the section on 'Reach and recruitment'. Within schools, implementation targets were broadly met; addressed in detail in the section on 'Fidelity, contextual factors, and sustainability' below.

Structures, systems, and processes

Evidence was collected on SSAT's M&E data and systems, to understand how they currently support the implementation and scaling of EFA. This section outlines the findings from this evidence, and is split into two parts: the first section outlines how SSAT's systems support the 'schools programme' (i.e. their support and monitoring of schools completing EFA); and the second section outlines how SSAT's systems support the 'mentor programme' (i.e. their support and monitoring of mentors supporting the delivery of EFA).

Schools programme

Reach and recruitment data

Table 10 outlines the data currently being collected by SSAT relating to 'Reach and recruitment'.

Table 10: Data currently collected relating to reach and recruitment

Topic	Data collected	How collected?	When collected?
School information	- School name - School ID - School characteristics - Teacher contacts	- Auto from website to CRM - Manually entered - Manually entered - Auto from website to CRM and manual	When school enters pipeline (then updated ad hoc)
Awareness	Does school download a web doc?	Auto from website to CRM	Point of download
EOI	- Does school complete EOI webform? - Does school attend open day? - Does a teacher meet with SSAT? - Does a decision-maker meet with SSAT?	- Auto from website to CRM - Manually entered - Manually entered - Manually entered	- Point of completion - Post-open day - Post-meeting - Post-meeting
Pre-commitment	Does school receive a booking form?	Manually entered	After booking form issued
Sign up	Does school pay the invoice?	Finance system	When invoice paid

CRM, Customer Relationship Management; EOI, expression of interest; ID, identification; SSAT, Schools, Students, and Teachers Network.

With the data collected in Table 10 above, three analysis processes were identified:

- A bookings tracker showed progress towards SSAT's school recruitment targets.
- A live dashboard was automatically updated, which showed the number of expressions of interest, the type of interest that had been expressed, and what action had been taken by SSAT to progress that booking.
- Ad hoc periodic analysis was conducted on reach and recruitment data, depending on the needs of the programme (e.g. to see which schools had expressed interest in an area where an open day is scheduled).

Three systems were identified to ensure that analysis was acted upon:

- Targets for school recruitment were recorded in SSAT's budget, keeping them front and centre of staff minds.
- The bookings tracker was reviewed on a weekly basis by SSAT staff.
- The live dashboard was reviewed on a daily basis by SSAT staff.

Recruitment data—strengths

The current system had several strengths.

- It ensured that the most important school data for the purposes of recruitment were collected.
- The data was captured in a form that allowed it to be aggregated for quantitative analysis. Some minimal and helpful qualitative data was also collected to record meetings and progress of recruitment.
- There was a good data collection and storage system. This system ensured that data was collected at different stages of the pipeline, and ensured that the data was well managed (e.g. there is a single record per school, with linked records for staff within schools). The system was also flexible, with fields easily added to the CRM system if needed.
- Data collection was timely. For example, key contact information was collected as soon as a school joined the pipeline, and updates to data and progress notes were made at relevant stages of the pipeline.

Recruitment data—areas for improvement

Four possible areas for improvement were identified:

- There was some evidence to suggest that the current CRM data relating to school characteristics (e.g. Ofsted rating, FSM %, etc.) was incomplete. It was therefore suggested that SSAT collect complete data on key characteristics for each school. To save time for SSAT, this could be done through a mass upload and matching process using DfE admin data. This would allow SSAT to analyse what types of schools are being reached, and not reached. In time, SSAT may wish to target schools with specific characteristics (e.g. those that benefit more than other schools) and this data would support that goal.
- It is suggested that SSAT should start collecting regular feedback from schools on: i. reasons for sign up (e.g. a quick web survey or question in the first EFA mentor session); ii. reasons for staying (e.g. an annual web survey or question in the end of year reflection session); iii. reasons for not signing up (e.g. a quick web survey); and iv. reasons for leaving (e.g. a quick web survey and/or exit interview). This would allow SSAT to better understand schools' motivations for participating or not, which, in turn, could help SSAT to tailor their processes accordingly.
- SSAT may wish to consider conducting periodic aggregated analysis of reach and recruitment data (including the proposed new data suggested above). Ideally, this would be done as part of a planned development process for the recruitment activities (e.g. an annual review). This would allow SSAT to identify trends that are not easy to spot through more live analysis, potentially providing additional insights to support improvements or at least prompt further investigation.
- It is recommended that SSAT develop a formal M&E plan for reach and recruitment. This may include: i. targets for reach and recruitment (e.g. targets for specific school types, beyond those specified by funders like the EEF); ii. the questions that SSAT would like to answer about reach and recruitment; iii. the data that needs to be collected to answer those questions; iv. the required analysis of that data; v. when that analysis should be completed; vi. who is responsible for completing the analysis; and vii. what could be done with the findings (e.g. whether any results trigger certain actions or feed into an annual development process etc.). This would allow SSAT to record current good practice, making the organisation more resilient to staff changes and making it easier to induct new staff.[8]

---

[8] As part of a separate piece of work funded by the DfE and the EEF, BIT has since worked with SSAT to create a comprehensive M&E plan.

Implementation data

Table 11 outlines the data currently being collected that relates to implementation of the schools programme.

Table 11: Data currently collected relating to implementation of the schools programme

Topic	Data collected	How collected?	When collected?
Timing	• Start date in school • End date in school	• Mentor conversation record Word document • Mentor conversation record Word document	• Mentor Session 1 • End-of-programme review session
Completion of key activities	• TLCs done? • Length of TLCs correct? • Peer observations done? • Wrap-around support done?	• Implementation survey completed by mentor (hard copy and webform options)	• End of years 1 and 2
Completion of 'impact monitoring'	• Impact monitoring carried out by school?	• Implementation survey completed by mentor (hard copy and webform options)	• End of years 1 and 2

TLCs, Teacher Learning Communities.

For the data collected in Table 11, one analysis process was identified. In this process, the programme lead reads each conversation record and each implementation survey that is submitted and qualitatively reviews the contents. This review is done without a formal framework.

There were no formal systems or processes identified to ensure that the data collected and analysed above was acted upon. The new version of the 'EFA Programme QA [Quality Assurance] Process' notes that: 'a low score on the implementation survey may trigger a QA call at the end of Y1. However, no further written plan was identified for this survey.

Implementation data—strengths

Three strengths were identified in the current systems and processes.

• While the decision for mentors to complete the implementation survey on behalf of the school could be seen as a compromise in data quality (because it is second hand information), it has the benefit of a very high completion rate (100% in previous years), which is unlikely to be matched if schools were given the responsibility.
• The implementation survey is mostly quantitative with clearly defined categories for scores. This makes it easy to analyse large quantities of data (though no aggregated analysis is conducted at present).
• The implementation survey had one qualitative question for gathering extra key information. This generated useful findings, but also minimised the amount of qualitative data to be reviewed, saving a lot of analysis time.

Implementation data—areas for improvement

Six possible areas for improvement were identified.

• SSAT could seek to ensure that all mentors use the webform for the implementation survey, and to remove the paper option. This would save SSAT considerable time in data processing and analysis, which will be increasingly important as the programme continues to scale.
• SSAT may wish to provide mentors with the implementation survey at the start of the year in their induction pack. This would reinforce the core components of the programme for mentors (as these are specified clearly in the implementation survey). Observations of launch events suggested that some mentors did not have a firm grasp of the intervention's essential features, and gaining familiarity with the survey may help to address this.

- It is recommended that SSAT ensure that all web surveys feed into the CRM. This may require using a third party survey app that is supported by the CRM software. This would allow SSAT to manage all programme information in one system, which would save staff time by avoiding duplication, and could also reduce analysis time (see next suggestion).
- With all web surveys linked to the CRM system, it is recommended that the CRM is used to: i. create auto-dashboards for the programme team to review results as they come in; ii. create report templates (that automatically produce key statistics and charts) for periodic reviews; and iii. create trigger emails for key programme staff when a survey/conversation record comes in and/or when a survey value is below a threshold (i.e. triggering a phone call). This would allow the programme team to see survey results as they come in (without having to read the raw data), easily analyse aggregate survey data to support programme reviews, and more efficiently react to responses that require a quick follow-up.
- SSAT should consider creating a formal M&E plan for implementation—see the section on 'Reach and recruitment data' above, for further suggestions on what this plan might include.
- It is recommended that SSAT begin collecting the following extra data on implementation: i. the proportion of teachers participating in EFA at each school; ii. feedback on adaptations, and barriers and enablers to implementation; and iii. data on implementation after the 2-year programme is complete. This would help SSAT to check whether the programme is being implemented at the whole-school level (as intended), provide insights that might support broader programme improvements, and build an understanding of whether the programme is likely to have the lasting impact in a school that is intended.

Quality data

Table 12 outlines the data currently collected by SSAT relating to programme quality.

Table 12: Data currently collected relating to programme quality

Topic	Data collected	How collected?	When collected?
Launch event quality	• Teacher perceptions of quality	• Launch event survey (hard copy and webform options)	• At end of launch event
Wider programme quality	• School lead perceptions of quality of programme	• End-of-programme evaluation survey completed by school lead (webform)	• End of year 2
	• Mentor perceptions of programme quality	• Mentor conversation records (Word document)	• After every mentor session
Mentor quality	• School lead perceptions of quality of mentor	• End-of-programme evaluation survey completed by school lead (webform)	• End of year 2
	• Head of programme's perception of mentor quality pre-placement	• Head of programme QA call notes at end of year	• End of years 1 and 2
		• Not clear	• During training

QA, quality assurance.

The head of programmes reviews the data outlined in Table 12 on an individual basis. This includes a review of all launch event evaluation forms, the first mentor conversation records for all schools, the complete set of conversation records for a school for more than 10% of schools in a yearly cohort, and all end-of-programme evaluation forms.

There was evidence of an analysis plan for the launch event evaluations data and the mentor conversation records. This plan also included recommended actions if the analysis raises any concerns. There was no record of an analysis plan for the end-of-programme evaluation surveys.

Quality data—strengths

Four strengths were identified for the current system and processes.

- Some data is captured in a form that allows it to be aggregated for quantitative analysis. This makes it simpler for SSAT staff to generate answers to questions they may have.
- A significant amount of qualitative data is collected on progress and quality of experience in each school. This provides rich information to support improvements (though is relatively time-consuming to analyse).
- Asking mentors to keep structured conversation records provides rich data on quality and implementation, but also helps to ensure that sessions happen, and that they are delivered as intended.
- The 'EFA Programme QA Process' document recorded the purpose of some key data collection and analysis, as well as some actions that should be taken when data raises concerns. Having this written plan helped to ensure that the QA process was followed and made the process resilient to changes in staff.

Quality data—improvement areas

Eight possible areas for improvement were identified.

- It is recommended that SSAT should formally interview mentors as part of the recruitment process, assessing them against the desired qualities listed in the QA framework, and taking notes from these interviews. Formalising the recruitment process will help to better screen and select potential candidates, raising the quality of the programme.
- SSAT may wish to use a semi-structured observation guide to observe the quality of potential mentors during their training phase. Discussions with SSAT staff suggested that observation of mentors during their training was important, and so formalising this process would help SSAT to structure and record the observations. In turn, this may help to better screen and select high-quality mentors.
- SSAT may wish to create a webform, linked to the CRM, for mentors to complete their conversation records. This would allow SSAT to collect, link, and review this data more efficiently. It would also allow some automated analysis of quantitative data collected through the form.
- SSAT may wish to create a webform for the launch event survey, ideally linked to the CRM, and remove the paper option. Using an online platform would allow SSAT to post a simple code on PowerPoint slides that participants could enter into their smartphone and complete the survey immediately with very little friction. This would help SSAT to collect, link, and review this data more efficiently. It would also allow some automated analysis of quantitative data collected through the form.
- It is recommended that SSAT add some multiple choice questions to the end-of-programme survey for school leads on the quality of core components of SSAT's support, for example, it would be helpful to gather feedback on the quality of: i. teaching resources; ii. facilitation resources for TLC leaders; iii. planning resources for school leads; and iv. the structure and content of mentoring support. This would help to provide aggregate data on the quality of different elements of the programme and to identify trends that might support programme development. The largely qualitative form currently used by SSAT does not allow for easy aggregation, and individually reviewing qualitative feedback will become harder as the programme scales.
- SSAT may wish to formally observe new mentors delivering launch events. This would allow SSAT to monitor the quality of launch events (a crucial aspect of SSAT's support for schools) and to check new mentors' understanding of the overall programme. This could also be set up in a way that can be easily scaled, for example, with experienced mentors conducting observations rather than the programme lead being responsible for all observations.
- SSAT should consider conducting periodic aggregated analysis of the launch event and end-of-programme evaluations. This would help to identify trends that are not easy to spot through reviews of individual survey responses, potentially providing additional insights to support programme improvements or prompt further investigation.
- While SSAT does have some basic documentation outlining analysis plans for quality-related data, this should be developed into a more comprehensive M&E plan. See the section on 'Reach and recruitment data' above, for further suggestions on what this plan might include.

Programme results data

Table 13: Data currently collected relating to programme results

Topic	Data collected	How collected?	When collected?
Impact on school	• School lead perception of impact on school	• End-of-programme evaluation survey completed by school lead (webform)	• End of year 2
Impact on teachers	• School lead perception of impact on teachers • Mentor perception of quality of school's impact monitoring	• End-of-programme evaluation survey completed by school lead (webform) • Implementation survey completed by mentor (hard copy and webform options)	• End of year 2 • End of years 1 and 2
Impact on students	• School lead perception of impact on students • Mentor perception of quality of school's impact monitoring	• End-of-programme evaluation survey completed by school lead (webform) • Implementation survey completed by mentor (hard copy and webform options)	• End of year 2 • End of years 1 and 2

Table 13 outlines the data currently collected by SSAT relating to the programme's results.

Two analysis processes were identified for the data in Table 13. Both the end-of-programme survey and the implementation survey are qualitatively reviewed by the programme lead. This review is done without a formal framework.

There were no formal systems or processes identified to ensure that the data collected and analysed (Table 13) was acted upon. There was no record of an analysis plan for the end-of-programme evaluation surveys. The new version of the 'EFA Programme QA Process' notes that: 'a low score on the implementation survey may trigger a QA call at the end of Y1. However, no further written analysis plan was identified for this survey.

Strengths—programme results data

Two strengths of the current approach were identified.

- EFA has already been subject to the EEF rigorous effectiveness trial with positive results (though this study does not rule out a null result with 95% confidence).
- This strong evidence of impact is combined with good monitoring of fidelity and quality, including a qualitative assessment of whether the school has its own good M&E processes in place, and the school lead's assessment of impact in their school. Schools use strategies to monitor indicators of impact, such as learning walks, lesson observations, student feedback, and teacher feedback. If the programme is implemented as intended, which is monitored on an ongoing basis, then it should achieve the intended outcomes. This system is likely the best possible system available. For example, it is very difficult to track outcomes for teachers well (this data would have to be self-reported, survey-based, and would likely have a low response rate). Student outcomes could be tracked (e.g. through Progress 8 scores), but this would be a statistically noisy measure and therefore not a useful indicator of programme impact. For these reasons, SSAT's current systems were considered to be a proportionate and efficient approach to impact management.

Mentor programme data

Table 14 outlines the data currently being collected by SSAT relating to the mentor programme.

Table 14: Data currently collected relating to the mentor programme

Topic	Data collected	How collected?	When collected?
Contact details	• Name • School • Telephone • Email	• Direct from mentor, stored on CRM	• During recruitment
Programme details	• Which schools they are supporting • Capacity for academic year • Where they can work • When they completed training	• From project manager during introduction conversations, stored in a spreadsheet • From trainer register, stored in spreadsheet	• During recruitment

CRM, Customer Relationship Management.

The research found that no substantial analysis was being conducted on the mentor data currently being collected. No systems were identified to ensure that the data was acted upon.

While there was no identified analysis of the data, the data was found to have two particular uses. First, mentors' contact details were stored on the CRM system so that mentor records could be linked to school records. Second, basic information on the work capacity of mentors was collected and reviewed to support the matching of mentors with schools.

It is recommended that SSAT start collecting and analysing data on mentor reach and recruitment, much as they do for schools. For example, SSAT should consider collecting the following data from mentors: reasons for signing up; reasons for staying; reasons for not signing up; and reasons for leaving. This would allow SSAT to monitor and evaluate mentor recruitment practices, helping to improve them over time, and leading to a stronger pipeline of high-quality mentors to support the programme.

As well as data on reach and recruitment, SSAT may wish to start collecting and analysing data on mentor training and support. In particular, SSAT may wish to collect implementation data (e.g. activities completed, start dates, and end dates), quality data (e.g. feedback from mentors, and observations of trainings) and outcomes data for mentors (e.g. self-reported learning, and assessments to test knowledge and skills). Collecting this data would allow SSAT to monitor and evaluate the mentor training and support, helping to improve it over time.

Summary of findings

Table 15: Summary of improvements to schools programme

	Reach and recruitment	Implementation	Quality	Results
What data is collected?				
How and when is it collected?				
How is it stored?				
How is it analysed?				

What systems and processes are in place to ensure that the analysis is acted upon?				

Notes: Green coloured cells indicate 'no improvements suggested'; Blue coloured cells indicate 'improvements suggested'.

The data currently collected by SSAT includes a mixture of quantitative and qualitative measures. The quantitative data currently captured is typically collected through online forms, which makes aggregation straightforward, and reduces processing and analysis time. Most quantitative surveys also included a limited number of qualitative questions, allowing for the capture of some additional feedback, while not placing excess burden on the reviewer. There was also some evidence of analysis plans for some surveys, with some suggested actions where certain criteria were met.

A range of suggested improvements were made for the school's programme. Table 15 summarises the findings relating to the school's programme's M&E data processes, noting where improvements were suggested.

In particular, no improvements were suggested for SSAT's existing approach to monitoring the results associated with the programme. The existing system was perceived to be a proportionate and efficient approach to impact management. A number of consistent recommendations were made across recruitment, implementation, and quality: i. ensure complete data is collected where possible; ii. collect qualitative and quantitative feedback from stakeholders; iii. conduct periodic analysis of the data collected; and iv. develop formal M&E plans relating to each area.

Table 16 summarises the findings relating to the mentor programme's M&E data processes, noting where improvements were suggested.

The research found that SSAT collected some limited data on mentors, but that no substantial analysis was conducted on the data currently collected. No systems were identified to ensure that the data collected was acted upon. The central suggested improvement was for SSAT to begin collecting and analysing data on mentor reach and recruitment, and on mentor training and support.

Table 16: Summary of improvements suggested for mentor programme

	Reach and recruitment	Implementation	Quality	Results
What data is collected?				
How and when is it collected?				
How is it stored?				
How is it analysed?				
What systems and processes are in place to ensure that the analysis is acted upon?				

Notes: Blue coloured cells indicate 'improvements suggested'; Pink coloured cells indicate 'nothing in place'.

Reach and recruitment

EFA sales process

Table 17: EFA sales pipeline

Phase	Data point
Awareness	Data point 1: Download a web document
	Data point 2: Fill out a form about EFA
EOI	Data point 3: Attend an open day
	Data point 4: Meeting with a member of the recruitment team
	Data point 5: Meeting with the programme lead (generally reserved for headteachers or someone who can actually make a decision about programmes)
Pre-commitment	Data point 6: Booking form sent to a school decision-maker
Started programme	Data point 7: School started delivery of the programme (i.e. delivering the first TLC)

EFA, Embedding Formative Assessment; EOI, expression of interest; TLC, Teacher Learning Community.

The EFA 'sales pipeline' consists of four phases. Table 17 provides a high-level overview of the pipeline, with the data points available to SSAT in each phase.

Changes to sales approach for scalability

One change to the sales approach was identified to support SSAT's efforts to scale. There was some evidence to suggest that they intended to make greater use of the ambassador school network (schools that have previously participated on the EFA programme). In particular, SSAT hoped to expand the number of prospective schools attending open days and engaging with school leads. It was hoped that this would allow schools to hear how other schools had implemented the programme in their context, and to hear some success stories associated with the programme. This, in turn, it was hoped would increase the number of schools expressing an interest in the programme and eventually signing up.

Recruitment outcomes

In total, 23 schools were recruited in the 2021/22 academic year, 14 of which began the programme in the 2021/22 academic year.[9] This represents approximately 4% of the 375 schools that interacted in at least one way with the sales pipeline (those that were 'reached'). At this conversion rate, if SSAT reached all c.4,000 secondary schools in England, only 160 of them would start the programme. This conversion rate therefore places a substantial restriction on the scaling potential of the programme.

Around 35% of the schools recruited this year had above-average levels of FSM. The majority of schools recruited (86%) were rated 'Good' by Ofsted. Around 64% of schools recruited were based in London or the south-east, with 21% being located in an 'Opportunity Area'. Tables 18 to 22 provide full breakdowns of the demographic characteristics of the schools recruited.

Table 18: Percentage of students eligible for FSM in schools reached and started

Percentage of students eligible for FSM	Schools reached N=375	Schools started N=14

---

[9] Four of these schools received funding from other sources (e.g. their local authority) to cover some of their programme fee.

<10	70 (18.7%)	4 (28.6%)
10–19	110 (29.3%)	4 (28.6%)
20–29	90 (24.0%)	3 (21.4%)
30–39	63 (16.8%)	2 (14.3%)
40–49	26 (6.9%)	0 (0.0%)
50–59	4 (1.1%)	0 (0.0%)
Missing	12 (3.2%)	1 (7.1%)

Notes: The average for schools nationally is 19.9%. 'Reached' means that a school interacted with at least one part of the recruitment pipeline described in Table 17. FSM, Free School Meals.

Table 19: Ofsted ratings for schools reached and started

Ofsted rating	Schools reached N=375	Schools started N=14
Outstanding	63 (16.8%)	3 (21.4%)
Good	181 (48.3%)	9 (64.3%)
Requires Improvement	38 (10.1%)	0 (0.0%)
Serious Weaknesses	3 (0.8%)	0 (0.0%)
Special Measures	5 (1.3%)	0 (0.0%)
Missing	85 (22.7%)	2 (14.3%)

Notes: The national average of schools not included in our sample that are either 'Outstanding' or 'Good' is 74.9%. 'Reached' means that a school interacted with at least one part of the recruitment pipeline. Ofsted, Office for Standards in Education, Children's Services and Skills.

Table 20: Geographical region for schools reached and started

Region	Schools reached N=375	Schools started N=14
East Midlands	23 (6.1%)	1 (7.1%)
East of England	32 (8.5%)	1 (7.1%)
London	73 (19.5%)	3 (21.4%)
North-east	19 (5.1%)	0 (0.0%)
North-west	55 (14.7%)	1 (7.1%)
South-east	39 (10.4%)	6 (42.9%)
South-west	21 (5.6%)	1 (7.1%)
West Midlands	71 (18.9%)	0 (0.0%)
Yorkshire and the Humber	42 (11.2%)	1 (7.1%)

Notes: The national average of schools not included in our sample that are either 'Outstanding' or 'Good' is 74.9%. 'Reached' means that a school interacted with at least one part of the recruitment pipeline.

Table 21: Size of schools reached and started

Number of pupils	Schools reached N=375	Schools started N=14
<200	2 (0.5%)	0 (0.0%)
201–400	14 (3.7%)	0 (0.0%)
401–600	26 (6.9%)	1 (7.1%)
601–800	45 (12.0%)	0 (0.0%)
801–1,000	69 (18.4%)	3 (21.4%)
1,001–1,200	69 (18.4%)	4 (28.6%)
1,201–1,400	47 (12.5%)	1 (7.1%)
1,401–1,600	51 (13.6%)	4 (28.6%)
1,601–1,800	27 (7.2%)	0 (0.0%)
1,801–2,000	10 (2.7%)	0 (0.0%)
>2,000	3 (0.8%)	0 (0.0%)
Missing	12 (3.2%)	1 (7.1%)

Notes: The national average of schools not included in our sample 967 pupils. 'Reached' means that a school interacted with at least one part of the recruitment pipeline.

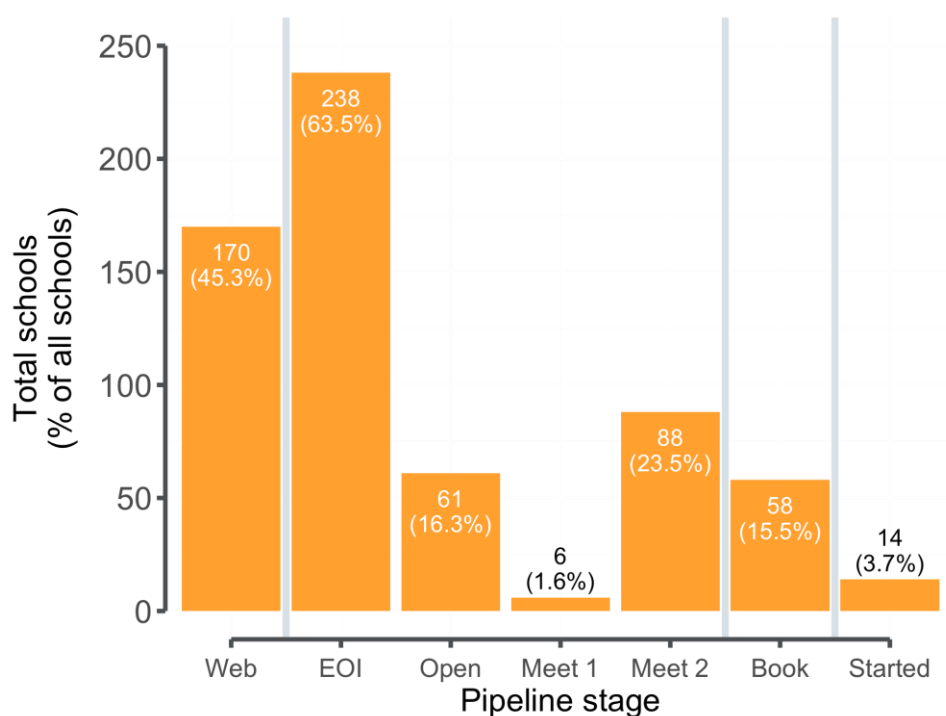*Table 22: Opportunity Area status for schools reached and started*

Opportunity Area	Schools reached N=375	Schools started N=14
Yes	24 (6.4%)	3 (21.4%)
No	350 (93.3%)	11 (78.6%)
Missing	1 (0.3%)	0 (0.0%)

Notes: The national average of schools not included in our sample that are in an 'Opportunity Area' is 4.6%. 'Reached' means that a school interacted with at least one part of the recruitment pipeline.

School response to sales process

In total, 375 schools entered the sales pipeline between January 2019 and June 2022, with 14 new schools starting the programme in the 2021/22 academic year. This was considerably lower than the 50 additional schools that SSAT were targeting for the 2021/22 academic year.

Figure 2 shows the number of schools that reached each stage of the sales pipeline in this time period.

Two mechanisms likely contributed to the low number of schools recruited to the programme. The first was the number of schools that entered the pipeline (n=170), which was insufficient. The second was the conversion rate to sign up for schools that entered the pipeline, which was 3.7% between January 2019 and June 2022. Table 23 summarises the factors that likely influenced each of these mechanisms, and the following sections contain additional details.

Figure 2: Number of schools reaching each stage of the sales pipeline.

Notes: N=375. This covers all secondary schools interacting with the pipeline between January 2019 and June 2022. Grey lines separate the four pipeline phases described in Table 16. EOI, expression of interest.

Table 23: Factors influencing the number of schools entering the sales pipeline and the conversion rate for schools in the pipeline

Factors affecting the number of schools entering the sales pipeline	Factors affecting the conversion rate for schools in the sales pipeline
• The school's prior exposure to 'trusted messengers' • The evidence base for EFA • Influence of school networks • Financial implications of the programme • Stage in the school's improvement journey • Functionality of the SSAT website	• Engagement with other EFA schools • Exposure to EFA resources • Alignment to existing school priorities • School's experience of engaging with SSAT • School leader's perception of the flexibility of the programme • School leader's concerns about potential impacts on teacher workload • Perceived need for support from SSAT • Possible loss of momentum for schools in the sales pipeline

EFA, Embedding Formative Assessment; SSAT, Schools, Students, and Teachers Network.

Factors affecting the number of schools entering the sales pipeline

Six factors were identified that influenced the number of schools entering the sales pipeline: i. the school's prior exposure to 'trusted messengers'; ii. the evidence base for EFA; iii. the influence of school networks; iv. the financial implications of the programme; v. the stage in the school's improvement journey; and vi. functionality of the SSAT website.

The school's prior exposure to 'trusted messengers'

Some school leaders noted that they were disproportionately influenced by the opinions and recommendations from other senior leaders, and particularly those that are working in contexts similar to theirs. Other school leaders explained

that 'education influencers' affected the programmes they looked into. If the programme was recommended by someone they followed on Twitter, or by a facilitator at a conference, they were more likely to look into it.

> *I'm quite influenced by the big teacher profiles that I follow on Twitter. If they recommend something I'll often then check it out.* (SLT)

When prompted, leaders also suggested that they would find marketing materials that highlighted a testimonial from another senior leader appealing.

> *I think testimonials are always good and do provide confidence when we are looking for programmes.* (Senior leader)

The evidence base for EFA

Some school leaders noted that they had been searching for an intervention that was tried and tested in a range of schools, and that had robust evidence of impact. Some school leaders had seen that the EEF had been involved in the previous research into EFA, and this gave them confidence in the strength and robustness of the evidence. School leaders suggested that amplifying the evidence base for EFA would likely increase the number of schools wanting to find out more about it.

> *It's like lemmings, if you hear that the programme worked in 140 schools then you're going to want to look into it as well.* (SLT)

The influence of school networks

Some school leaders noted that they had become aware of the EFA programme, and subsequently entered the sales pipeline, because the programme had been recommended to them through the multi-academy trust or local authority. In some cases, the fact that the recommendation came from a senior regional director made the school even more likely to then look into it.

The financial implications of the programme

In some cases, school leaders had been aware that they could access funding through their local authority for CPD, and that this subsidy first prompted them to research different CPD programmes. When they then came across the EFA programme, they noted that they were not phased by the cost of the programme because they knew they had this subsidy to use.

> *We were allocated a pot of money from the Local Authority for CPD, so that's what prompted me to actually look for an intervention to use.* (SLT)
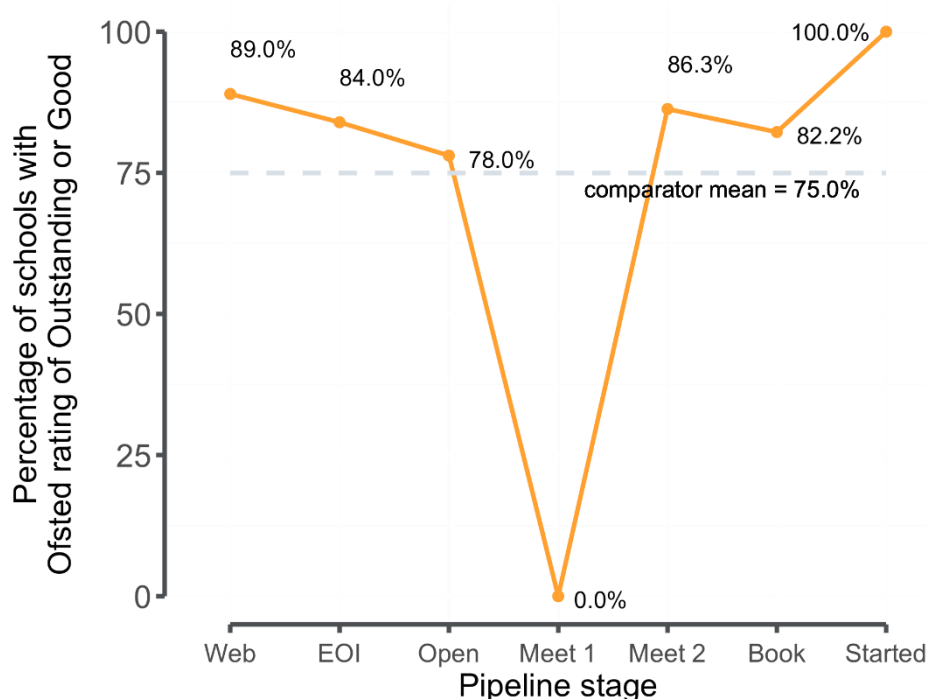
For schools that were not able to access such subsidies, there were some concerns around the cost of the programme. While leaders typically did not think that the cost of the programme was unreasonable, it was high enough that they felt they would have to pitch the programme to the rest of the SLT and generate broader support for the intervention. It was perceived that to do this effectively required a significant amount of time to research the programme. Some school leaders also noted the importance of being able to quickly find information about the cost of the programme, and that they had found it difficult to determine the full cost. It was inferred from this that other leaders may have been put off from finding out more because of this ambiguity.

> *The website said the programme cost £1.20 per student but I wanted to know the full cost of the programme to the school and that seemed hard to find. I think if the cost isn't 100% clear from the outset it might push some schools away.* (SLT)

The stage in the school's improvement journey

As noted in Figure 3, the percentage of schools in the pipeline that are 'Good' or 'Outstanding' is considerably higher than the national average.[10] This suggests that schools further on in their school improvement journey have the time and capacity to explore and research new programmes and have the capacity to take on a whole-school intervention like EFA. The implication is that EFA may be less attractive to schools that are classified as 'Inadequate' or 'Requires Improvement', despite being the schools that might benefit the most from such a programme. As shown in Figure 2 above, only six schools took part in the 'Meet 1' stage of the pipeline (an optional meeting with a member of the recruitment team). All six of these schools had 'Satisfactory' or 'Requires Improvement' Ofsted ratings. We do not have any data to explain this. Given the small numbers, it could have occurred by chance, and it may explain why none of the schools who opted for this meeting ended up participating in the programme (see Figure 4, below); i.e. the 0% conversion rate from this meeting could have been artefact of the schools' characteristics rather than the effectiveness of the meeting.

Functionality of the SSAT website

Some school leaders found the website to be difficult and cumbersome to use. In some instances, this had made it challenging to find the information needed to determine whether EFA might be appropriate for their school. Indeed, this may have contributed to a lower number of schools entering the sales pipeline than would have otherwise. However, it is important to note that this was not a universal view, with some leaders commenting that they had found the website to be well-signposted overall.

Figure 3: Proportion of schools that are rated 'Good' or 'Outstanding' by Ofsted at each stage of the sales pipeline. EOI, expression of interest; Ofsted, Office for Standards in Education, Children's Services and Skills.

Factors affecting the conversion rate for schools in the sales pipeline

Interviews with senior leaders identified eight broad factors that influenced the likelihood of a school in the sales pipeline converting to a full sign up: i. the school's engagement with other EFA schools; ii. the school's exposure to EFA resources; iii. the extent to which EFA aligned with existing school priorities; iv. the school's experience of engaging with SSAT; v. the school leader's perception of the flexibility of the programme; vi. the school leader's concerns about

---

[10] The Ofsted ratings presented are the latest received by the school by the end of the 2020/21 academic year (the most up-to-date data available at the time of reporting). The 'comparator mean' in Figure 3 is the percentage of the rest of the secondary schools in England (i.e. not those schools in the pipeline) that were 'Good' or 'Outstanding'.

potential impacts on teacher workload; vii. the perceived need for support from SSAT; and viii. a possible loss of momentum for schools in the sales pipeline.

The school's engagement with other EFA schools

School leaders noted that they were influenced considerably by their engagement with other EFA schools. They found it particularly helpful to hear from schools that had already completed the programme, and to hear how they had tackled various implementation challenges. Some leaders also noted that they had seen other teachers discussing the EFA programme on social media (primarily on Twitter), and that after seeing other teachers writing positively about the programme they were more likely to sign up.

In general, leaders noted that they wished to have the opportunity to interact with schools currently implementing the programme. Indeed, this is the purpose of the open days organised by SSAT. Unfortunately, none of the interviewees had attended one of these and so were unable to comment on the quality or usefulness of the event, but the leaders did perceive these to be a helpful offer from SSAT. However, due to Covid, many of the open days organised by SSAT had to be delayed or cancelled.

The school's exposure to EFA resources

Some leaders noted that they had had the opportunity to review some of the TLC content before making the decision to sign up for the programme. Often this opportunity had arisen because another school in their multi-academy trust was using the programme and were willing to share the resources. This opportunity helped confirm that the programme materials were high quality and reduced the perceived risk of proceeding. In one instance, a school leader had been able to test out the EFA content with a group of staff in their school to ensure that it had a positive reception.

The extent to which EFA aligned with existing school priorities

Some school leaders noted that their decision to sign up for the programme depended on the extent to which they saw the programme as aligning with the school's broader priorities. Some also commented that they were seeking a professional development programme that would help to embed previously taught CPD. For example, some school leaders thought that the programme would directly build on techniques from 'Teach Like a Champion' (Lemov, 2021) and instructional coaching. These were areas that the schools had already focused on in the past, and the leaders liked the fact that EFA would continue to develop these concepts within their staff. It should be noted that it was up to individual school leaders to consider how the programme might align with other previous CPD priorities. Where the alignment was less obvious, some school leaders noted that the programme felt quite different to anything their school had undertaken before, which led to concerns about the reaction to the programme from staff, making them less likely to sign up.

The school's experience of engaging with SSAT

Some school leaders spoke about how calls with members of the SSAT team were particularly influential on their decisions to sign up. Leaders spoke about the usefulness of being able to ask practical questions about implementation, and to hear answers from people that were highly familiar with the programme. Some leaders noted that conversations with the programme lead at SSAT were particularly impactful, and that they had appreciated the opportunity to discuss the challenges they were foreseeing around implementing the programme in their context, and to brainstorm a range of solutions. This is also reflected in Figure 4, where it can be seen that the conversion rate for schools that had a call with the programme lead (labelled 'Meet 2') is higher than for other sales activities. However, it should be noted that data from the sales pipeline suggests that some interactions with SSAT may be less effective—in particular, as seen in Figure 4, the conversion rate from 'Meet 1' was 0% (with the caveat about the composition of this group noted in the discussion of Figure 3 above).

> *To be honest, the biggest influence on deciding to sign up was a call I had with [the programme lead]—she was knowledgeable about the programme and could answer all of my questions and after that chat I was fully convinced.* (SLT)
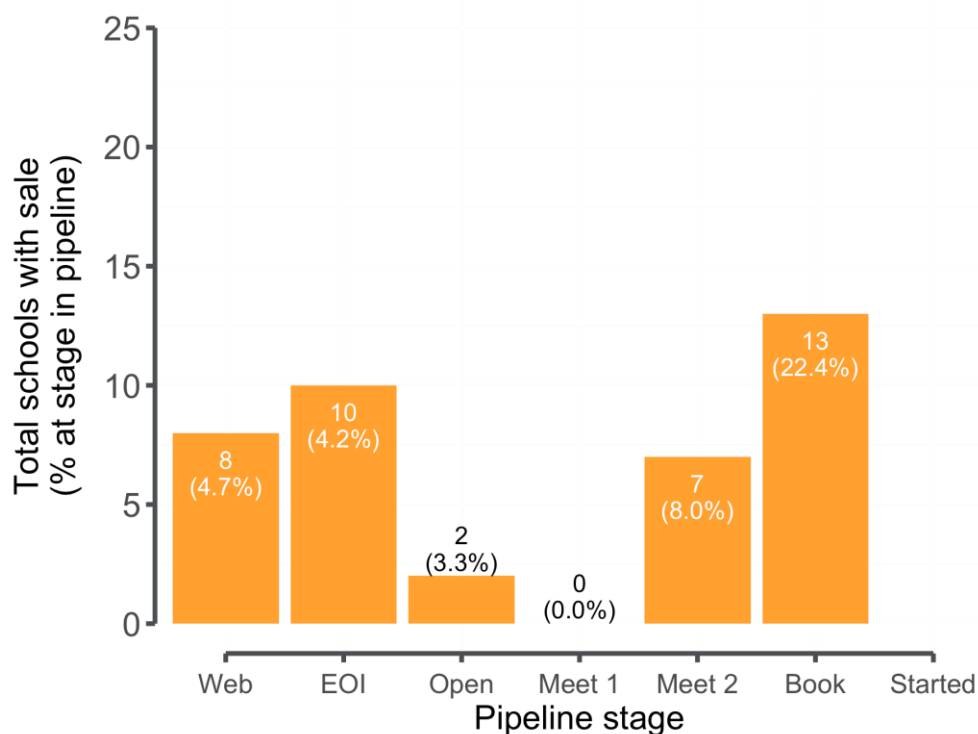
Figure 4: Number of schools signing up for the programme from each phase of the sales pipeline. EOI, expression of interest.

Observations of the open days organised by SSAT showed them to be informative and useful to the schools attending, and also provided a helpful forum for schools to ask questions that they had about the programme. It also provided an efficient way for a lot of schools to hear from the programme lead, which, in theory, may reduce the demand for individual calls with the programme lead at a later date. However, only 16% of schools in the pipeline attended an open day. It is important to note that many open days had to be delayed or cancelled entirely due to Covid, which helps explain the relatively low proportion of schools that did attend an open day. However, some school leaders were unaware that any open days were planned, and so SSAT may wish to explore ways to raise the profile of these events.

While there was considerable praise for SSAT's support throughout the sales process from senior leaders, some leaders did comment that they found the communication approach to be cumbersome and inefficient. These leaders felt that they received too many emails from SSAT, and that these emails were often very long and that it was unclear what action was required. Taken together, this put them off engaging with the communications which, in turn, made them less likely to sign up for the programme.

The school leader's perception of the flexibility of the programme

Some schools were wary to commit to a 2-year programme (which was considered to be substantially longer than other CPD programmes). Some leaders also had concerns about waning teacher engagement with the programme and thought that the repetition of 18 TLCs over 2 years could lead to staff losing interest.

The school leader's concerns about potential impacts on teacher workload

Some school leaders noted that the requirement for TLCs to last for 75 minutes would lead to an increase in teacher working hours, and that this requirement made them less inclined to sign up for the programme.

> *In our school we only have 60 minutes allocated to CPD, so the 75 minute requirement would have meant we were going over teacher working hours. We're in a really unionised borough so that would have caused some major issues.* (SLT)

Other leaders commented that they were most concerned about the potential rise in workload for the teachers working as TLCs. With many schools still recovering from the impact of Covid-related school closures, there was significant concern about burdening teachers with additional responsibilities. It appeared that in some cases schools liked the programme and wanted to participate but felt that they would be able to engage with it better at a more stable time.

The perceived need for support from SSAT

Some schools noted that they had considered whether they needed to pay for the programme (and actually sign up), and whether it may have been possible to implement the core components of the programme themselves without being part of the formal programme. Some leaders discussed their reluctance to sign up for the programme stemmed from the fact that they wanted to implement a well-evidenced intervention, but that they wanted to retain flexibility and autonomy to implement the programme in a way that suited their school.

A possible loss of momentum for schools in the sales pipeline

As shown in Table 24, the median time from entering the pipeline to a sale was 440 days (~14 months). While it does likely take a considerable amount of time for schools to make a final decision about a CPD programme and to actually procure the programme, this long time period could lead to a loss of momentum within the school (as well as potential changes in leadership/decision-makers and other circumstances in the school that may affect the decision). Research from behavioural science suggests that people respond to deadlines, and that setting shorter deadlines could be an effective way to maintain a sense of momentum and urgency (Ariely, 2002).

Table 24: Time from entering pipeline to sale (N=14 schools)

Time	Days
Mean	395
Minimum	0
25th percentile	203
Median	440
75th percentile	582
Maximum	777

Given that just 22% of schools that received the booking form went on to start the programme in the 2021/22 academic year, it would also appear that there is room for improvement in terms of the frequency and style of reminder messages being sent to schools.

Fidelity, contextual factors, and sustainability

Essential features of intervention

The essential features of the intervention, as perceived by those involved with programme delivery and implementation, varied considerably depending on the role of the individual being asked. SSAT programme staff generally considered implementation to involve three key activities: teachers participating in TLCs and reflecting on their practice; teachers observing one another between TLCs; and producing personal action plans to support them as they try new techniques in their classroom.

Table 25: Summary of essential features of intervention

Summary of findings
1. SSAT programme staff perceived the programme to involve three key activities: TLCs; peer observations; and personal action plans 2. School staff perceived TLCs and peer observations to be central to the programme, but not personal action plans

SSAT, Schools, Students, and Teachers Network; TLCs, Teacher Learning Communities.

In general, all school-based participants—the school lead, the TLC leaders, and classroom teachers—considered TLCs and peer observations to be central to the programme. In particular, school staff consistently commented that an essential feature of the programme was the way that it facilitated dialogue and reflection between staff cross-departmentally. Teachers noted that schools, and secondary schools in particular, can be highly siloed, with staff interacting regularly with those in their departments and hardly at all with teachers outside their department.

> *Central [to the programme] is collaboration with colleagues that you might not otherwise be collaborating with; cross-departmental, cross-curricula. For me that is the strongest element of the programme.* (Classroom teacher)

No data was collected from school-based participants on how exactly this cross-departmental collaboration served the aims of the programme. However, one hypothesis is that it may be linked to teachers seeing real-life examples of formative assessment strategies being trialled in a wide variety of subject contexts. For example, some teachers noted that prior to the programme they felt that some strategies only worked in certain subjects. However, the cross-departmental nature of the programme challenged this assumption and showed teachers the creative ways that the techniques were being applied across subjects.

Beyond the implementation of the TLCs, school-based participants typically perceived the peer observations to also be a core part of the intervention. Around 91% of endline survey respondents reported organising peer observations as part of their TLC sessions. However, there was some uncertainty around the recommended frequency of observations, with 20% of participants unsure of how frequently they should be observed. This uncertainty was also reflected in interviews, with some classroom teachers perceiving the EFA programme to largely involve attending an after-school TLC session periodically.

> *The main thing is attending the TLCs when they happen. There's not a huge amount to do between the sessions.* (Classroom teacher)

This observation was particularly pronounced in schools that had weaker leadership of the programme, and where the mechanisms of the intervention were not fully understood by the school lead. It was also observed in schools where the school lead did not have the capacity to follow-up with individual teachers about not completing the other important features of the intervention, such as peer observations and personal action plans (see the section on 'Monitoring intervention fidelity' below, for further details).

Personal action plans, considered a core feature of the intervention by SSAT, were not perceived to be an essential feature of the intervention by some school-based participants. Some staff did not see the value in the approach, while others did but felt that they did not need to complete it as it would not be checked by the school lead.

> *I think we're supposed to complete an action plan, but it's not normally checked by anyone so you don't necessarily have to do it.* (Classroom teacher)

This inconsistent approach to personal action planning was reflected in data from the endline staff survey. While 94% of respondents reported having ever completed a personal action plan, only 66% of respondents thought that a personal action plan was required after every TLC session (as expected by SSAT).

The length of the TLC sessions was also identified as a core feature of the intervention in interviews with some school leads. It appeared that the emphasis on TLC sessions needing to be 75 minutes had been understood by these leads. However, this requirement was adjusted in some instances though—see the section on 'Appropriate adaptations' below, for further information.

> *We have definitely stuck to the programme. Our sessions are always 75 minutes, and so I think fidelity to the programme has been good.* (TLC leader)

Overall, there was some confusion and uncertainty from school staff about what the most important elements of the programme are, and which elements of the programme are adaptable. School staff were largely aware that the programme had a strong evidence base, but this had led to some schools feeling uncertain about changing elements of the programme for fear of removing an aspect of the intervention that made it effective. Some of these schools noted that they would benefit from further explicit guidance that made clear the aspects of the programme that should: i. never

be changed; ii. be changed with good reasons and with consent from their mentor; and iii. be adapted to suit the needs of their school.

> *I have a fear…because we're using the evidence base to say, 'If this material is delivered in this way, it has this effect'. So I very clearly don't want to rock the boat. But I would like more guidance on what is allowed.* (School lead)

Adaptation of intervention

This section is broken down into three parts. The first part considers adaptations to the programme that were made by schools and were considered by SSAT to be appropriate and/or suitable for scaling the intervention. The second part considers the process for schools to agree modifications to the programme. The third and final part then explores a range of factors that affected whether schools were able to make effective adaptations. Table 26 summaries the findings presented in each of these three parts.

Table 26: Summary of findings relating to adaptation of intervention

Appropriate adaptations	Process for agreeing modifications	Factors influencing effective adaptation
- Changes to programme language (e.g. not using the word 'observation') - Adjustments to TLC session length - Increasing the size of TLC groups - Adjustments to TLC session content (e.g. changing formative assessment strategies recommended) - Use of technology to automate programme elements (e.g. personal action planning)	- Typically, the school lead suggests adaptations to the EFA mentor (either in regular check-in or by email/call) - Where necessary, the EFA mentor escalates the request for modification to the programme lead - Some school leads felt empowered to make minor modifications without escalation to the EFA mentor	- Opportunities for regular feedback to the school lead about ongoing implementation - Quality of relationship between the EFA mentor and the school lead - Responsiveness of the EFA mentor to the school lead's emails - Proactive engagement from the EFA mentor to the school lead - Expertise and experience of the EFA mentor - EFA mentor's perception of the school's fidelity to the programme

EFA, Embedding Formative Assessment; TLC, Teacher Learning Community.

Appropriate adaptations

Some schools adapted the language of the programme to make it more familiar and amenable to their staff. In particular, some schools chose to not use the language of 'peer observations' and opted for replacements such as 'peer drop-ins' or 'peer feedback'.

> *We have peer coaching visits here. We call them peer coaching visits because it fitted with our ethos better than 'observation'.* (TLC leader)

In many cases, this was due to a school's historical use of high-stakes observations to monitor teacher performance. Some teachers explained that this historical use of observations had led to stress and anxiety among staff, and that this led to considerable wariness when the EFA programme was first being introduced.

> *When they first started talking about observations as part of the EFA programme, I thought 'here we go again', and was thinking it was another way to monitor what we are doing in our classrooms.* (Classroom teacher)

In instances where the school's leadership were aware of this anxiety, they were able to address these concerns head on at the start of the programme. Some schools leads noted that they chose to explicitly tell staff that they were changing

the language of the programme away from 'observation', and to provide the rationale for this, as a way of addressing staff concerns.

> *I knew our staff would have a problem with the term 'observation' given how it's been used in quite a negative way to monitor teacher performance in the past. So I decided to tell staff that we were changing the language to 'peer drop-in' and emphasised that it was a low-stakes activity which was purely about teacher development, and that the outcomes of it would not be fed back to SLT.* (School lead)

As noted in the section on 'Essential features of intervention' above, school leads were aware of the requirement for TLC sessions to be 75 minutes. However, there was some evidence of schools making adaptations to the structure of the sessions to make the timings fit with their existing school timetable. Indeed, endline survey data showed that only 34% of responding schools had TLCs that lasted for 75 minutes, with a majority of schools (56%) reporting their TLCs lasting for an hour or less. This may be because the school opted to split the minutes across multiple sessions (see case study in Box 1, below), or it could have been that some schools spent a shorter time on TLCs than SSAT expected them to.

Box 1. Case study: Adjusting the delivery structure of the TLC sessions to fit within directed hours

In one case study school, the school's leadership and the EFA school lead decided that it was not feasible to complete a 75-minute training session after-school hours, as this would not fit with the school's policies on working hours. In light of this, the school lead made the following adaptation to the delivery format:

- Teachers attended a 15-minute morning meeting in their TLC groups, in which the TLC leaders introduced the formative assessment techniques that were the focus of that week's TLC. Typically, this introduction focused on the theory and concepts behind the techniques and teachers were sometimes given a small reflection task to consider
- Teachers then attended a 60-minute after-school meeting in their TLC groups on the same day. This followed the usual structure of a TLC, involving reflection and action planning

This model was perceived to be highly effective by teachers at the school. In particular, teachers noted that it provided them with time to reflect on the theory throughout the day and to even implement the strategy with one of their classes that day. Teachers commented that this meant that the after-school discussion in the TLCs was then richer as a result.

This delivery model—with the TLC being split into two parts—was considered risky by SSAT, as it added additional organisation and monitoring for the school lead. However, the EFA mentor also considered the adaptation to have been a success largely because of the talent and skill of the school lead.

EFA, Embedding Formative Assessment; SSAT, Schools, Students, and Teachers Network; TLC, Teacher Learning Community.

Another adaptation observed in some schools was increasing the size of the TLC groups beyond the level prescribed by SSAT. This was usually in response to high levels of staff absence driven by Covid (see the section on 'Impact of Covid' below, for further information) or a high number of part-time staff, and the school lead's desire to ensure a minimum number of staff attending each TLC to ensure a richness and diversity of conversation.

Some schools also made small adjustments to the content of the TLC sessions. The most minor type of content adaptations involved small tweaks to the presentation and form of the session content, usually incorporating the school's branding into the PowerPoint slides, if possible. This was perceived to improve buy-in from staff, as it made it appear that the programme had been tailored and made relevant to the school and was not just an off-the-shelf package from an external organisation.

Beyond minor adjustments to the presentation of content, some schools also adapted the content itself. This was usually done by the school lead to ensure that the content was relevant to the teaching staff and pupils in the school. For example, one school changed references in the strategies to ones that their pupils would be more aware of.

> *One of the strategies is 'call a friend', which is a 'Who Wants to be a Millionaire?' reference. We realised that none of our kids watch that programme and don't really understand it. So we changed it to 'DM your friend'.* (TLC leaders)

Another school noted that they adjusted the language of the formative assessment strategies to mirror and align with previous CPD initiatives that the school had pursued. This was perceived to be important as it meant the formative assessment strategies did not feel completely new to teachers. This, in turn, helped to retain staff buy-in to the programme (see the section on 'Alignment to other CPD priorities' below, for further information).

> *It doesn't matter how great something is, if you lose the staff buy-in, then you've lost it. All of these changes have been to ensure that staff are supported, that they feel like they're listened to…so that the buy-in remains high.* (School lead)

Some schools reported that the strategies included in the TLC sessions felt somewhat outdated, and not fully reflective of current trends and findings in education research. Some of these schools opted to make more substantial adaptations to the strategies, and this was often informed by the school lead's personal research and interest.

> *I'd said to her [EFA mentor], 'Some of these materials seem a bit archaic', and she was like, 'Well, if they're not appropriate then think about what you can do that fits properly but links to what you do', and I talked to her at length about TLAC [Teach Like a Champion[11]] and Rosenshine [Principles of Instruction[12]]. She was like, 'You could drop that in, you could leave that out', etc., and I shared a couple with her to make sure she was okay and she was like, 'Spot on, that's fine, but you don't mess with the protocol'.* (School lead)

When observed, these updated strategies were appropriate and engaging for teaching staff. However, it is important to note that the success of this adaptation depended substantially on the school lead's ongoing engagement with education research and their ability to apply it. It is recommended that SSAT explore updating the strategies to reduce the burden placed on school leads.

Some schools made adaptations to the process for managing teachers' personal action planning, often by making use of technology. These schools noted that a purely paper-based system (as recommended by SSAT) was time-consuming, inefficient, and made it harder to access. In one case study school, teachers completed the action plan as an online form. This meant that the school lead could rapidly ascertain, which teachers had completed their action plans, helping their monitoring of fidelity. It also meant that the school lead could then send a reminder message to teachers telling them what their target was for that cycle. This likely increased the likelihood of teachers trying out their desired strategy in the classroom.

Importantly, for this adaptation, the school lead considered the risks of the change and took steps to mitigate these. For example, there was a concern that by using technology, it may lead to staff browsing their laptops during TLC sessions and not being fully present in the discussion. The school lead stressed to staff that the default in all sessions was for laptops to be put away, and that they should only be used at very specific points in the session (e.g. during the action planning section). The success of many adaptations seemed to depend on the ability of the school lead to consider the risks and to take steps to mitigate these.

Process for agreeing modifications

The process for agreeing modifications varied considerably, often based on the intensity of the proposed modification and on school-level factors. For more substantial changes to the programme content, many school leads reported asking their mentor whether they could make the change.

School leads tended to be more likely to consult their mentor before making a modification where they reported having a positive relationship with their mentor (the importance of the school lead–mentor relationship is also discussed in the sections on 'Factors influencing effective adaptation of the intervention' below and 'SSAT factors' below). Some school leads also noted that they discussed potential modifications to the programme in the occasional check-ins that the mentor instigated with them, highlighting the importance of the mentor actively contacting school leads.

> *A few times my mentor just dropped me an email to see how things were going and I had actually been thinking about making a change, so I just replied and asked what they thought about it. I probably wouldn't have got in touch out of the blue, it was just because they had already emailed me.* (School lead)

---

[11] Lemov (2021).
[12] Rosenshine (2012).

Mentors varied in their willingness to permit changes to the programme. Some mentors took a relatively hard-line approach and encouraged schools to stick as closely to the original programme as possible, while others were considerably more lenient, allowing more substantial changes. There were some suggestions that newer, less experienced mentors tended to be stricter in their implementation, possibly based on a shallower understanding of the core mechanisms of the programme. In contrast, more experienced mentors, and particularly school-based mentors, were perceived to be more willing to permit changes to the programme. This was perhaps due to their deeper experience of implementing the programme within schools, which gave them a deeper appreciation for the elements of the programme that can be altered if necessary.

Some mentors noted that they often escalated requests for modifications to the SSAT programme lead. In general, these mentors tended to be less confident and less experienced with the programme, and therefore needed additional support to respond to such requests.

In some schools, the school lead felt empowered to make minor modifications to the programme without escalation to their mentor. This was particularly likely in schools where the school lead had a strong understanding of the core mechanisms of the EFA programme, which gave them the confidence to make small changes as they knew these would not interfere with the key elements of the programme. This improved understanding of the core mechanisms was often developed through an interest from the school lead (e.g. they had read the EEF evaluation of the programme) or was developed through previous interactions with their mentor.

> *My mentor spent quite a bit of time with me at the beginning of the programme explaining the key bits of EFA, the bits that we really can't change, so I knew what they were. Beyond that, my mentor encouraged me to make any small changes I needed to make it work in our school.* (School lead)

Factors influencing effective adaptation of the intervention

A range of factors that influenced the effectiveness of adaptations to the intervention were identified in the research.

The first—and potentially most significant—factor was whether there were regular opportunities for teaching staff and TLC leaders to provide feedback to the school lead about the ongoing implementation of the programme. In schools where there were mechanisms in place for the school lead to receive feedback on the programme in real-time, effective and useful adaptations could be made that improved staff buy-in and overall intervention fidelity.

> *With our staff, and with COVID, and with all of the huge pressure that's been on the profession, I needed to balance and make sure that I was sticking with the programme because it works, and the research shows that it works, but doing it in a way that was supportive of the staff and maintained the buy-in...All of these changes informed by feedback from staff have been to ensure that staff are supported, that they feel like they're listened to.* (School lead)

In one school, the school lead collected feedback on the implementation of the programme after the first three TLCs. The feedback was collected through an anonymous Google Form and revealed that some staff were struggling to find time within the 4-week period to complete the peer observation. This prompted the school lead to consider a range of strategies and adaptations to the programme that could make this process easier, eventually allowing teachers a full half-term period to organise their peer observations.

In another instance, the school lead sought feedback on the content of the TLCs before delivering them to the whole-staff community. This allowed the school lead to identify issues with the language or content which, in turn, allowed them to make sensible adaptations ahead of time.

In another case study school, the school lead met with the TLC leaders immediately after each TLC to gather their feedback on how it went, and to generate suggestions for appropriate adaptations to improve the delivery of the sessions in the future.

> *I meet them after the TLCs normally and they'll just...listen to them and they just said, don't like this, do like this, can we change that, can I change this...I just encourage them to be honest because the whole point about implementation is that it's an iterative process, isn't it?...Nobody implements anything as is, do they?* (School lead)

While some of these feedback-generating mechanisms were more formal (e.g. a written Google Form for all staff to complete), many of them were informal, relying on casual conversations between the school lead and the TLC leaders. This required a positive relationship to exist between the school lead and the TLC leaders, and also for the school lead to be approachable and open to feedback.

Beyond opportunities for feedback, the quality of the relationship between the EFA mentor and the school lead was also considered to be an important factor affecting a school's ability to make effective adaptations to the programme. Where the school lead perceived the mentor to be available and accessible, they were more likely to ask questions about possible modifications. This meant that the mentor was able to provide support to ensure that any deviations were appropriate, or to push back on deviations that the mentor thought the school should not be making.

> *I'm normally in really close contact with the school leads. Sometimes schools will say, 'oh, I don't particularly want to do that for a particular session'. That's where I would say we really don't recommend that, because the programme was written…to follow [a] particular order, not jump around from one thing to the next…By and large, the schools that I've worked with have said, 'okay, fine, we can now see that rationale'.* (Mentor)

In particular, the speed with which the mentor replied to messages was noted by school leads as being particularly important. Given that school leads are often busy and time-stretched members of the SLT, timely responses to queries helped them to incorporate the mentor's advice and feedback into any adaptations that they were making.

Some school leads also appreciated their mentor actively offering support, rather than only responding to questions. This positive engagement from the mentor made the school lead feel like they were not burdening the mentor with questions, and made the school lead more likely to ask for support. In turn, this support helped them adapt and change the programme to better suit the needs of the school.

> *She's always making offers to support me—'Do you want a Teams meeting?', 'Do you have any questions?' I sometimes think it's quite a British thing to not ask for help when you need it, but the fact she offers her support so readily makes me know it's there so when I do need it I can ask for it.* (School lead)

However, in some schools the limited capacity of the school lead made it difficult for them to build a positive relationship with their mentor, meaning that they were unable to fully take advantage of the support they could offer.

As well as the relationship between the mentor and the school lead, the quality of the mentor themselves was considered to be another important factor. In particular, where mentors had a strong grasp of the essential features of the programme, and understood how to balance intervention fidelity with flexibility, they were able to provide good advice to schools that helped them make effective adaptations. School-based mentors appeared to be particularly good at this, with their experience of delivering the programme giving them an insight into adaptations that had worked for them in the past. Importantly, these mentors saw the programme as necessarily flexible, rather than rigidly prescribed.

Even where the mentor understood the need for flexibility, some schools had significant concerns about not wanting to change an evidence-based programme. In some instances, this prevented the school from making appropriate adaptations that could have increased its effectiveness.

The school's fidelity to the programme, as perceived by their mentor, was also considered to be a factor influencing their ability to make effective adaptations. Where mentors saw their schools as implementing the programme with low fidelity, they were less willing to permit adaptations. This was largely due to the mentor having concerns that making adaptations might add to the workload of the school lead, which could reduce fidelity even further.

> *Sometimes a school that isn't implementing the programme as well will ask me if they can make some change. In these cases, I'm usually a bit less likely to say yes and will try and push them to just stick to the programme. Often making changes to the programme requires time and effort from the school lead, and if they can't give it that time then those changes will be implemented badly and that could worsen the overall implementation.* (Mentor)

As noted in the section on 'Appropriate adaptations' above, some schools found novel and efficient ways to incorporate technology into the intervention (e.g. online forms for action planning with ongoing reminders sent to staff about their

plans). The ability for schools to make effective technology-related adaptations largely depended on the technical ability of the school lead.

> *I've had quite a bit of experience building simple web tools for our school so it was very easy for me to build the system to track teacher's personal action plans online.* (School lead)

Some schools wanted to incorporate technology further into the programme but found themselves unable to do so, often due to concerns about privacy and data protection. In one case, a school wanted to record the TLCs to allow staff that had missed the session to catch up on the content, but they had decided it was not possible for data protection reasons.

Box 2. Case study: Aligning EFA to previous whole-school CPD efforts

In one case study school, the school lead made a series of adaptations to the language and content of the TLC sessions. The school had prioritised other CPD initiatives in recent years, with a particular focus on techniques from Lemov's (2021) 'Teach Like a Champion' and Rosenshine's (2012) 'Principles of Instruction'. The school lead was keen to ensure that any techniques or strategies introduced as part of the EFA programme aligned to the techniques that teachers had already learnt. In this way, it was hoped that the EFA techniques would reinforce the previous CPD.

With this in mind, the school lead reviewed all of the TLC content and changed the language of the PowerPoints and strategies to match the equivalent strategies in Lemov's (2021) 'Teach Like a Champion' or in Rosenshine's (2012) 'Principles of Instruction'. She also produced a document that explained to TLC leaders how the content of each TLC linked up with previous CPD in the school, so that the TLC leaders could make those connections explicit in the sessions.

Feedback from staff within the school was positive about these adaptations. Staff appreciated that the strategies being introduced felt familiar and welcomed the reminder of the importance of them. Some staff also noted how in the education sector it can sometimes feel like new ideas and techniques are being pushed onto teachers all the time, which makes it hard for teachers to embed the techniques over time. Staff liked that the same techniques were being recommended over time within the school, with staff suggesting that they felt more positively towards the programme because they were not being overloaded with new concepts and strategies.

> *[The school lead] made sure that the programme matched up with everything that we've been working on as a school in the past few years. This made EFA feel like a natural continuation of all of that rather than something brand new…That definitely made people feel more positively about the programme and helped with overall engagement.* (TLC leader)

EFA, Embedding Formative Assessment; CPD, Continuing Professional Development; TLC, Teacher Learning Community.

Encouraging intervention fidelity

This section outlines the factors identified that encouraged intervention fidelity. The factors are broken down into SSAT factors and school-level factors. Table 27 summarises the factors identified.

Table 27: Summary of factors that encouraged intervention fidelity

SSAT factors	School-level factors

• Quality of school lead's relationship with the EFA mentor
• Proactive engagement from the EFA mentor to the school lead
	• Supporting TLC leaders
	  o Reducing burden on TLC leaders
	  o Assigning a deputy TLC leader to each TLC group
	• Whole-school encouragement
	  o References to EFA embedded into staff communications
	  o Technology-based reminders
	  o Competition between TLC groups
	• Performance management processes
	  o Linking engagement with EFA to teacher performance reviews and progression
	• Building staff buy-in

	o Positioning EFA as response to staff feedback o Piecemeal introduction of EFA programme components o Developing staff intrinsic motivation for formative assessment

EFA, Embedding Formative Assessment; TLC, Teacher Learning Community.

SSAT factors

The quality and frequency of the school lead's relationship with their mentor was identified as being an important factor for encouraging intervention fidelity.

Where schools leads held strong relationships with their mentors, they were also more likely to seek support with issues within the school, and to seek advice on how to encourage intervention fidelity. For example, some school leads noted that they contacted their mentor to explain that certain teachers were struggling to engage with certain elements of the programme, and the mentor was able to provide some advice about how to handle the situation. School leads noted that they were more likely to proactively reach out in this way to their mentor when they felt their mentor would be non-judgemental, and when their mentor had made them feel like they did not mind being bothered with small requests such as this (further information on the importance of the school lead–mentor relationship can be found in the sections on 'Process for agreeing modifications' above and 'Factors influencing effective adaptation of the intervention' above).

> *My mentor also made clear that I could email anytime with any concern—small or big—and they would be able to help me out. So I did that a few times!* (School lead)

School-level factors: Supporting TLC leaders

A wide range of school-level measures were identified to encourage intervention fidelity. These included actions taken before the programme began, as well as measures that were taken while the programme was being implemented.

First, some schools took measures to reduce the burden placed on TLC leaders. For example, some schools allocated TLC leaders with protected time to complete their tasks. This meant that they had the time to adequately follow-up with staff in their groups, to complete all paperwork, and to prepare for TLC sessions. This was also reflected in the endline survey data, with 93% of TLC leaders noting that they felt supported by their school to engage with the EFA programme.

> *We told TLC leaders that they could have some time off in lieu in the summer term because of all of the extra workload that the EFA programme might give them. This definitely made TLC leaders feel better about doing all the tasks required... .* (School lead)

TLC leaders across all case study schools commented on the high workload involved with the role, with some suggesting that they did not have enough time to complete all the activities to a high-enough standard. Some TLC leaders suggested that this may have reduced overall intervention fidelity.

> *One of the parts of the role that I find most overwhelming is the volume of admin there is to do after the sessions. Sometimes it's hard to scan all the action plans and [email] people who weren't there. Obviously, those things are really important for the programme but sometimes I just don't have time.* (TLC leaders)

Beyond assigning additional time to TLC leaders, some schools took other measures to support TLC leaders to do their role, which in turn encouraged intervention fidelity. Some school leads explained that they sought to make the role of the TLC leaders as easy as possible. This was usually done by taking some of the administrative tasks off them, for example by managing all of the printing required for the TLC sessions.

Another measure taken by some schools to support TLC leaders was to assign every TLC group with a deputy TLC leader. This adaptation to the programme meant that every group had two individuals responsible for the delivery of the content, as well as the associated administration. Where this was done, TLC leaders typically felt less overwhelmed by the workload and had the time and capacity to engage positively with the programme.

School-level factors: Whole-school measures and encouragement

Beyond support for TLC leaders, the schools with the highest fidelity to the programme sought to regularly remind staff about the programme in between TLC sessions. For example, some schools ensured that references to EFA, and the specific strategies being worked on in that cycle, were mentioned in whole-staff briefings on a weekly basis. Other schools referred to the programme in staff bulletins or on staff notice boards. Classroom teachers in these schools reported finding the regular reminders helpful, noting that it reminded them to try and embed techniques into their lessons, and to also schedule their peer observations.

> *We just try and make sure that everyone is talking and thinking about the programme all the time. There's lots about EFA on the noticeboard in the staff room, and we're also putting it in the Teaching & Learning newsletter which is pinned up in every faculty office. So it [the EFA programme] is just very present when you're walking around the school. Even if teachers don't read it all, there's that constant reminder that you should be doing things.* (School lead)

As noted in the section on 'Appropriate adaptations' above, some schools used technology in innovative ways to remind staff about the programme between TLC sessions. This allowed some schools to automatically send reminders to staff to encourage them to try the technique they had committed to trialling.

Some schools also attempted to use competition between TLC groups to encourage intervention fidelity. In one school, staff were encouraged to pin a note on a whole-school noticeboard once they had completed their peer observation for that cycle. The TLC group that had the highest number of notes stuck on the board were then rewarded.

> *We got all of our staff to write a little post-it note after their peer observation saying what they had learnt and to stick it on these notice boards. Then the TLC group with the most posts on it would get cake at the next TLC session. It was the TLC leaders that really drove this because they wanted the cake!* (School lead)

As noted in the quote above, it appeared that it was the TLC leaders (rather than the school's leadership) that drove engagement with the competition. The TLC leaders themselves were keen to win the competition, and so took it upon themselves to follow-up with staff, encouraging them to complete their EFA activities. From conversations with the school's leadership, this was in keeping with the culture of the school, which was perceived to be competitive and ambitious.

> *We're a very competitive school. When I was a TLC lead, there were big fun rivalries between the TLC leaders to see who could win the cakes. It meant that we had the TLC leads going round encouraging people and saying to members of their team 'Have you done your peer coaching visit? What have you been practising recently? Make sure you get it done'. That kind of cajoling and encouragement is really important. It moves the programme from a whole-school thing to a more personal thing.* (School lead)

This type of healthy competition was perceived to have been highly effective by classroom teachers in the school. They commented that it made the whole process more fun, and they felt that they did not want to let their group down by not doing their peer observation.

Some schools also thought carefully about how to make the TLC sessions themselves as pleasant an experience as possible for the staff attending them. In doing so, it was hoped that staff would be more likely to attend and engage positively in the sessions. Some schools thought very deliberately about composing the TLC groups to balance personalities, and to ensure that the group would be positive and fun.

Beyond group composition, some schools also provided TLC groups with free food, drinks, and snacks. This was perceived to make the sessions feel less formal, and to encourage discussion. It also made the sessions more attractive for staff members.

> *If you think about how a TLC is, it's not like an information giving thing, it's not a meeting, it's not minuted. It's a group of like-minded professionals sharing practice and moving forward in their teaching together. It's a relatively social occasion and tea, coffee, and cake helps that. It lubricates the conversation, doesn't it?* (School lead)

School-level factors: Performance management processes

Some of the schools with the highest fidelity to the programme took more formal measures to encourage intervention fidelity. For example, some schools chose to include teachers' engagement with the EFA programme in the school's official performance management system. In this way, teachers felt that they had to engage with the programme, otherwise they would risk a poor performance review, which could have an impact on their progression and salary increases. Some of the schools that took these measures stressed that they were only doing so to mitigate some of the problems that Covid had caused for the programme. In these cases, there were concerns that lockdowns and school closures had led to a drop in momentum for the programme (see the section on 'Impact of Covid' below, for more information), and that making it part of the performance management process was a way to signal to teachers that it was still important and should be a priority for them. This active approach meant that staff got back on track with the programme very quickly after returning to in-person teaching.

However, some respondents felt that including EFA in performance management processes could lead to negative outcomes. Some staff suggested that such strong accountability measures could crowd out some of the intrinsic motivation that staff had, and that some staff might begin to view it as a 'tick-box' exercise rather than really engaging with the ethos of the programme.

> *If it comes into performance management, and staff are held formally accountable for it, does it begin to feel a bit forced? Teachers' attitudes towards it will change because they're now doing something because they're being told they have to, rather than because they want to do and believe in it. There is already so much compliance-focused things in education. I think it is important to make teachers feel trusted and treat them like professionals.* (Mentor)

School-level factors: Building staff buy-in

Lastly, some schools sought to encourage intervention fidelity by trying to build staff buy-in to the programme. These schools recognised that fidelity to the programme, and its success in terms of changing teacher behaviour, rested on having engagement from staff.

For example, some schools sought feedback from staff on the teaching and learning support that they wanted, and the specific structure that this support should take. In these instances, staff suggested that they wanted more time to work with their peers, more time for coaching, and wanted more lower-stake opportunities to observe one another. The school's leadership then identified EFA as being an appropriate programme that could meet these objectives and could then pitch the programme as being a direct response to suggestions and feedback from staff. In this way, staff felt that they had played a part in EFA being the programme that was selected, which increased their buy-in to the programme, and their willingness to engage with it thereafter.

> *We asked staff what they wanted from a CPD programme…It meant that I could then go in and say 'Right, here's what you asked for. Here's what the research says. This is what we're going to do'. It was really exciting because when we launched it, the buy-in was brilliant, people were so excited, there was a great buzz. And that engagement has then lasted throughout the programme.* (School lead)

Some schools also took measures to introduce the various elements of the programme slowly, to allow staff time to embed them. For example, one school encouraged staff to begin doing peer observations, and to start meeting in TLCs before the EFA programme actually began so that they would get used to working in this way. Then when these concepts were introduced as part of the EFA programme, staff felt reassured because it was things they were already doing in their day-to-day practice. In turn, it was perceived that this helped to generate staff buy-in to the programme, which encouraged them to engage positively with the programme thereafter.

Some schools also took additional steps, beyond what is recommended by the programme, to build staff understanding of the mechanisms and processes that help to make EFA effective. By doing so, it was hoped that teachers' understanding of the programme would increase, along with their intrinsic motivation to engage with it. In turn, it was thought that this would translate into higher fidelity. In one school, this involved the delivery of a bespoke launch event with guest speakers (see case study Box 3 below). In other schools, this involved the school lead spending more time early on in the programme discussing and explaining the rationale and theory behind the programme. Measures such as these depended heavily on the commitment and enthusiasm of the school lead to the programme (see the section

on 'EFA school lead' below for further discussion on the importance of the school lead for the programme's implementation and adoption).

Box 3. Case study: Developing a bespoke launch event to gain buy-in from staff

One case study school developed a bespoke launch event to introduce the programme to all of the staff within the school. The event was a full day and took place on an INSET day before the start of the academic year.

The day began with an introduction from the school's headteacher who described how the programme would support the school's ongoing priorities and how it aligned with the principles and values of the school. For the school in question, the EFA programme was a careful and considered choice. The timing had been thoughtfully planned as the school had already been working to ensure the foundations for the programme were in place. Staff spoke very positively of the introduction given to the programme by the headteacher, noting that it was clear from the very beginning that the school's leadership were firmly behind the programme, and believed wholeheartedly that it would be effective in their context.

The headteacher's address was followed by two keynote speeches from thought leaders in the education sector. The first speech discussed the need for formative assessment and provided a summary of much of Dylan William's research in the area. The second speech came from an author and educator who also has a substantial following on social media, and he discussed how formative assessment plays an important role within the following four areas: how the curriculum is portrayed; how students' thinking is exposed; how to motivate through achievement; and how to support positive behaviour. Both speakers blended research and their own experiences in the classroom to generate an engaging and exciting talk. After this, the two speakers did a Q&A session with staff about what had been discussed, which led to some engaging debates between staff.

Following the whole-staff element, teaching staff then went into their TLC groups and completed the first TLC of the year. Classroom teachers noted that the sessions with the external speakers had convinced them that formative assessment was an appropriate focus for the coming 2 years, and that they were excited to get started. In this way, the engagement in the very first TLC was reported to have been particularly high. Overall, staff spoke very positively about the training, noting that the external speakers gave the event a feeling of importance, and signalled the investment the school was making in the EFA programme.

> *The launch event was amazing, probably the best CPD we've ever had. It gave the whole programme a big bang of a start.* (Classroom teacher)

> *Imagine if we didn't have that launch and we just had to tell staff what they're going to be doing. It would have been really hard for us alone in small groups to build up the momentum. Having all the teachers in the theatre listening to people, getting inspired and then doing it, I think proved really, really effective.* (TLC leader)

CPD, Continuing Professional Development; EFA, Embedding Formative Assessment; INSET, in-service training day; Q&A, Question and Answer; TLC, Teacher Learning Community.

Monitoring intervention fidelity

Fidelity to the intervention was monitored by SSAT (through the EFA mentor), the school lead, and the TLC leaders. The ways in which fidelity was monitored by these actors are summarised in Table 28.

Table 28: Summary of findings on how fidelity was monitored

SSAT and the EFA mentor	School lead	TLC leaders

• Virtual informal check-ins with the school lead
• In-person school visit at end of first year of implementation, including:
  o Observations of TLC sessions
  o Meeting with TLC leaders and classroom teachers
  o Lesson observations
• Completion of Fidelity Survey
	• Recording teacher attendance at TLC sessions
	• Observations of TLC sessions
	• Recording teacher completions of peer observations and personal action plans
		• Recording teacher attendance at TLC sessions
		• In some instances, recording teacher completions of peer observations and personal action plans

TLC, Teacher Learning Community.

Monitoring intervention fidelity: SSAT and the EFA mentor

The EFA mentor monitored intervention fidelity in two main ways. First, mentors typically held a series of informal virtual check-ins with the school lead throughout the year. These were organised on an ad hoc basis, and often in response to queries from the school lead, or the mentor reaching out. Some mentors incorporated a RAG (Red-Amber-Green) rating system into these check-ins to allow the school lead to self-assess the school's fidelity to the programme up to that point. This was perceived to help the mentor build an understanding of the school's implementation, and to identify areas that the school may need additional support in going forward.

> *[With the school lead], we went through the action plan and she RAG-rated the things that were in that plan. That then guided the discussion, so she could say, so these things are all green and going well, but these things here are amber/red and need more work. This helped me get a sense of how implementation was going in the school and also helped prioritise my support going forward.* (Mentor)

Mentors suggested that they were able to judge fidelity to the programme (to an extent) through these check-ins. The school lead's self-assessment provided some insight into how the programme was going, but some mentors also noted that they could pick up on overall intervention fidelity based on the confidence and knowledge that the school lead demonstrated in the sessions. These mentors were typically more experienced and had worked with a wide range of school leads in the past.

> *I've worked with quite a few schools now, and you can kind of just tell through your conversations with the lead how the programme is going. It's not 100% all of the time, but I can get a pretty good indication based on how confidently they talk about the programme, how well they can identify challenges and so on.* (Mentor)

However, mentors did recognise the limit of these virtual calls as a method of monitoring fidelity and noted that they visited every school at least once as well, usually at the end of the first year. As such, in-person visits were the second main method for monitoring intervention fidelity by mentors. Mentors noted that the visit provided an opportunity for the mentor to observe TLC sessions, to meet with the TLC leaders, to meet with classroom teachers, and to observe a range of lessons. In doing so, they were able to determine whether the programme was being implemented appropriately, and whether the key elements of the programme were in place.

> *The end of year visit is a chance to have a much meatier conversation, where I can actually see some evidence on the various bits and pieces of the programme. So on the day I'll be asking 'Can you show me an example of this?', 'Can I see a TLC?', 'Can I see a lesson and see some EFA strategies actually being used?'* (Mentor)

Following this visit, and at the end of the first year of implementation, the mentor completed the 'Fidelity Survey', in which they rated the school on various aspects of their implementation. This report was then submitted to SSAT and the programme lead for review. The aggregate outputs from the Fidelity Survey can be found in Table 29 below. Data was only collected by mentors for half of the new schools by the end of the 2021/22 academic year. The small sample size limits the value of interpreting the percentage figures.

Table 29: Output from Fidelity Survey

Question	Overall (N=7)	
	Raw responses	Proportion
TLCs are meeting approximately once per month (mostly every 3–5 weeks) over the course of the year?		
No	0	0%
Yes	7	100%
TLCs held are approx. 75 minutes		
No	1	14.30%
Yes	6	85.70%
Peer observations are taking place regularly (not necessarily every month)		
No	0	0%
Yes	7	100%
Personal action plans are being completed after each session		
No	1	14.30%
Yes	6	85.70%
The school is monitoring the impact of the project effectively (score 0–3)		
0	0	0%
1 = Minimal monitoring, some feedback collected	2	28.60%
2 = A variety of monitoring strategies are used and feedback is shared	1	14.30%
3 = Impact is measured in a wide variety of ways, e.g. through learning walks, lesson observations, student and teacher feedback, which is shared and acted on to maximise the impact of the programme	4	57.10%
The school has fully committed to the project providing wrap-around support (score 0–3)		
0	0	0%
1 = Minimal support in place, which has not been maintained over time or infrequent	1	14.30%
2 = Staff have been given regular support in between meetings for peer observations and practice shared through other meetings/briefings/communications	2	28.60%

3 = Staff are supported beyond TLC meetings, with support/time to complete peer observations. The programme has a high profile with staff and students. There is regular input, e.g. through briefings, meetings, newsletters, celebrations, and sharing practices	4	57.10%

TLCs, Teacher Learning Communities.

This preliminary data shows that schools appear to be implementing the core components of the programme (75-minute long TLCs, peer observations, and personal action planning). However, there appears to be more variation in the extent to which schools are monitoring the impact of the programme, and the extent to which schools are committed to the programme.

Some mentors noted that the visit to the school reveals crucial information about their fidelity to the programme, many of which cannot be ascertained from the virtual calls with the school lead that they had been having. SSAT's scaling strategy budgeted for one visit to each school in the first year of implementation. Some mentors queried this, noting that an additional visit (some recommended after the first three TLCs) would be particularly helpful. It was thought that this additional visit would allow the mentor to identify concerns sooner rather than later and help to course-correct if they needed to.

Monitoring intervention fidelity: School lead

School leads monitored intervention fidelity in a range of ways.

First, some school leads noted that they consistently monitored teachers' attendance at the TLC sessions. This was usually done through a formal register taken by the TLC leader. This was then passed on to the school lead who kept an ongoing record of teacher's attendance.

Teachers appeared to be aware of this, with 57% of endline survey respondents reporting that their attendance at TLCs was monitored. However, 41% of respondents did not know whether their attendance was monitored. This suggests that in a significant proportion of schools it was not made explicit that attendance is being monitored, or that attendance was simply not being monitored.

As well as monitoring teacher attendance at sessions, some school leads also noted that they sought to monitor teachers' engagement and participation within the sessions as well. Some school leads had deliberately not allocated themselves to a TLC group so that they would be able to move between the groups, and check that everyone was on task and that discussions were engaged and fruitful. This also provided a way for the school lead to monitor the TLC leader's delivery of the session content, and helped them identify TLC leaders that needed additional support.

Some of the most effective school leads relied heavily on their TLC leaders to monitor fidelity, and to then report any issues to them. In general, this approach reduced workload for the school lead, and distributed some responsibility for the programme to other members of staff.

Some school leads also had systems in place to monitor whether staff were completing their personal action plans on a regular basis. In particular, some school leads noted that they had shifted the action planning process online, which made it very straightforward to monitor whether staff had completed their plans.

> *Everything in our school is online, so I made EFA go onto Google Drive too. Everyone has to upload their action plan into the relevant folder, and then I can go in and check that they're all there.* (School lead)

Some school leads also sought to monitor the completion of peer observations. For example, in one case study school, the school lead had asked staff to book cover in order to go and observe their partner. The school lead then met the cover supervisor to check who had booked in cover and who had not. Interestingly, this had proved to be a poor data source, as most teachers had not formally requested cover to complete the observation and had instead relied on members of the department covering them for short periods of time.

Some school leads also attempted to monitor outcomes associated with the programme, though this was typically more challenging to do. In some cases, school leads explained how they tried to track whether the actual classroom practices

of teachers were changing as a result of the programme, and how students in the school were responding to this. In general, schools with higher fidelity to the programme had a more active approach to monitoring such outcomes, and recognised the importance of ensuring regular teacher practice was being altered, in addition to the core elements of the programme being followed.

> *We monitor what's then changing in the classroom through our quality assurance schedule of learning walks. Our focus is checking for understanding, ensuring learning intentions are clear. Are the students receiving feedback that will move them forward? Do we see paired work? Do we see self-assessment? So looking for the five pillars within what we're seeing.* (School lead)

However, some schools' efforts to monitor outcomes were perceived to be ineffective, and a potential distraction for the school lead. Some classroom teachers suggested that their schools were putting considerable resources and effort into attempting to monitor outcomes, and that this came at the expense of not monitoring fidelity in other ways. In particular, it was suggested that some schools were hoping to have evidence of impact within a few months and tried to engineer ways to show this. For example, some schools made teachers take pre- and post-assessments that tested their understanding of various formative assessment strategies, before and after the TLCs. The school lead then attempted to demonstrate the improvement in staff understanding as an outcome of the programme. Some respondents from the school noted that this measure relied on self-judgement and was likely an inflated measure of progress. Overall, these attempts to prove impact in a short period of time were less apparent in schools where the school lead and SLT fully understood the mechanisms of the programme, and prioritised implementation of the programme's core features, knowing that the benefits would accrue over a longer period of time.

While some school leads took a more active approach to monitoring fidelity, other school leads decided to take a more hands-off approach. This was typically done for one of three reasons. First, some school leads described a school culture that resisted monitoring teacher behaviour and performance. For example, some schools had had a historical high-stakes culture in which SLT closely monitored staff, often through regular observations, which had led to resentment and distrust from staff (see the sections on 'Appropriate adaptations' above and 'School culture' below, for more information). To keep staff bought into the programme some school leads therefore thought it necessary to be more relaxed in their tracking of fidelity. However, schools with a culture that made it difficult to monitor fidelity typically had lower fidelity to the programme overall.

Other school leads took a less active approach to monitoring fidelity because of a strongly-held belief that staff should be empowered and trusted, and that their engagement with the programme should come from their own intrinsic motivation rather than from fear of external monitoring.

> *I am very reticent to…really hard monitor what's going on with the programme. I think that this is supposed to be a high-trust model, where people are working with each other and not being watched…As a leader…part of me squirms with that because…how can I be sure everyone it [sic] doing it? But I've got to be ok with that.* (School lead)

The third reason that often led to a less active approach to monitoring fidelity was the impact of Covid (see the section on 'Impact of Covid' below, for more information). School leads noted that Covid led to high levels of staff absence, which meant lots of staff missing TLCs for legitimate reasons. It also made it very challenging for staff to conduct their peer observations, as there were usually a small number of opportunities within the timetable for staff to do their observations with their partners, and if one of them was off work that day, it was hard to rearrange the observation for another time.

Reasons such as these made it harder for staff to fully engage with the programme, and some school leads therefore decided to take a less active approach to monitoring. In general, this was perceived to have maintained staff buy-in to the programme but did lead to a loss of fidelity.

> *Because of Covid we gave a bit of leeway to staff because we knew how hard it was to engage with the programme with everything else going on…So I think they're still bought into the programme, but I definitely don't think implementation during that Covid period was as great as it could have been. (*School lead)

While a less active approach to monitoring fidelity was welcomed by some staff in some schools, this was not a universal opinion. In fact, some staff in schools where the monitoring of fidelity was weaker felt that they were being asked to do

lots of work for the EFA programme but were not receiving any credit or recognition for doing it. Feedback from staff at these schools suggested that they did not particularly mind their compliance with the programme being checked, provided that it meant their efforts would be recognised. Monitoring of fidelity by the school's leadership could therefore be a way of celebrating staff engagement with the programme, helping to build goodwill and commitment to it in the longer term.

> *I have struggled with engagement on my team. They're not not doing the activities, but they're just not invested, maybe. Some people have voiced concerns that they're putting all this effort in, doing all these things, but there's no check-in to see is it having the desired effect, or what is the mood around it with the pupils, or they feel a little bit like they've been left to do this thing with us, and there's no accountability [chuckles] of whether it's actually working. (*TLC leaders)

Monitoring intervention fidelity: TLC leaders

The extent to which TLC leaders monitored fidelity to the programme varied considerably. In some schools, the TLC leaders did not see this as a core part of their role. They believed their role was to deliver the session content at the TLCs, but that the school lead was responsible for all other components of the programme.

> *I need to be completely honest, I don't monitor in between the sessions. I could do I suppose, but I've never been explicitly told that I need to. We were under the impression that [the school lead] did all of that.* (TLC leaders)

The implication of this was that the school lead had to lead on monitoring fidelity between TLC sessions. Given that the school lead was often a senior teacher with other responsibilities, this sometimes meant that they did not have the time to monitor fidelity as closely as they would have liked. This led to some staff feeling that their compliance with the programme would not be checked, or if it was, it would be checked in a light-touch way.

This approach contrasts to schools where TLC leaders believed that they did have an important role to play in monitoring the fidelity of staff in their group. This typically involved TLC leaders keeping track of the attendance of members of their TLC group, as well as noting their engagement within the sessions. These TLC leaders also explained that they checked whether staff had completed their personal action plans, and whether they had completed their peer observations.

If the TLC leaders identified any issues with fidelity from individual members of their group, they took it upon themselves to follow-up directly, and only flagged more serious concerns with the school lead. This was seen as an effective and efficient way to monitor staff engagement with the programme, as it meant that the school lead had the capacity to fully address the issues that were flagged to them. It was also welcomed by classroom teachers, who usually preferred the TLC leaders (who were often regular teachers or middle leaders) to check whether they had completed various activities, as opposed to the school lead (who was usually a senior teacher).

Factors affecting adoption and implementation of EFA

Eleven factors were identified that affected the adoption and implementation of EFA. Implementation refers to the deployment and configuration of the EFA programme's core components, whereas adoption relates to the extent to which school staff embrace the EFA programme and incorporate it into their daily practice. These 11 factors are outlined in the sections below and summarised in Table 30.

Table 30: Summary table of factors affecting adoption and implementation of EFA

Factor	Key finding(s)
Alignment to other CPD priorities	● Teachers were more likely to embrace the programme when it was perceived to link with other CPD priorities the school had pursued
Competing school priorities	● Schools that used EFA as their primary professional development programme for staff tended to see quicker adoption of the programme ● Schools where EFA was one of several programmes that staff were expected to engage with negatively impacted adoption of the programme
Impact of Covid	● In the most extreme instances, Covid led to some schools pausing the programme

	entirely ● Covid led to higher staff absences, which made it more challenging for staff to organise peer observations ● Some evidence that cross-departmental working that happened during Covid made some teachers more receptive to the TLC format
School leadership	● Where SLTs were more bought into the EFA programme, they were able to drive quicker adoption of the programme in their schools ● Supportive SLTs regularly attended TLC sessions, which helped improve implementation of the programme ● Where SLTs recognised teacher engagement with EFA, adoption was quicker
School culture	● Historical systems that excessively monitored teacher performance led to some teachers feeling wary of the EFA programme, hindering adoption ● Where schools had an 'open-door' culture, adoption of the programme was faster
EFA school lead	● The school lead was one of the most significant factors determining the success of the programme in a school ● Where school leads had clear passion and enthusiasm for the programme, implementation was more effective ● The most effective school leads were able to have challenging conversations with disengaged staff or SLT, helping to improve adoption of the programme ● Where the school lead had greater capacity to lead the programme, implementation was more effective
School size	● There was some suggestions that smaller schools resulted in more cohesive staff bodies, which increased their willingness to engage with TLCs
Staff	● Staff with higher intrinsic motivation to improve educational outcomes for disadvantaged students tended to be more open to evidence-based initiatives (such as EFA) ● Some more experienced staff perceived the formative assessment approaches taught within EFA to be too basic, leading to lower levels of engagement ● Schools with a higher proportion of part-time teachers faced greater implementation challenges
Students	● Poor student behaviour made teachers less willing to trial new formative assessment approaches in their classrooms, hindering implementation
TLC leaders	● The ability of TLC leaders to facilitate active discussions in their TLC groups influenced staff engagement with the programme
Workload and progression	● High workloads made it difficult for some teachers to complete peer observations, and in some cases, prevented them from attending TLC sessions

CPD, Continuing Professional Development; EFA, Embedding Formative Assessment; SLT, senior leadership team; TLCs, Teacher Learning Communities.

Alignment to other CPD priorities

School staff were more likely to engage positively with the programme, and ultimately adopt it, where the programme was perceived to align with previous CPD priorities that the school had focused on. Classroom teachers explained that the strategies taught as part of the programme often mirrored the language of previous CPD initiatives that the school had focused on. This made the programme feel more manageable and reduced the feeling of teachers being overloaded with new concepts to work on. It was suggested that this made it more likely that teachers would then test out the new strategies in their classrooms. This alignment often did not happen by chance and was often the result of considerable work by the school lead as noted in the section on 'Factors influencing effective adaptation of the intervention' above.

Some schools also had considerable experience with coaching and had trained some or all of their staff in how to conduct coaching conversations. This aided implementation of the programme as staff were better able to engage in the TLC sessions. For example, schools with this kind of coaching culture reported their staff finding it easier to critically

assess the effectiveness of the strategies they had tested in the TLC sessions. Teachers in schools with less experience of coaching took a bit longer to get comfortable reflecting on their practice in this way.

Experience with coaching was also seen as useful for TLC leaders, as it gave them some of the skills needed to be an effective session facilitator. Some TLC leaders who had previous coaching experience noted that they felt more confident in their ability to encourage reflection from teachers in their sessions. This view was also shared by classroom teachers, who noted that TLC leaders that had previous experience with coaching had been effective TLC leaders from the start of the programme.

Competing school priorities

Schools that used EFA as the primary professional development programme for their staff tended to see quicker adoption to the programme and more effective implementation overall. Teachers in these schools noted it was clear to them what was expected of them, and appreciated the fact that they were being given the time to focus on one programme. In contrast, some schools had EFA as one of a number of programmes that staff were participating in. In these schools, teachers noted that there was sometimes some confusion about what they should be doing, and that it was not always clear, which activities contributed to which initiatives. A further implication of this was that staff had less time to dedicate to EFA, sometimes resulting in fewer formative assessment strategies being trialled in their classrooms, and difficulties finding time to observe their peers.

Box 4. Case study: EFA as part of a multi-year school improvement plan

In one of the case study schools, the school's leadership had developed a 5-year plan to improve teaching and learning. This was designed in such a way to lead to incremental and sustainable changes in teaching behaviour, that would ultimately yield improved student outcomes. The first 2 years of the plan had focused on changes to the school's curriculum and improving pupil behaviour. These two areas were identified as being essential foundations that had to be improved before the school could begin seeking to change teaching practices inside classrooms.

Following the focus on curriculum and pupil behaviour, the school shifted focus to improving teacher practice, and the EFA programme was procured as the vehicle to do this. The school's leadership explained how the EFA programme would help them achieve the goals set out in their school improvement plan, and noted that participation in the programme would be the primary 'Teaching and Learning' focus for the past year.

Teachers in the school were highly complementary of the approach taken by the school's leadership. They noted that they appreciated understanding how EFA supported the school's overall goals, and how it built on the priorities of previous years. They also noted that the school's approach to have one or two focuses in each academic year gave them the time and capacity to properly engage, which they believed led to more substantial improvements occurring than if they had been working on all of the priorities at the same time.

EFA, Embedding Formative Assessment.

Some schools where implementation of the programme was particularly strong had developed a multi-year school improvement plan and had identified a good moment in this process to introduce EFA. The perceived benefit of this approach was that schools took an incremental approach to improving teacher behaviour, building skills slowly over a long period of time, rather than imposing a number of initiatives at the same time. Further information on this approach can be found in the case study in Box 4, above.

Impact of Covid

The spread of Covid throughout the UK severely impacted the adoption and implementation of the programme in some schools. The most extreme consequence was for some schools to decide to pause the programme throughout this period.

> *Then Covid hit. What was this amazing buzz and really amazing buy-in went to staff falling ill, being absent, being called all over the place to cover, and so all of those peer drop-ins and all of the peer work was scuppered because of illness. People…were starting to get really stressed. So the headteacher and I made the decision to pause.* (School lead)

Even where schools chose to persist with the programme, staff absences (as a result of Covid) made it particularly difficult to complete some of the core elements of the programme, such as the peer observations. Given the challenge of finding time to complete the peer observations, and the general stress that teachers were under through this period, some schools decided to not expect staff to complete the observations, and rather encouraged them to be prepared to share some reflections on their implementation of a given formative assessment strategy in the next TLC session.

Some staff also noted that they were excited to test out new strategies within the programme, but that to do so required time to think about how best to integrate it into their lesson sequence, but to do this required sufficient time to plan. There was a general view that teachers had less headspace and capacity due to Covid (e.g. some teachers noted that they had fewer free periods as they had to do more 'emergency covers'), and that this gave them insufficient time to plan and implement the strategies learnt in the TLC.

There was one positive consequence of Covid identified. Interestingly, some staff suggested that the experience of working in 'bubbles' through the pandemic had actually got staff used to working cross-departmentally, as they were often teaching on a particular corridor surrounded by teachers from across the school. These staff reported that this experience made the TLCs—and their focus on reflecting on classroom practice with teachers from across the school—less daunting, as they were more used to working with colleagues in other departments.

School leadership

The research found that a school's leadership both positively and negatively impacted the adoption and implementation of the EFA programme.

First, the extent to which the school's leadership was bought into the EFA programme had a significant impact on school staff's likelihood to embrace the programme and embed it into their daily practice. In some schools, it was clear that the headteacher and other senior leaders had engaged with the research behind the EFA programme, and had carefully thought about how the programme could help the school achieve their medium-term ambitions. For example, in some schools, the headteacher gave a welcome address to staff during the launch event, in which they articulated the importance of EFA, and how it was appropriate for the school at that particular point in their school improvement journey. This type of explicit support emphasised the importance of the programme and helped develop teachers' own beliefs in the importance of the programme.

Supportive SLTs were also more likely to attend and engage in the TLCs, which supported the implementation of the programme. By having senior leaders participating in TLCs, just like all other teachers in the school, it made it clear that all staff—irrespective of seniority—can improve their teaching practice. In contrast, senior leaders did not attend TLCs in some schools, which fostered frustration and resentment from staff. In turn, this made it harder to effectively implement the programme.

> *SLT are supposed to attend the TLCs but they don't turn up sometimes. They're not that present, not really engaging with it, and so everyone felt pretty annoyed about it because we were being asked to do this thing that they didn't seem to care about.* (Classroom teacher)

In addition to attending TLC sessions, effective leadership teams recognised the importance of acknowledging and celebrating teacher's engagement with the programme. This was sometimes done by senior teachers thanking staff for their participation in whole-staff briefings, and sometimes through small interactions that SLT had with individual teachers. In such cases, staff felt that their participation was being noted and appreciated, which made them more willing to continue to participate positively. In contrast, some teachers in other schools noted that they felt that the school's leadership were not checking for compliance with the programme, and so were unable to provide meaningful recognition of teacher's participation. This led to staff feeling that the programme was not a priority for senior teachers, and so in turn they assigned less importance to the programme, reducing intervention fidelity.

> *There was not really any acknowledgement of what we were doing, no notice of whether we had done the things we had to do. I think a lot of people just wanted some recognition for that, and when they didn't get that, they started to disengage a little.* (TLC leader)

School culture

School culture played a pivotal role in the adoption and implementation of the EFA programme in a number of ways.

First, the school's culture towards monitoring teacher behaviour affected implementation of the programme. Some schools had had historical policies of using high-stakes observations of lessons as a way to monitor teacher performance, a policy that caused substantial stress and anxiety for many staff. Even in some cases where the school had moved away from such measures, the prospect of observations caused significant concern for teachers. In some cases, due to these historical policies and the general culture within the school, some teachers perceived the EFA programme as another way for the school's leadership to monitor teacher behaviour.

> *We used to have a system where all teachers would be observed in this quite intense way by SLT every few weeks. The feedback was really inconsistent, and the whole system created a lot of mistrust and staff became really resistant to it, it used to stress people out, and then they began to feel quite hostile towards it. So the idea of doing more observations as part of the EFA programme initially caused some concerns.* (TLC leader)

This contrasted to other schools where observations had not been used as a high-stakes measure of performance, with teachers in these schools often much more comfortable to host visitors in their classrooms. This so-called 'open-door culture' was observed in a number of schools, and it was suggested that this culture made the adoption of the EFA programme simpler and more efficient in the short-run given staff experience and comfort around conducting low-stakes observations of one another.

> *We had a culture of observation here already. It was normal for people to wander in and out of classrooms, it's completely normal. So the peer observations as part of EFA just didn't feel like a big deal.* (Classroom teacher)

Teachers in some schools had a high level of mistrust towards SLT. Some teachers suggested that they felt uncomfortable expressing their views, or being totally honest about their teaching practice, in front of SLT, often for fear of judgement. In other cases, some teachers believed that SLT members participated in TLC sessions to monitor their engagement, rather than because they too could improve their teaching practice. This level of mistrust was reported to lead to more stilted discussion in TLC sessions, as teachers were reluctant to share. Overall, this was perceived to reduce the effectiveness of the programme in the school.

> *I imagine the SLT that comes in is a spy [laughs], and they're checking that we're doing what we're supposed to…maybe that makes people less willing to share.* (TLC leader)

Lastly, some schools had a school culture that was informed by high levels of unionisation among the staff. Staff in such schools were reported to be more wary of measures and initiatives, in case they increased workload. The implication of this was that schools had to be very careful in the way in which EFA was presented, and it had to be clear that it would not be implemented outside of teacher-directed time. In some cases, this resistance to new initiatives led to some pushback from staff towards the EFA programme, which led to the adoption of the programme by all staff taking longer than in schools with lower levels of unionisation.

> *We're a very unionised school, and sometimes the staff can be quite militant. So you have to be really careful how you phrase things and what you do. For example, I would love the EFA meetings to run in addition to the regular meeting schedule because they're effectively sharing good practice. They could be seen as training, all of those things, but I can't do that. It has to be in the directed time meeting schedule…When we first introduced the programme, some staff weren't too happy because they thought it was an additional thing for them to do.* (School lead)

EFA school lead

As noted in previous sections (e.g. in the sections on 'Factors influencing effective adaptation of the intervention', 'Encouraging intervention fidelity', and 'Monitoring intervention fidelity: School lead'), the school lead's skills, experience, and attitudes played a critical role in the adoption and implementation of the EFA programme. Feedback from EFA

mentors suggested that the abilities and competency of the school lead was often the most significant factors determining the success of the programme within a school.

> *In my opinion, the school lead is the most significant part of the entire implementation process. If they're good then the programme will run so much better. If they're not, the programme will really suffer, no matter how good everything else is.* (Mentor)

The first factor common to some more effective school leads was their passion and enthusiasm for the programme. Such school leads recognised that the success of the programme depended on their ability to persuade and convince other staff of the benefits of the programme, and that showing high levels of enthusiasm for the programme was an effective way to do this. In general, where school leads had this high level of passion for the programme, they were more able to persuade teachers within the school of the benefits of the programme, which increased teachers' intrinsic motivation to engage with the programme, and in turn improved adoption of the programme.

> *I think the passion and commitment of the leader is the most important thing. If you have somebody who's leading it [the programme] who is absolutely committed, almost evangelical about what they're doing, then that enthusiasm and passion and commitment will impact its delivery and effectiveness.* (TLC leaders)

In some instances, this passion for the programme was rooted in their overall determination to have impact in the education sector. In particular, it was clear that some of the most effective school leads were driven by a desire to tangibly improve pupil outcomes, particularly in disadvantaged communities. This commitment informed their overall leadership of the programme, as they understood that implementing the programme effectively had the potential to improve the quality of education—and subsequent life outcomes—for pupils in their school.

> *I'm also really passionate about working in schools where there is a substantial amount of disadvantage and challenges that our students are facing. I'm really passionate about closing that disadvantage gap for our students, and this programme and formative assessment, as a pedagogy, is shown to be most effective, so that's why we chose it. That commitment is what gets me out of bed in the morning and motivates me to do my job well!* (School lead)

Beyond having a passion for the programme, effective school leads often had the skills and personality to inspire their fellow teachers. For example, in some schools, the school lead had the public speaking skills necessary to provide the rationale for the EFA programme, and could explain how the programme built on their previous work as a school, and how the programme would help the school achieve their medium-term goals. Teachers commented on the power of such addresses and noted that the ability to situate the EFA programme as a means to achieving shared goals was highly effective for motivating teachers to engage with the programme.

The ability of the school lead to have difficult conversations with staff was also noted as being an important factor influencing the implementation of the programme. In schools with higher fidelity to the programme, it was apparent that the school lead was more comfortable chasing up members of staff that were not fully engaging with the programme. In doing so, the school lead was able to address issues that were preventing staff from engaging, as well as constantly reminding staff of the importance of the programme.

Some of the most effective school leads were also able to challenge the school's leadership if they felt they could be doing more to support the programme. In some instances, the school lead explained that they had faced issues of SLT members not fully engaging in the programme, but that they had followed up with them to address the problem. It is important to note that the ability to challenge the SLT was more straightforward when the school lead themselves was a member of the SLT.

> *I've had a few obstreperous members of SLT which I've had to deal with. I just said 'no, you're a senior member of staff, put your game face on, you can't not turn up to sessions'. That's just something that I've had to do.* (School lead)

In contrast, in some schools with lower fidelity to the programme, the school lead was less comfortable challenging poor engagement with the intervention. For example, in some schools, teachers reported that the TLC sessions sometimes ran for less than 75 minutes, and that the school lead would not question the fact that teachers were leaving the sessions early.

Yes, people do leave significantly early sometimes, but I don't really want to interfere with that. They might have good reasons for doing so. It might just be that they've broken up the meeting and that people are going off to write up their action plans, I don't know. But I try not to be too dictatorial. (School lead)

The overall capacity of the school lead was also highly consequential for the implementation of the programme. In particular, school leads that had fewer school-wide responsibilities and more time to dedicate to the leadership of EFA, were able to better organise and implement the programme. Practically, this often meant that the school lead had more time to make appropriate adaptations to the programme, to encourage intervention fidelity, and to monitor intervention fidelity—all of which aided adoption and implementation. One way that some schools sought to increase the capacity of the school lead was to assign them with a deputy, who also held responsibility for the implementation of the programme.

Lastly, some of the most effective school leads were also particularly receptive to feedback. It was noted that they would actively seek out feedback on implementation of the programme, which helped them to make timely adaptations where necessary.

School size

Some respondents suggested that adoption and implementation of the programme was simpler in smaller schools. It was suggested that teachers knew one another better in smaller schools, and this led to teachers feeling more comfortable in cross-departmental TLC groups. In turn, this resulted in teachers being more willing to share in the sessions, and more willing to take risks (in terms of trying out new teaching strategies) in front of their peers.

Some respondents also suggested that in smaller schools the school lead may have a better understanding of staff personalities, and that this would help them to allocate effective TLC groups.

Staff

Certain characteristics of school staff influenced the adoption and implementation of the programme. First, some schools were reported to have large numbers of staff that were particularly motivated to improve educational outcomes for disadvantaged pupils. It was suggested that staff that had this strong 'moral purpose' were more open to new initiatives, and engaged with them more actively, provided that they were convinced that it would likely benefit the pupils in the school. In some cases, it was also suggested that these teachers were often newer to the teaching profession than others. Where a school had high numbers of teachers with such strong convictions, adoption to EFA was sometimes faster, and implementation was sometimes easier, as the teachers were adequately persuaded that the programme's desired outcomes were aligned with their personal beliefs.

The staff at [school name] are amazing. At the launch day [event] they were asking great questions, really invested in the programme and evidence. There's a real purpose behind what they all do. [School name] has a high proportion of disadvantaged students, and there's a real sense of moral purpose in the school in terms of what they're trying to achieve. That shapes the dynamic of the staff body, and makes them more open to new ideas I think. (Mentor)

In contrast, some staff bodies were more resistant to the programme (and to new educational initiatives in general). Some of these schools were in traditionally high-achieving schools, often out of major urban areas, and with a stable staff body (i.e. many of the staff had been at the school a long time). It was suggested that having such a stable staff body made the school more resistant to new programmes, in part because they had seen so many programmes before and sometimes felt cynical about their effectiveness.

In particular, it was noted that some more experienced staff members believed that the teaching practice they had used up to that point had served them and their pupils well, and that they did not need to be taught new formative assessment strategies. This, in turn, made them more resistant to the new ideas and concepts being taught as part of the EFA programme, and made them less willing to participate.

There are a few staff whose attitude was, 'I don't need this. I've been teaching for 25 years. I've taught the same way for 25 years. I get good results. Why on earth are you trying to teach me to suck eggs?' Yes? One in particular who went, 'Yes, well, all research is just rubbish. It's all a fad. It'll all change in a couple

of years, so why would I bother to do this when in two years' time…?' and described academic research as 'stuff off the internet'. (School lead)

In schools where there was discontent from particular staff, it was usually a small minority of staff that expressed their views in this way. However, as some respondents noted, it was often the teachers that were most frustrated with the programme that were most forthcoming in expressing their views. This gave the impression that the resistance to the programme was sometimes greater than it actually was.

Some more experienced teachers reported finding the content of the TLC sessions too basic, and that it did not push them on in their own teaching practice sufficiently. It was suggested by some respondents that the content was reminiscent of what they had studied in their Postgraduate Certificate in Education (PGCE) and in their very early days of teaching, and were now hoping for more advanced knowledge. This attitude sometimes made them more resistant to the programme, as they perceived it to be covering things that they already knew.

This issue—of some experienced staff feeling frustrated by the level of the content—was mitigated effectively by some school leads. In some cases, the school lead particularly emphasised that the core of the programme centred around sharing best practice between teachers, rather than introducing them to new topics. They were able to therefore convince these more experienced teachers that their role in the TLC was to share the knowledge they had accumulated over their careers with less experienced members of staff. Some school leads also opted to adapt the content of the programme in response to some of these views (see the section on 'Appropriate adaptations' above, for further discussion).

Lastly, it was suggested that schools with a high proportion of part-time teachers found adoption and implementation of the programme more difficult. In such schools, it was reported that it was harder to ensure high levels of attendance at the TLC sessions because some teachers were not working on specific days, and that it was also harder to arrange peer observations given the reduced timetable that part-time teachers were working.

Students

Student behaviour was an important factor influencing the implementation of the EFA programme. In some schools, poor pupil behaviour was noted as a particular challenge, and there was some suggestion that this was exacerbated by high levels of pupil poverty. Some respondents also noted that pupil behaviour had deteriorated after Covid-related school closures.

As noted in earlier sections, a core component of the EFA programme is teachers trying out new formative assessment strategies in their classrooms and then reflecting on these afterwards. However, some teachers explained that poor pupil behaviour made them less willing to test out new strategies in their classrooms. These teachers were concerned that something might go wrong when they first tried out a new teaching approach, and that this was a risk they felt they could not take because of poor pupil behaviour.

TLC leaders

Some characteristics of TLC leaders influenced the effective adoption and implementation of the EFA programme. In particular, the section on 'Encouraging intervention fidelity' above, outlined examples of measures taken by TLC leaders that supported implementation of the programme, and the section on 'Alignment to other CPD initiatives' above, provided a discussion of how TLC leaders with experience of coaching often led to more effective TLC sessions. Beyond this, the ability of selected TLC leaders to facilitate an active discussion was also identified as a factor affecting the quality of implementation. In one school where overall fidelity to the programme was perceived to be lower, TLC leaders explained that they found it hard to justify keeping staff in the sessions when they had nothing else to say, and would let staff leave the session early. The issue of TLC groups running out of things to say was not identified in other schools. This highlights the importance of carefully selecting the TLC leaders, and ensuring they have the skills to facilitate a healthy discussion.

Furthermore, some schools noted that selecting TLC leaders that were personable, well-liked by their peers, and had the ability to have difficult conversations where necessary, supported overall implementation of the programme.

Workload and progression

Some teachers noted that they had a particularly high workload, with a large number of lessons to teach each week, which did not give sufficient time to fully engage in the programme. In particular, teachers noted that a heavy timetable made it particularly challenging to complete the peer observations, as it was difficult to find time where both them and their partner were available. This was also reflected in the endline survey data, with 19% of respondents noting that they did not have time to arrange an observation.

A small proportion (10%) of endline survey respondents noted that other school responsibilities (such as teaching extra lessons) prevented them from attending or engaging in the TLC workshops. Unfortunately, it was not possible to interview any teachers in this position to understand how this affected their engagement with the programme. However, at a minimum, it is likely that missing the TLCs meant that these teachers were not introduced to the new formative assessment strategies, and they did not have time to reflect on their implementation of other strategies in the preceding weeks.

Indicators of embeddedness

Both mentors and school leads identified a range of potential indicators of embeddedness. In particular, they noted that they were hoping to see tangible changes in classroom practice, with teachers increasingly confident in using a range of the strategies and teaching approaches from the EFA programme. Both mentors and school leads explained that they usually assessed whether this was happening through whole-school observations or on learning walks.

Beyond collecting observation data, some school leads and mentors noted that they found it helpful to review pupil books to look for evidence of formative assessment. If they found evidence in books (such as low-stakes quizzes or examples of hinge questions) then that acted as a further indicator that EFA was becoming increasingly embedded in school practice.

Other school leads also noted they looked for evidence of formative assessment strategies being embedded into Schemes of Work. It was suggested that when a technique was embedded into the curriculum in this way, teachers would default to using it. In particular, where teachers shared lesson planning within the department, it was noted that if the formative assessment strategy was incorporated into the shared lesson plan, then most teachers in the department would likely use it as they all used the same resources. For some school leads, this was a strong indicator of embeddedness.

Some mentors and school leads suggested that rather than looking for evidence of embeddedness in teaching practice or pupil work, they instead attempted to assess embeddedness through conversations with teachers in the school. They noted that it was still a relatively early stage of implementation (within the first year) and that they felt it was too early to see significant changes in teaching practice. Instead, an early indication of embeddedness was the extent to which teachers were comfortable talking about using the techniques in their classrooms, and the range of techniques that had been applied.

However, some school leads noted seeing some evidence of some formative assessment strategies being used was insufficient evidence to determine the embeddedness of the programme. There was a concern raised by some school leads that staff were good at implementing the strategies they had learnt in the previous TLC, but in doing so, then forgot to implement the strategies they had learnt before that. Some school leads noted that for EFA to be embedded in teacher practice required them to be able to recall and implement the full range of strategies they had been introduced to, not just the ones they had learnt about most recently.

Factors affecting embeddedness of intervention

Given the relatively early stage for many of the schools involved in this evaluation, it was too soon to generate much evidence on the various factors affecting the embeddedness of the intervention. Despite this, two factors were suggested by respondents at this stage.

First, the stability of leadership of the EFA programme was deemed to be particularly important for the programme's sustainability. In some schools, the leadership of the programme had to change for various reasons. While the new school leads were often competent, there was inevitably some loss of momentum in the handover process. It was also

suggested that the member of staff that had originally procured the programme was a particularly strong champion of the programme within the school. This may be because they had to convince the school's leadership and other staff of the programme's benefits. Whereas the member of staff taking over the leadership of the programme did not have the same history and association with it.

Second, the level of staff turnover within the school was also perceived to affect the embeddedness of the intervention. It was noted that schools that had large numbers of staff leaving at the end of the year would likely find it harder to embed the programme over time. It was suggested this was likely due to the new staff having to learn the basics of the programme. The length of the programme (2 years) was also noted to be helpful for embeddedness, as it provided sufficient time to actually change teacher behaviour—these benefits would occur for fewer teachers if large numbers left.

Sustainability of the programme following scale-up

There was evidence to suggest that the programme had changed teacher's classroom behaviours at this stage of implementation. Around 76% of endline survey respondents noted that they had used EFA strategies within their teaching, and a further 62% of respondents felt that their general teaching practice had changed as a result of their participation in the EFA programme. Around 47% of teachers reported feeling confident using the teaching techniques and resources provided by the programme, and 79% of respondents intended to carry on using a range of formative assessment techniques in their everyday teaching—a figure that is likely to grow as teachers spend more time on the programme. These figures were also reflected in feedback from school leads, some of whom noted that they had seen tangible changes in the teaching behaviours of teachers at their schools since the programme began. Taken together, this evidence suggests that teachers may continue to use the strategies after the end of the programme.

> *We've definitely seen big changes in teaching since starting. On learning walks, I'm seeing a lot more formative assessment happening now which is great. (*School lead)

However, there was limited evidence that teachers want the specific components of the EFA programme to persist after the end of the scale-up phase. When asked, 36% of endline survey respondents reported wanting the TLCs to continue following the end of the 2-year programme (though it is important to note a further 30% had no view either way). A similar assessment was made of peer observations, with 44% of respondents wanting these to persist after the end of the programme. These figures suggest that teachers are not fully convinced of the usefulness and effectiveness of the TLCs and peer observations, and suggest that it may be challenging to expect teachers to continue to engage with these after the end of the programme.

Conclusion

Table 31: Summary of findings

Research topics	Findings
Strategy	SSAT had a target to reach an additional 50 schools in the 2021/22 academic year, but were unable to meet this target, with 14 schools starting the programme in the 2021/22 academic year. Further details on possible reasons for this can be found in 'Reach and recruitment' below. SSAT introduced a new CRM system in the 2021/22 academic year, helping to digitise and automate a series of time-consuming manual processes. This helped reduce the burden on key SSAT staff, and also allowed them to better evaluate their marketing methods (e.g. measuring open/click rates). The size of the team responsible for the EFA programme was also expanded to provide additional capacity during the scale-up phase No significant changes were made to the core content of the programme (TLC sessions) to support scaling. The most significant change to the programme during the research period was a shift to online delivery due to Covid. Some schools had to complete their launch events online, and some meetings between school leads and mentors that usually happened in-person were conducted virtually
Structures, systems, and processes	SSAT collected a range of qualitative and quantitative M&E data on the programme—some of these measures were collected through online forms, which made analysis more straightforward and reduced the burden on SSAT staff, which in turn supported the scaling of the intervention SSAT's current M&E approach relating to the monitoring of programme outcomes was deemed to be a proportionate and efficient approach to impact management (given the existing impact evidence from an effectiveness trial), and no further improvements were suggested in this area. Some potential improvements to programme M&E were identified across school reach and recruitment, programme implementation, and programme quality. It is suggested that SSAT: i. ensure complete data is collected where possible; ii. collect more structured quantitative feedback from stakeholders; iii. conduct periodic analysis of the data collected; and iv. develop formal M&E plans relating to each area The research found that SSAT collected some limited data on mentors (their programme support staff), but that no substantial analysis was conducted on the data currently collected. No systems were identified to ensure that the data collected was acted upon. The main potential improvement identified here was for SSAT to begin collecting and analysing data on mentor reach and recruitment, and on mentor training and support. To support all these suggested improvements, BIT has worked with SSAT to develop a comprehensive M&E plan and a new set of feedback surveys to support this
Reach and recruitment	In total, 23 schools were recruited in the 2021/22 academic year, of which 14 schools actually began the programme—this was considerably lower than the target of 50 new schools that SSAT were aiming for. Around 35% of the schools recruited had above-average levels of FSM and the majority (86%) were rated 'Good' or 'Outstanding' by Ofsted. Around 64% of the schools recruited were located in London or the south-east, with 21% located in an 'Opportunity Area'[13] A total of 375 schools entered the EFA sales pipeline between January 2019 and June 2022, with 14 schools from this pool beginning the programme in the 2021/22 academic year—implying a sales conversion rate of 3.7% A range of factors were identified that help explain this conversion rate, including concerns from schools about increasing workload for their staff (particularly in the wake of Covid),

---

[13] See https://www.gov.uk/government/publications/social-mobility-and-opportunity-areas.

	concerns about the flexibility of the programme, and a potential lack of momentum in the sales approach
Fidelity, contextual factors, and sustainability	SSAT staff suggested that the intervention contains three essential features: i. teachers participating in TLCs and reflecting on their practice; ii. teachers observing one another between TLCs; and iii. producing personal action plans to support them as they try new techniques in their classroom. In general, school leads, TLC leaders, and classroom teachers considered the TLCs and peer observations to be central to the programme (and stuck faithfully to these components), but there was some evidence to suggest a more inconsistent approach to personal action planning

A wide range of adaptations to the programme (made by school leads, TLC leaders, and classroom teachers) were identified—including small changes to the language used within the programme, changes to the delivery format, changes to the size of the TLC groups, as well as more substantive changes to the programme content. Two factors particularly facilitated the effective adaptation of the intervention: i. regular opportunities for TLC leaders and school staff to provide feedback on the programme to the school's leadership; and ii. the quality and regularity of the relationship between the school lead and the EFA mentor

Intervention fidelity was encouraged by SSAT via the EFA mentors but was largely the responsibility of the school lead and TLC leaders. Some schools developed innovative systems to reward and incentivise adherence, while others incorporated EFA into their performance management processes. Some of the schools with the highest fidelity took additional measures to build staff buy-in to the programme, which in turn encouraged adherence

Fidelity to the intervention was monitored by SSAT (through the EFA mentor), the school lead and TLC leaders. The EFA mentor monitored intervention fidelity through a series of informal (often virtual) calls with the school lead, coupled with an in-person visit to the school at the end of the first year of implementation. The level of monitoring conducted by school leads and TLC leaders varied across schools. Schools with the highest fidelity to the programme tended to have school leads and TLC leaders that actively monitored staff attendance at TLC sessions and their completion of peer observations and personal action plans. Some schools took a less active approach to monitoring fidelity given the additional pressures that many schools had been under in the wake of Covid, which often resulted in lower overall fidelity to the intervention.

A wide range of factors were identified that affected the adoption and implementation of the programme—including the alignment of the programme to other CPD priorities, the ongoing impact of Covid, the school's leadership, the school's culture, and specific characteristics of a school's staff and pupils |

BIT, Behavioural Insights Team; CPD, Continuing Professional Development; CRM, Customer Relationship Management; EFA, Embedding Formative Assessment; FSM, Free School Meals; M&E, monitoring and evaluation; Ofsted, Officed for Standards in Education, Children's Services and Skills; SSAT, Schools, Students, and Teachers Network; TLCs, Teacher Learning Communities.

Discussion

There was evidence that SSAT had developed and implemented a range of measures to support the scaling up of the EFA programme. This included changes to personnel, growing the EFA team, and introducing a range of (digital) efficiencies to support scaling. However, even with these changes in place, they were unable to reach their target number of schools for the 2021/22 academic year. There was some evidence to suggest that this was—at least partly—driven by the lasting impact of Covid on the UK's education system, with some schools more reluctant to begin new programmes, and in particular intensive programmes such as EFA, during a period of such instability. Other factors— such as a school's exposure to EFA before signing up, and their concerns about the flexibility and potential burden of the programme—were also identified, and help to explain the sales conversion rate of 3.7%. It is important to note the challenges faced in recruiting schools to participate in this research that had considered the programme but ultimately not chosen to sign up. This means that there were likely factors that help explain the sales conversion rate that were not identified in this report.

Despite these challenges, 14 new secondary schools were successfully recruited and began the programme in the 2021/22 academic year. There was some evidence of variable levels of fidelity to the programme across these schools, which was driven by a wide range of factors. In particular, some of the schools with the highest levels of fidelity to the programme implemented a range of innovative and school-specific measures to encourage staff to engage positively with the programme, and coupled this with regular, consequential monitoring of fidelity. It was clear that the lasting impact of Covid continued to be a challenge for schools, with some schools having to focus on other priorities, or deciding to monitor staff engagement less. However, some schools were able to maintain momentum behind the programme, and continued to build staff buy-in, throughout a year in which the education system was still severely affected by Covid. This was often driven by the personal skill and experience of the school lead, and their ability to make appropriate adaptations to the programme that reflected the burden that teachers were under.

The ability to make appropriate adaptations to the programme was facilitated by the level of feedback that the school's leadership was able to gather on the implementation of the programme (which helped them make responsive changes), but also the relationship between the EFA mentor and the school lead. SSAT's scaling strategy allowed for one in-person visit for the mentor at the end of the first year of implementation. This was deemed to perhaps be too late in the year and meant that the mentor could only properly gauge a school's fidelity to the intervention at this late stage. It is therefore suggested that SSAT explore the option of bringing the in-person mentor visit to earlier in the academic year.

Recommendations

Based on the research activities and findings, 47-specific recommendations have been made to SSAT covering the full range of research topics over this past year of research. These recommendations have been delivered as quickly as possible during the year, as part of the rapid feedback described in the section on 'Approach to feedback and reporting' above. Many of these recommendations focused on SSAT's sales process and included a range of measures that could help to increase the number of schools reached and recruited. Given the failure to meet their scaling targets in the 2021/22 academic year, it is recommended that SSAT particularly prioritise implementation of these recommendations. The full list of specific recommendations can be found in Appendix B.

Based on the findings in this interim report, the following recommendations should be considered by those wishing to scale educational interventions:

- **Ensure scaling strategies include a comprehensive plan for recruitment.** Organisations wishing to scale an intervention should first ensure there is sufficient demand for the intervention and should develop a comprehensive plan for how they plan to reach their target number of stakeholders within the scale-up phase. A wide range of factors can affect the ability of an organisation to recruit stakeholders; these should be carefully considered in advance and mitigations planned where appropriate. One improvement to the marketing message to schools might be to address the mismatch between the perceived and actual level of acceptable adaptation. Some schools in the sales pipeline think that the programme is less flexible than it actually is.
- **Collect and analyse data that helps to explain the types of settings reached, and the possible reasons why some stakeholders choose to not sign up.** Collecting data on the specific characteristics of the schools that enter the sales pipeline helps build an understanding of the types of setting being reached, and which ones are not being adequately reached. Quantitative data should be supplemented with qualitative feedback from stakeholders to understand their reasons for signing up or not. This would allow organisations to target stakeholders with specific characteristics, which may improve recruitment and quality outcomes (e.g. allowing the targeting of school types that might particularly benefit from the intervention).
- **Automate and streamline internal processes where possible.** Scaling an educational intervention will likely place additional burden on key staff, and all internal processes that can be automated or streamlined should be in advance of the scale-up phase. In some cases, this will require investment in new technology.
- **Identify and encourage the facilitators of effective adaptation of the intervention.** Effective adaptation is critical to ensure the intervention can be adopted in a wide range of contexts. Providers should carefully consider the components of the intervention that can be adapted if necessary, and how stakeholders can be supported to make any adaptations as effectively as possible. For example, findings in this evaluation found that effective adaptations to EFA were facilitated by regular, semi-formal feedback mechanisms between school staff and the school lead—in the context of EFA, such feedback mechanisms should therefore be encouraged.

- **Distribute responsibility for encouraging and monitoring fidelity to the intervention.** Intervention fidelity will likely be higher when encouraging and monitoring fidelity happens at multiple levels. For example, this responsibility could be shared between the lead organisation, the lead contact in the educational setting, and any staff that support the implementation of the intervention.

Future research and publications

The findings in this interim report were based on data collected in the first year of the scale-up, which focused particularly on implementation of the programme. The next phase—in the second year of the scale-up—will focus on the embeddedness of the intervention, though findings will be generated across the full range of research questions. The next phase of the research will include further review of SSAT's administrative data, further surveying of school staff, school observations and interviews, and interviews with SSAT staff about strategy, organisational capacity, and processes. An additional four schools will also be recruited to participate in light-touch case studies. The next phase of the research will also include a comprehensive cost evaluation of the programme.

There was some evidence at this interim stage to suggest that school-based mentors could be particularly effective mentors, particularly in supporting schools to understand the core features of the intervention and helping them make appropriate adaptations. This hypothesis could be further explored in the next phase of the research.

The upcoming research activities will inform a fourth feedback output that will be submitted to SSAT in July 2023, with a focus on school embeddedness and sustainability. There will then be a final findings and post-mortem workshop held in October 2023, with the final report for the scale-up evaluation completed by February 2024.

References

Ariely, D. and Wertenbroch, K. (2002) 'Procrastination, deadlines and performance: Self-control by precommitment'. *Psychological Science*, 13(2): 219–24.

Barker, P.M., Reid, A. and Schall, M.W. (2015) 'A framework for scaling up health interventions: Lessons from large-scale improvement initiatives in Africa'. *Implementation Science,* 11: 12.

Black, P. J. and Wiliam, D. (2009) 'Developing the theory of formative assessment'. *Educational Assessment, Evaluation and Accountability*, 21(1): 9.

Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology'. *Qualitative Research in Psychology*, 3(2): 77–101.

Department for Education. (2022). 'Opportunity areas'. Available at: https://www.gov.uk/government/publications/social-mobility-and-opportunity-areas (accessed 30 March 2021).

Education Endowment Foundation (EEF). (2015) 'Feedback'. London: Education Endowment Foundation. Available at: https://bit.ly/3vWrcnr (accessed 25 March 2021).

Education Endowment Foundation (EEF). (2018) 'Embedding formative assessment: Evaluation report'. London: Education Endowment Foundation. Available at: https://d2tic4wvo1iusb.cloudfront.net/documents/projects/EFA_evaluation_report.pdf?v=1630925723 (accessed 12 September 2022).

Gorard, S., See, B. H. and Siddiqui, N. (2014) 'Anglican schools partnership: Effective feedback'. Available at: https://bit.ly/3faV0a4 (accessed 16 March 2018).

Newman, M., Kwan, I., Schucan Bird, K. and Hoo, H.T. (2021) 'The impact of feedback on student attainment: A systematic review'. London: Education Endowment Foundation.

Ritchie, J., Lewis, J., McNaughton Nicholls, C. and Ormston, R. (Eds.). (2013) 'Qualitative research practice: A guide for social science students and researchers'. London, Thousand Oaks, CA: Sage Publications Ltd.

Rosenshine, B. (2012). 'Principles of instruction: Research-based strategies that all teachers should know'. *American Educator,* 36: 1, 12–39. Available at: https://www.aft.org/periodical/american-educator/spring-2012/principles-instruction (accessed 17 July 2022).

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'. London: Education Endowment Foundation. Available at: https://bit.ly/3d4Lwum (accessed 25 March 2021).

World Health Organization (WHO). (2010) 'Nine steps for developing a scaling-up strategy'. Geneva: World Health Organization. Available at: https://bit.ly/39s6nXD (accessed 30 March 2021).

World Health Organization (WHO) and ExpandNet. (2009) 'Practical guidance for scaling up health service innovations'. Geneva: World Health Organization. Available at: https://expandnet.net/PDFs/WHO_ExpandNet_Practical_Guide_published.pdf (accessed 30 March 2021).

Appendix A: Detailed research questions

The detailed list of research questions here expands upon the main research questions set out in the body of the plan.

Topic 1: Strategy

1.1.   What is SSAT's strategy for scaling up the EFA programme?

- ○   *How is SSAT set up to deliver the EFA scale-up?*

1.2.   How does SSAT's strategy for scaling evolve over time?

1.3.   What factors influence changes to the scaling strategy?

1.4.   What role does the EEF play in helping SSAT to achieve readiness for scaling up?

Topic 2: Fidelity

2.1.   What are the essential features of the intervention, and what adaptations are appropriate (and required to support scaling)?

2.2.   How does the approach taken to scaling support or hinder fidelity?

2.3.   How is intervention fidelity managed?

- ○   *What measures are taken (both at the SSAT and school levels) to encourage fidelity of implementation?*

- ○   *How is intervention fidelity monitored?*

- ○   *What is the process for agreeing modifications to the intervention?*

- ○   *What action is taken when essential features of the intervention are not consistently implemented?*

2.4.   What are the barriers to, and enablers of, the effective adaptation of EFA?

Topic 3: Structures, systems and processes

3.1.   What challenges are there organisationally when making a sizeable change in the scale of implementation of the EFA programme, and how are these overcome?

- ○   *What changes are made in organisational structure and processes to make the intervention scalable?*

3.2.   How well do SSAT's monitoring and evaluation (M&E) systems support data-based decision-making and how can they be improved?

- ○   *Are they flexible enough to respond to changing needs?*

Topic 4: Reach and recruitment

4.1.   What is SSAT's sales process/pathway to sales for the EFA programme?

- ○   *What changes are made in the approach to recruitment to make the intervention scalable?*

4.2.   How many and what types of schools are SSAT reaching and successfully recruiting?

- ○   *What is the pace of scale-up and what factors affect this?*

4.3.   How do schools respond to the sales approach?

  ○ *What are the barriers and facilitators to schools signing up to the programme, and how could the barriers be overcome?*

Topic 5: Contextual factors

 5.1. What school characteristics affect the adoption and implementation of EFA and how (e.g. culture, school-type, leadership, subject(s) taught, characteristics of individual teachers and mentors)?

 5.2. What are the facilitators and barriers - in the context outside of schools - to scale-up of the EFA programme (e.g. education policy, funding, networks between schools)?

Topic 6: Sustainability

 6.1. What indicates that EFA has been embedded in school practice?

 6.2. What are the facilitators of, and barriers to, embedding the EFA programme in a school?

 6.3. How viable is it for schools to sustain the use of EFA on an ongoing basis after the end of the scale-up? What are the factors that affect this?

 6.4. Is EFA being institutionalised at levels other than the school? What are facilitators of, and barriers to, 'vertical' scaling up, and how can the barriers be addressed?

Topic 7: Cost

 7.1. What is the cost of implementing EFA over 3 years during the scale-up?

  ○ *What is the overall cost of implementing the programme as part of the scale-up for a) SSAT b) schools?*

  ○ *What is the cost of implementing the programme in year 1 for a) SSAT b) schools?*

  ○ *What is the cost of implementing the programme in year 2 for a) SSAT b) schools?*

 7.2. How acceptable is the overall cost of implementation to a) SSAT and b) schools?

 7.3. Is scaling up becoming more cost efficient over time?

 7.4. Is it financially sustainable for SSAT to continue EFA delivery across schools that the programme is scaled-up to, without the support provided by the EEF?

Appendix B: Full list of recommendations

BIT provided a large set of recommendations to SSAT throughout the first year of the scale-up. These were integrated into feedback slide decks submitted to SSAT. The full set of recommendations have been copied below.

Related research topics	Date	Recommendation
Reach and recruitment	July 2021	Review the objectives and content of the online open days. ● Communicate objectives of the event clearly to Host Schools (to ensure presentations and discussion centres on the impact of EFA rather than programmatic details) ● Develop a suggested structure for the event, including suggested talking points for each section ● Embed a Quality Assurance process in which the slides developed by the Host School are reviewed by SSAT before the event
Reach and recruitment	July 2021	When drawing on evidence for sales materials, ensure the evidence is presented in a simple, clear and compelling manner
Reach and recruitment	July 2021	Generate evidence on the effectiveness of different sales materials and approaches (e.g. the printed sales materials), potentially generating ROIs in terms of costs and conversions
Reach and recruitment	July 2021	Build additional staff capacity to support sales so the organisation is less reliant on a single "voice"
Reach and recruitment	December 2021	The current data in the CRM on school characteristics (Ofsted rating, % FSM etc) is incomplete. You might want to collect complete data on key characteristics for each school. To save time, this could be done through a mass upload and matching process using DfE admin data
Reach and recruitment	December 2021	Start collecting regular feedback from schools on: ● Reasons for sign up, e.g. a quick web survey or question in first EFA mentor session ● Reasons for staying, e.g. an annual web survey or question in end of year reflection session ● Reasons for not signing up, e.g. a quick web survey ● Reasons for leaving, e.g. a quick web survey and/or exit interview
Reach and recruitment	December 2021	Conduct periodic aggregated analysis of reach and recruitment data (including the proposed new feedback data). This should be done as part of a planned development process for your recruitment activities (e.g. an annual review)
Reach and recruitment	December 2021	Create a formal M&E plan for reach and recruitment that includes: ● Targets for reach and recruitment (possibly developing these by having targets for specific school types, beyond those specified by funders like the EEF) ● What questions you want to answer about reach and recruitment ● What data you collect to answer those questions ● What analysis you do on the data ● When you do that analysis ● Who is responsible for the analysis ● What you do with the results of the analysis (e.g. whether any results trigger certain actions, whether the results feed into an annual development process…)
Reach and recruitment	January 2022	Ensure sales materials reference EFA's ability to embed other CPD and Teaching & Learning initiatives

Reach and recruitment	January 2022	Ensure sales materials reference CPD practices that EFA is aligned to, making clear that EFA is complementary (rather than duplicating)
Reach and recruitment	January 2022	Review and tweak messaging on EFA's time commitment, noting how time burden can be minimised and why the two-year commitment is so important
Reach and recruitment	January 2022	Clearly explain the benefits of a longer term programme in sales materials, and directly address concerns about waning staff engagement
Reach and recruitment	January 2022	Provide schools with guidance on how other schools have addressed practical challenges like ensuring EFA does not increase total working hours (e.g. allocating 'twilight CPD' hours to EFA)
Reach and recruitment	January 2022	Provide schools with guidance on how other schools have ensured EFA does not result in overly burdened TLC Leaders
Reach and recruitment	January 2022	Ensure the benefits of SSAT support are clearly articulated in marketing materials
Reach and recruitment	January 2022	Review content of emails and check that it follows behavioural science principles, such as ensuring the 'Call For Action' is always in the top line. This could be done with support from BIT under the EEF Capability Building Fund
Reach and recruitment	January 2022	Ensure sales materials reference the potential of a subsidy
Reach and recruitment	January 2022	Ensure that the total cost to a school is clear in all relevant communications
Reach and recruitment	January 2022	Emphasise the range and diversity of schools that EFA has been implemented successfully in
Reach and recruitment	January 2022	Develop a set of FAQs to provide to prospective schools. Many of the recommendations made in this document could be addressed through these FAQs. Suggested questions might include: • Can I see examples of the EFA materials? • Can I try the EFA materials out in my school before deciding whether to sign up? • How does EFA relate to other CPD initiatives? • How long is the programme? • Why is the programme two years? • Can the programme be made shorter than two years? • What are the benefits of receiving support from SSAT for implementation of EFA? • Why are the TLC sessions 75 minutes? • Will EFA lead to an increase in workload for staff?
Reach and recruitment	January 2022	Share sample EFA materials with schools that express interest in the programme
Reach and recruitment	January 2022	Explore options to formalise the opportunity for schools to 'try out' EFA before officially signing up (e.g. by providing SLT with a 'taster pack' to run)

Reach and recruitment	January 2022	Increase the number of 1-1 calls that Programme Lead leads with prospective schools
Reach and recruitment	January 2022	Incentivise existing schools to recruit peers in their network
Reach and recruitment	January 2022	Identify individuals in SSAT's network with a public profile that would be willing to write/tweet about the EFA programme
Strategy (mentors)	July 2021	Formalise and embed the learnings from the mentor recruitment process (e.g. by updating job descriptions, developing a structured interview guide)
Strategy (mentors)	July 2021	Shift to a hybrid model of mentoring on a permanent basis, with more online mentoring and less face-to-face visits. This could help ease recruitment and matching burdens as EFA scales
Strategy (mentors)	July 2021	Review ongoing support & CPD offer for mentors ● Generate evidence on the effectiveness of blogs as a method for mentor learning ● If blogs written by mentors continue to be a core element of supporting mentor development, develop a Quality Assurance process for the blogs to ensure that they are sharing genuinely useful practices ● Explore alternative ways for mentors to share best practice. One suggestion from interview participants was to hold a series of mentor meetings in regional hubs ● Interview participants suggested the benefit of the Education Lead hosting webinars during the year to update mentors on changes and additions to the resources. This would also be a helpful way to reinforce fidelity to the programme

Strategy (mentors)	December 2021	Start collecting and analysing data on mentor reach and recruitment in the same way that you do for schools. For example:
		● A recruitment database and pipeline
		● Feedback form mentors on:
		  ○ Reasons for sign up
		  ○ Reasons for staying
		  ○ Reasons for not signing up
		  ○ Reasons for leaving

Strategy (mentors)	December 2021	Start collecting and analysing data on mentor training and support in the same way that you do for schools. For example:
		● Implementation data:
		  ○ Activities completed, participation of mentors in each
		  ○ Start date
		  ○ End date
		● Quality data:
		  ○ Feedback from mentors
		  ○ Observation of trainings
		● Outcomes for mentors:
		  ○ Self-reported learning
		  ○ Assessments to test knowledge/skills

Fidelity	July 2021	Review the objectives for the online launch event for participating schools. ● For example, given that school staff only have one hour with their mentor now (rather than a full day), it may be appropriate for staff to work through the online materials (inc. watching pre-recorded videos on the rationale for the programme) before the meeting. This would allow the meeting with their mentor to focus on their questions and concerns, and could deal with the more practical aspects of running effective TLCs ● Central to the EFA programme is the idea that effective delivery (of formative assessment) requires practice and support, not just learning why

		it is important. The same is likely true for effective delivery of TLCs, and it could be appropriate to provide TLC Leads with the opportunity to practice facilitating a TLC discussion
Fidelity	July 2021	Include a series of multiple choice questions for school staff to complete after working through the online training. This would help the mentor to assess where there may still be gaps in understanding and where to follow-up with further support or resources
Implementation	December 2021	Ensure all mentors can use and do use the webform for the implementation survey, and remove the paper option
Implementation	December 2021	Give the implementation survey to mentors at start of year in their induction pack
Implementation	December 2021	Make all web surveys feed into your CRM and retire the 'master spreadsheet'. Depending on your system, this may require using a third party survey app that is supported by your CRM
Implementation	December 2021	After linking all web surveys to your CRM, use the CRM to: ● Create auto-dashboards for the programme team to easily review results as they come in ● Create report templates (that automatically produce key statistics and charts) for periodic reviews ● Create trigger emails for Programme Lead when a survey/conversation record comes in and/or when a survey value is below a threshold (i.e. triggering a phone call)
Implementation	December 2021	Create a formal M&E plan for implementation that includes: ● Targets for implementation ● What questions you want to answer about implementation ● What data you collect to answer those questions ● What analysis you do on the data ● When you do that analysis (e.g. review part of each form when it comes in, end of year review of aggregated data…) ● Who is responsible for the analysis ● What you do with the results of the analysis (e.g. whether any results trigger certain actions, whether the results feed into an annual development process…)
Implementation	December 2021	Begin collecting the following extra data on implementation: ● The proportion of teachers participating in EFA in each school ● Feedback on adaptations, barriers and enablers to implementation ● Data on implementation after the 2 year programme is complete (this should be linked to the development of the ongoing support model)
Structures, systems, and processes	July 2021	Scaling may necessitate an increase in headcount and company size. SSAT should consider how to maintain effective communication practices as the company grows, and should be aware of this as a potential risk area that comes with growth
Quality	December 2021	Formally interview mentors, assessing them against the desired qualities listed in your QA framework, and take notes from these interviews
Quality	December 2021	If observation of potential mentors during their training is an important part of your QA process (as suggested in the 'EFA Programme QA process' doc) then we suggest using a semi-structured observation guide, covering the qualities that you're looking for, to help structure and record your observations
Quality	December 2021	Create a webform, linked to your CRM, for mentors to complete their conversation records
Quality	December	Create a webform for the launch event survey, ideally linked to to your CRM, and

	2021	remove the paper option. A platform like Slido, for example, allows you to post a simple code on your power that participants can enter into their smartphone and complete the survey immediately with very little friction. (If you don't have the in-house expertise to update your CRM and survey systems in this way, then you may consider using some of your scale-up funding for development support)
Quality	December 2021	Add some multiple choice questions to the end-of-programme survey for School Leads on the quality of the core components of the support you provide, e.g.: ● Teaching resources ● Facilitation resources for TLC Leaders ● Planning resources for School Leads ● Structure and content of mentoring support
Quality	December 2021	Start formally observing new mentors delivering launch events. Set up this observation programme in a way that can be easily scaled (e.g. with experienced mentors conducting the observations, rather than the Head of Programmes being responsible for all observations)
Quality	December 2021	Conduct periodic aggregated analysis of launch event and end-of-programme evaluations. This should be done as part of a planned programme development process (e.g. an annual review)
Quality	December 2021	Develop your 'EFA Programme QA process' doc into a more complete M&E plan for quality that includes: ● What questions you want to answer about quality ● What data you collect to answer those questions ● What analysis you do on the data ● When you do that analysis ● Who is responsible for the analysis ● What you do with the results of the analysis (e.g. whether any results trigger certain actions, whether the results feed into an annual programme development process…)

Appendix C: Ethical review

BIT has an internal ethics review process, which follows the ethical principles for research developed by the Government Social Research Profession. This project was assessed as low risk along the following dimensions:

- Research methods: Standard research methods commonly applied within the substantive area of the research
- Participants: Non-vulnerable adults
- Subject matter: Research related to a politically and socially uncontroversial area
- Experience: BIT had extensive experience conducting research in the education sector and using the planned research methods.

The only medium-risk aspect of the research related to the nature of the data. Given the use of surveys and interviews, individual-level data was collected (that is not routinely collected). However, all participation was voluntary and involved adults, and all participants were provided with information about the research in order to make an informed decision about whether to participate and what data they were willing to provide.

Informed consent

Interview participants were provided with an information sheet explaining why the study was being conducted, what their participation involved, how their data would be processed, including that all information would remain confidential unless there were concerns about risk to themselves or others, and their rights in relation to withdrawing consent. They were also given a verbal explanation prior to beginning the interview, and had the chance to ask any questions. If they agreed to go ahead, audio recorded consent to participate was taken.

For surveys, information about the purpose of the survey and how the data would be processed formed the first page of the survey. Teachers were also asked to tick a box confirming that they had read this information and consented to take part. They were also made aware of the process through which they could withdraw their data (up until the point of analysis).

For in-school observations, as it was a non-controversial topic and the researchers were not there to observe the behaviour of specific students, the school was asked to act in loco parentis. Schools chose whether to inform parents or seek consent from them in relation to the observation.

Participant welfare

The focus of this evaluation was not a sensitive topic and participants engaging directly in research activities (i.e interviews and surveys) were not classified as vulnerable. The need to minimise the burden of taking part for schools was recognised. This was achieved through using existing M&E data wherever possible, only asking teachers to complete two voluntary short surveys per year, and conducting a maximum of two observations per school.

It was not anticipated that the researchers would encounter any situations that would require action to be taken in relation to safeguarding and/or distress, but this was always a possibility when working with schools. If a safeguarding issue arose in any aspect of the evaluation, all researchers were trained in BIT's policies on safeguarding and conducting research safely. They were able to refer to and take action in line with the following BIT policies and procedures: Adult Safeguarding Policy, Child Safeguarding Policy, Lone Working Procedure and Emergency Crib Sheet for Field Researchers. All researchers complied with all relevant school procedures.

Appendix D: Data protection

For this study, BIT acted as the data controllers, or data controllers in common with schools or SSAT, for personal data shared or collected during the project. BIT processed personal data under the 'legitimate interests' condition (Article 6(1)(f)) of the GDPR. It was necessary in BIT's 'legitimate interests' to process personal data in order to conduct an evaluation of EFA's scale-up that had been commissioned by the EEF. The research project fulfils BIT's core business aims including undertaking research, evaluation and information activities in sectors that will deliver social impact. In line with BIT's privacy by design approach, the researchers asked SSAT to anonymise or pseudonomise MI where possible. The privacy notice for this study is available online at: https://bit.ly/3uxdu8Y.

Appendix E: Project team

Project team member	Project role
Dr Patrick Taylor	Principal investigator
Rizwaan Malik	Lead researcher
Pujen Shrestha	Research and analysis advisor
Dr Neus Torres Blas	Research and analysis advisor
Dr Alex Sutherland	Senior quality assurance
Anna Bird	Quality assurance and senior project advisor
Eleanor Collerton	Project advisor and data support
Kim Bohling	Previous principal investigator