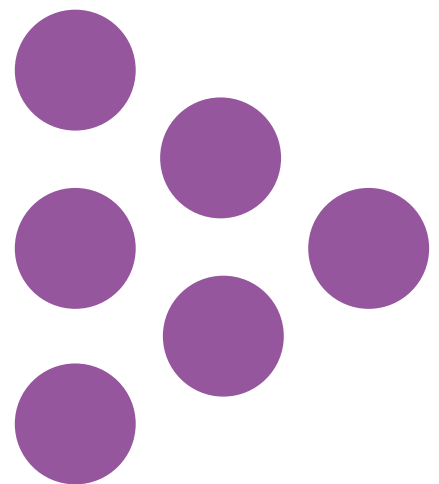

Study Plan

Evaluation of the EEF's Accelerator Fund Pipeline

National Foundation for Educational Research (NFER)

**Sheffield Institute of Education (SloE) at Sheffield Hallam
University**

17th June 2022



Evaluation of the EEF's Accelerator Fund Pipeline

By the National Foundation for Educational Research,
The Mere, Upton Park, Slough, Berkshire SL1 2DQ
www.nfer.ac.uk

Contents

1	Introduction	1
1.1	Description of the AF pipeline activity	1
1.2	Evaluation approach	3
1.3	Structure of the study plan	3
2	Ethics, consent and data management	5
3	Overarching evaluation (OE)	6
3.1	Description of activity	6
3.2	Evaluation objectives	6
3.3	Research questions (RQs)	6
3.4	Research design	7
3.5	Data collection	8
3.6	Analysis and reporting	13
3.7	OE evaluation timeline	15
3.8	Personnel/team	16
4	Early-stage development programmes strand	17
4.1	Description of strand activities	17
4.2	Evaluation objectives	18
4.3	Research questions (RQs)	19
4.4	Research design	19
4.5	Analysis and reporting	22
4.6	Timeline	23
4.7	Personnel/team	24
5	Trials strand	25
5.1	Description of strand activities	25
5.2	Evaluation objectives	29
5.3	Research questions (RQs)	29
5.4	Research design	29
5.5	Analysis and reporting	32
5.6	Timeline	32
5.7	Personnel/team	33

6. Scaling strand	34
6.1 Description of strand activities	34
6.2 The programmes and previous trial outcomes	35
6.3 Evaluation objectives	36
6.4 Research questions (RQs)	37
6.6 Analysis and reporting	39
6.7 Timeline	40
6.8 Personnel/team	41
7 Capability Building strand	42
7.1 Description of strand activities	42
7.2 Evaluation objectives	43
7.3 Research questions (RQs)	43
7.4 Research design	44
7.5 Analysis and reporting	47
7.6 Timeline	48
7.7 Personnel/team	48
8 School Support strand	50
8.1 Description of strand activities	50
8.2 Evaluation objectives	50
8.3 Research questions (RQs)	51
8.4 Research design	51
8.5 Analysis and reporting	55
8.6 Timeline	55
8.7 Personnel/team	56
9 Risks	57
10 Summary GANTT chart and project milestones	62
11 References	62
Appendix A – First ToC draft for Early-Stage Development Programmes strand	63
Appendix B – First ToC draft for Trials strand	64
Appendix C – First ToC draft for Scaling strand	65
Appendix D – First ToC draft for Capability Building strand	66
Appendix E – First ToC draft for School Support strand	67

1 Introduction

1.1 Description of the AF pipeline activity

Throughout the academic year 2021 to 2022, the Education Endowment Foundation (EEF) has been expanding its activities, with the aim of strengthening the development of its pipeline of evidence-based programmes, and to support more schools to use evidence in their practice. The AF pipeline work supports the EEF’s mission to break the link between family income and educational achievement.

The work is supported through the Department for Education (DfE)’s ‘Accelerator Fund’ (AF). This fund aims to support recovery from the Covid-19 pandemic, in this instance, by targeting three poorly affected Regional School Commissioner (RSC) regions¹: the North; East Midlands and Humber; and the West Midlands, supporting schools in these areas to use evidence-based maths and literacy programmes to aid their recovery. The AF funding is being used to enhance the EEF’s usual pipeline activities within a specific period – the academic year 2021-22. The funding will support a range of EEF-funded programmes along the EEF pipeline, from early stage development, through to implementation at scale, by strengthening the EEF’s development, support and implementation processes. For brevity, from this point onwards, we refer to the AF-funded pipeline as ‘the AF pipeline’. The AF pipeline contains five strands, which are summarised below.

Table 1: AF pipeline strands and their stage of development

Description of strands	AF implications for EEF activities
Early-stage programme development: early-stage programmes (thematic) (See Chapter 4)	New design, delivery and formative evaluation led by Research Schools (RSs), supported by EEF and external partners
Trials: promising interventions undergoing efficacy/effectiveness trials (See Chapter 5)	Conduct of trials aligns with EEF usual activity but concentrated recruitment in target regions
Scaling: effective interventions supported to scale in readiness for the next stage of the pipeline (See Chapter 6)	Expansion of usual activity via funding/ support for programme scaling at various stages in the pipeline

¹ Reports published by the DfE on pupils’ learning progress from Spring 2020 to Summer 2021. Available from: <https://www.gov.uk/government/publications/pupils-progress-in-the-2020-to-2022-academic-years?msclkid=96438878b58d11eca9571d88dc833986>

Description of strands	AF implications for EEF activities
Capability building: effective interventions supported to develop further (potentially in preparation for scaling) (See Chapter 7)	New support provided by external partners
School support: training for schools to improve awareness of the EEFs work and signposting to evidence-based programmes (See Chapter 8)	New CPD offers led by RSs (a light-touch version of EEF's usual regional school offer)

As noted in the table above, different strands of the AF pipeline have varied relationships to existing or previous EEF activity. The strand most similar to existing activity is the trials strand. However, in the AF pipeline this strand builds on, but is distinctive from, previous activity in its focused recruitment to trials in a small number of regions around identified areas of need.

Reach

Table 2 shows the reach of AF pipeline activity in terms of the target number of schools for each strand. The numbers in the table represent **'total schools'** rather than 'unique schools', as any one school can take part in more than one strand.

EEF and the DfE agreed that 80% of schools would be drawn from the three RSC regions and that 20% could be drawn from outside the regions. These criteria were set so that two effectiveness trials could recruit a nationally representative sample of schools, and so that one efficacy trial (which is recruiting colleges) could recruit 50% of colleges outside the regions, to increase its pool, given the relatively small number of colleges nationally.

Table 2: Recruitment targets for school participation in AF pipeline strands

Work Strand	End December 2021 (schools)	End March 2022 (schools)	End July 2022 (schools)
Early-stage programme development		75-125	
Trials	200-300	100-150	270-330
Scaling		150-200	300-400
Capability building	N/A	N/A	N/A
School support		200-300	400-600
TOTAL - target number of schools	200-300	800-1,200	1,245-1,755

The target number of ‘unique schools’ by July 2022 is smaller than the figures presented in Table 2 – between 845 and 1155.

There are approximately 6,831 schools across the three regions (1,721 in the North; 2,566 in East Midlands and Humber; and 2,544 in the West Midlands). The target for ‘total school’ reach by end July 2022 is therefore approximately 18-25 per cent. For ‘unique school’ reach, it is approximately 12-17 per cent.

1.2 Evaluation approach

Given the AF pipeline objectives and the short timescale of the project, a formative implementation and process evaluation (IPE) is appropriate.

The evaluation aims to understand the promise, feasibility and scalability of the AF pipeline and gain insight, which EEF can use to develop its work in this area. The evaluation is not able to assess the impact of activities on schools and pupils; however, it will gauge the extent to which AF pipeline activity is beginning to influence schools’ decisions to use evidence-based programmes, or their likelihood of doing so in the future.

The evaluation is structured into six parts – an overarching evaluation (OE) and strand evaluations for each of the five strands of the AF pipeline. Evaluations of the five strands feed into the OE, with the latter also comprising overall evaluation project management activity (see OE tasks below).

NFER and SloE will jointly conduct the evaluation, with NFER as the lead partner. NFER and SloE will jointly conduct the OE and the ‘Early-stage programme development’ strand, NFER the ‘Capability building’ and ‘School support’ strands and SloE the ‘Trials’ and ‘Scale-up’ strands.

The evaluation will be conducted in tandem with the development of the AF pipeline, including of elements that are particularly new for the EEF, such as early-stage programme programmes, scaling and capability support. The evaluation team will thus be learning from EEF’s strand teams – gathering information about intentions and operations as the AF pipeline develops. Thus, the overall approach taken is one of formative and developmental evaluation, particularly in relation to the newer aspects.

As part of this process, the evaluation team will hold early-stage discussions with the EEF strand teams about the type of management information data that will be collected to monitor progress. At the time of writing this study plan, these discussions are only just starting to take place. Hence, it is not possible to precisely pre-specify the types of analysis that will be conducted on this data, although we provide an overview of likely approaches in each of the following chapters.

1.3 Structure of the study plan

This study plan follows a different format to usual EEF study plans, because it is not a typical evaluation. Section 2 focuses on ethics, consent and data management; then sections 3-8 have the following structure for each strand of the pipeline in turn:

- Description of strand activity
- Key evaluation objectives and research questions
- Research design

- Data collection
- Analysis and reporting
- Evaluation timeline
- Details of the evaluation and EEF teams.

Section 3 focuses on the OE. Sections 4, 5, 6, 7 and 8 focus, in turn, on early-stage programme development, trials, scaling, capability building, and school support.

Section 9 summarises study risks and countermeasures. Section 10 provides an overarching study timeline as a Gantt chart.

2 Ethics, consent and data management

Ethics and consent

NFER and SloE, separately, will follow their organisational procedures to ensure all research activity has undertaken ethical review appropriate to their policies and their strand responsibilities. As EEF is the commissioning body, EEF staff are not treated as research participants for the purposes of ethics and consent.

Processes for data collection from non-EEF participants will be subject to NFER ethics review in the overarching evaluation survey and strands 4 and 5. Processes for data collection from non-EEF participants in strands 2 and 3 will be subject to SloE ethics review. For strand 1, undertaken collaboratively, processes for data collection will be subject to review by both organisations following their organisational processes.

Regardless of strand, all participants in the evaluation will be provided with information on the study and a privacy notice relevant to processing their data. Participants can withdraw from data processing at any time during the evaluation and instructions will be provided in the privacy notice for how they can request their data is not processed.

Data protection and management

NFER/SloE will ensure all work streams are compliant with the Data Protection Act 2018 (DPA) and General Data Protection Regulation (GDPR). Both organisations have ISO27001 and Cyber Essentials Plus certifications and registration with the Information Commissioner's Office.

The DfE and EEF are joint data controllers who have overarching responsibility for the Accelerator Fund programme. NFER and SloE will be independent data controllers with responsibility for the evaluation. They will make decisions about how and what personal data is used in accordance with the purposes set by the DfE and EEF.

NFER has identified specific grounds, known as legal basis, for its processing of personal data in the course of the evaluation. NFER has identified 'legitimate interests' as the legal basis. The evaluation fulfils NFER's core business purposes of undertaking research, evaluation and information activities and has broad societal benefits by contributing to improving the educational experience of children and young people. SloE has identified 'public task' as its legal basis for processing of personal data in the course of the evaluation.

The evaluation will not involve the collection or handling of special category personal data.

All personal data collected in the evaluation will be transferred to the EEF data archive within three months of submitting the final report.

DfE are kept informed of the progress of this evaluation and are involved with developing the research questions. They are consulted on key decisions for the evaluation and are invited to review and provide feedback on key outputs including the study plan, research instruments, Theory of Change and the report.

3 Overarching evaluation (OE)

3.1 Description of activity

See section 1.1 for a description of AF pipeline activity.

3.2 Evaluation objectives

The OE has one key aim – to consolidate data from each AF strand and provide an overall picture of the capability of the AF pipeline to meet its desired outcomes. The specific outcomes, discussed with the EEF and DfE, have been identified as follows:

1. **Process-level outcomes** – improved EEF development, support and implementation processes.
2. **School-level outcomes** – enhanced school awareness, acceptance and take-up of evidence-based programmes. There will be variation in the extent to which this outcome is feasible, according to the strand of pipeline activity.
3. **Learning ('design')-level outcomes** – enhanced learning for the EEF about process effectiveness and activities that need to be refined to improve performance across EEF's wider work (with particular regard to new processes such as early development work and capability building).

This aim will be addressed through the following evaluation steps:

1. Develop an overarching AF-pipeline theory of change (ToC).
2. Develop core questions that will guide the development of research instruments across all strand evaluations. This will ensure consistency/comparability of strand data, whilst allowing scope for bespoke areas of questioning, as relevant to each strand.
3. Coordinate and manage the work of the strand evaluations. The OE leads (see Section 3.6) will coordinate external-facing work, liaison with EEF, and the work of the strand leads/research team. They will also quality assure all work and outputs.
4. Conduct cross-strand analysis and synthesis.
5. Produce findings. There will be a summative findings presentation in September and a single final report in October. The report will draw together strand and OE findings (which will each be presented individually to EEF through a series of PowerPoint presentations in the September meeting) into a single research report.

3.3 Research questions (RQs)

In order to address the OE aim, we will explore the following RQs:

OE RQ1: How was the AF pipeline developed, supported and implemented? Was this with fidelity to its intentions? (This RQ links to the process-level outcomes)

- OE RQ1.1: Overall, were strands developed, supported and implemented as intended and did strands progress as expected?
- OE RQ1.2: What factors helped or hindered the delivery of strands? Which of these were common or specific to individual strands?

OE RQ2: What progress did the AF pipeline make towards its intended short-term outcomes (school awareness, acceptance and take-up of evidence-based programmes)? (This links to the school-level outcomes)

- OE RQ2.1: How many schools in AF regions signed up to EEF programmes or used EEF resources (across strands)? To what extent was there variation across, and between, the three AF regions?
- OE RQ2.2: What were schools' reasons for getting involved or not getting involved? What might encourage non-engaging schools to get involved? To what extent did this vary by strand?
- OE RQ2.3: How did schools in the AF regions respond to EEF programmes and resources? How beneficial did they find the process, and the programmes/resources/training? To what extent did this vary by strand?

OE RQ3: What progress did the AF pipeline make towards its intended medium-term outcomes (indicative impact on school behaviour/decision making)? (This links to the school-level outcomes)

- OE RQ3.1: What indications are there that AF-funded activity is beginning to influence schools' decisions to use evidence-based programmes, or may do so in the future?
- OE RQ3.2: What common and/or strand-specific factors contributed to this?

OE RQ4: How effective is the AF pipeline model? (This links to learning (design)-level outcomes)

- OE RQ4.1: Does the model support its objectives to improve the uptake of evidence-based programmes, and schools' decisions to use evidence-based programmes?
- OE RQ4.2: How effectively do the different strands build on or link with each other? Are there any tensions or weak progression points?
- OE RQ4.3: To what extent is cohesion between strands an important factor in supporting future progress on the pipeline and improving school uptake of programmes? What mechanisms (for example capability building or external ToC development support) enable or hinder effective cohesion?
- OE RQ4.4: What lessons are there for the usual EEF pipeline model in relation to improving the quality and impact of funded programmes?

We have agreed with the EEF not to have a RQ related to the potential longer-term outcomes of the AF pipeline, because these cannot be measured within the lifespan of this evaluation.

3.4 Research design

The OE will primarily synthesise evidence gathered through the strand evaluations, rather than collecting data itself. However, there will be a discrete OE school survey, which will gather data from participating² and non-participating schools in the three regions (see Task 4 below).

² All schools involved in any of the AF pipeline strands, or in more than one strand.

3.5 Data collection

Initial theory of change (ToC) development meetings with EEF and DfE

We started the evaluation with two virtual meetings with members of the EEF's AF pipeline strand teams and DfE colleagues, to scope a ToC for the AF pipeline. The aims of these meetings were to:

- articulate the underlying theory of the AF pipeline (including assumptions underpinning the sequencing and logic of intended outputs and outcomes)
- make causal mechanisms explicit
- identify contextual and moderating factors that may impact on AF pipeline effectiveness, the conditions that are important for success and weaknesses or tensions that could stall progress
- identify timings for the achievement of outputs or outcomes within strand
- identify appropriate RQs for the OE.

Further development of ToC

In March/April, we will undertake a scan of practice and theory approaches, including examples of existing pipeline evidence (see for example, Asmussen *et al*, 2019; Coldwell, Maxwell and Sharples, under review; Gough, Maidment and Sharples, 2018; Martin and Williams, 2019). This, and the outcomes of the two meetings already held, will inform a draft ToC for the AF pipeline, including a logic model depiction, which we will share with the EEF and DfE for comment. Following this, we will hold a workshop in the summer with EEF and DfE colleagues (including the NFER and SloE study advisors (see Section 3.6) to refine and finalise the AF pipeline ToC.

There will be a final opportunity to revisit the ToC in the autumn, towards the end of the evaluation, during or around the time of the proposed September feedback event (see Section 3.4).

Design of core questions

In parallel with the continuing ToC discussion, we will develop core questions that will guide the development of research instruments across all strand evaluations. These questions will address themes that are common and need to be explored across all the strands. For example: What was delivered in the strand during the timescale of the AF pipeline? Was this in line with expectations? Was the strand delivered with fidelity to its intentions? To what extent did the strand meet its key evaluation objectives? What were the mechanisms, facilitators, barriers and contextual issues influencing the promise, feasibility and scalability of the approach?

These questions will ensure consistency/comparability of strand data, whilst allowing scope for bespoke areas of questioning, as relevant to each strand. Questions will be shared with, and approved by, EEF and the DfE prior to being converted into appropriate research instruments by each strand lead and research team.

Design and administration of a school survey

In order to collect data that will contribute to answering RQs 2 and 3, we will design and administer a short single time-point survey with all primary, secondary and all-through state-

funded schools in the three AF pipeline regions in the summer term, 2022³. This survey will be designed in collaboration with the EEF – in particular, drawing on the expertise of the Dissemination and Impact team.

Survey design

We anticipate that the survey will contain two short banks of questions:

1. The first will aim to deduce the relative importance of different sources of information in school leaders' decision making (including, but not limited to, evidence-based sources). These questions will become gradually more specific – moving from the generic (broad source 'types', for example government guidance, academic research) to the specific (named programmes or interventions – including, but not limited to, those produced by the EEF). This will help us assess how important evidence-based sources and programmes are, relative to other sources of information, in school decision making, thereby helping to answer RQ3. In our experience, this elicits more useful information than simply asking respondents whether or not they use these types of sources (which tends to encourage socially-desirable responses). In this type of design, it is important not to 'prime' respondents, so we would advise against branding the survey as an EEF, or AF pipeline survey, specifically.
2. The second will focus on school leaders' awareness, and degrees of 'use' of, EEF resources, opportunities and programmes. NFER designed a survey for the EEF in 2018, which explored schools' awareness of a variety of EEF resources. These included, for each question, an EEF website screenshot of the resource or programme in question and then a simple question in the following format (this example was of a guidance report).

Before this survey I had [Please tick one box only]

- a. Never seen this report
- b. Seen this report, but not read it
- c. Seen this report and read it
- d. Seen and read this report, and used it to support my practice.

We could use a format similar to this, or there are a number of alternative question formulations that we could consider (for example, levels of awareness or use on a Likert scale). Analysis of responses to these questions will help to answer RQ2.

The purpose of inviting *all* schools in the three regions to participate is to enable analysis of survey responses according to whether or not schools were included in AF pipeline activities and to consider any variation in response. As mentioned in Section 1.1, there are approximately 6,831 schools across the three regions (1,721 in the North; 2,566 in East Midlands and Humber; and 2,544 in the West Midlands). Of these, EEF anticipate that 1,245-1,755 schools in total may engage with one or more AF pipeline strands, although 20 per cent of these schools may be outside the three regions being covered by the survey. Each school will be linked to NFER's Register of Schools (ideally by URN) to enable comparison of the two samples for equivalence, and to assess whether any school characteristics are

³ One efficacy trial targets early years and another targets post-16 colleges so the survey will not cover these demographics but this is a very small amount of activity in the context of AF activities.

associated with higher or lower levels of awareness and use (we will pre-specify these characteristics with the EEF).

The survey will be designed in April, built and tested in May and administered in June (kept open, if necessary, until the end of the summer term – mid July). Further details on the analyses that we will undertake are outlined in Section 3.4. Information on administration of the survey is provided below.

Survey administration

We anticipate that the survey will be sent to one individual in each school, ideally a school senior leader (unless, in the case of schools taking part in AF pipeline activities, the key contact in the school is a non-senior leader (for example, a literacy or numeracy coordinator). We will ask the EEF to supply us with details of the schools engaged in AF pipeline activities, along with key contact information. The key contact details will enable us to liaise directly with someone in the school who is engaged in the AF work, and this will help to achieve our predicted response figures.

From our current work in schools we are aware of the on-going pressures and challenges with regards to managing both staff and pupil absences, which has been effecting schools' ability to participate in research projects. Our current intelligence (based on our monitoring of response-rate data from recent NFER surveys of similar potential respondents) is that we can anticipate approximately a ten percent response rate to the survey. We also predict that the response from primary schools will be higher than that of secondary schools – this is something that has always been the case when running a survey across both phases.

The survey will open shortly after the May half-term (Wednesday June 8) and will remain open until just before the end of the summer term (Friday July 15). During this five-week period we will send out an initial email invitation containing a survey link, followed, on June 17, with an online reminder then, on June 23, with a postal reminder, containing both a survey link and a paper copy of the questionnaire with pre-paid return envelope. Our final two reminders will be via email only during weeks beginning July 4 and July 11.

EdCo will support the NFER by providing an additional sample source for the senior leader survey. They provide sales and marketing support to the education supply sector, including direct mail to schools through named contacts. A decision was made to use EdCo to enhance the survey response. Using their contact database, EdCo's role is to invite a defined group of senior leaders within the three named RSCs to complete the online survey. EdCo do not guarantee a response as they are a list rather than a research panel. The EdCo sample will be given two weeks to complete the survey, which was recommended by EdCo, and they will send one initial invite and one reminder. Following this, NFER will use their register of schools and a list of schools and contacts provided by EEF to contact all schools in the defined areas who have not completed the survey. This will involve inviting schools to complete either a paper or online survey. The survey will be open for a further four weeks with a reminder going out each week.

Each school will receive their own unique survey link to allow us to monitor schools' responses and remind appropriately. The unique link will also allow us to follow up with any schools who start but do not finish the survey. By using a unique survey link, we can also match the survey data with specific school characteristics.

By using a mixed-method survey approach, we will aim to increase the responses across all schools phases. Our initial email will go to either the contact person named in the details received from the EEF or, for schools not in the list, we will use NFER's Register of Schools database, which contains the main school email address. The invite will then ask for the survey to be passed on to the most relevant person. The invitations will highlight the importance of the research and why we need schools like theirs to participate. All schools responding to the survey will also be eligible to be entered into a prize draw to win one of ten £200 book tokens or Amazon vouchers for their school.

Following the launch of the survey, we will review the survey statistics on a daily basis noting if there are any particular questions where teachers are dropping out of the survey, and also keeping an eye on the average survey length. This information will be passed directly back to the survey Lead. Weekly progress reports will also be shared with the EEF, which will highlight the number of surveys completed; additional breakdowns for the responses can also be provided in this report.

The survey will be programmed in our survey software Questback. The survey will be mobile enabled so that respondents can complete on all devices.

For those schools that choose to complete the survey on paper, they will be able to return the questionnaire in a pre-paid envelope, which will be provided with the questionnaire. Once returned to the NFER the questionnaire data will be captured and then later both online and survey responses will be merged together into the final data set.

On any occasion that we receive more than one survey response from a school (whether that be more than one online response, more than one paper response, or a mix of online and paper responses), we will apply clear criteria for prioritising which survey response to process. Our quality assurance checks will ensure that we only process one survey response for each school URN. Our criteria will be as follows. We will give priority to the individual's seniority, irrespective of survey type (online or paper). A headteacher/principal will have priority over a school senior leader, a school senior leader over a middle leader and so on. In the event that we receive more than one response from individuals with the same level of seniority, we will prioritise an online over a paper response. This is because the online survey technology enables clearer routing and contains formulae, which ensure that respondents answer only appropriate questions. If the mode of completion is the same, we will prioritise the most complete survey, or if there is no distinction in completeness, the first returned survey.

Table 3 provides a summary of the primary data collection methods that will be used across the strand and OE evaluations, mapped against the OE key research questions. Further details on each of these tasks, as well as other activities relating to the research design, are provided below.

OE RQ1: How was the AF pipeline implemented? Was it implemented with fidelity to its intentions? (This RQ links to the process-level outcomes)

OE RQ2: What progress did the AF pipeline make towards its intended short-term outcomes (school awareness, acceptance and take-up)? (This links to the school-level outcomes)

OE RQ3: What progress did the AF pipeline make towards its intended medium-term outcomes (indicative impact on school behaviour/decision making)? (This links to the school-level outcomes)

OE RQ4: How effective is the AF pipeline model? (This links to learning (design)-level outcomes).

Table 3: Data-collection methods mapped to Research Questions

	Strand qualitative interview and observation data (see strand study plans in Sections 4-8)	Strand MI data analysis (see strand study plans in Sections 4-8)	OE survey
OE RQ1	✓	✓	
OE RQ2	✓	✓	✓
OE RQ3	✓	✓	✓
OE RQ4	✓		

3.6 Analysis and reporting

Synthesis of strand data

We will synthesise findings from the five strand evaluations using a thematic framework derived from the OE research questions (and the core question areas). We will also summarise strand-specific outcomes, drawing together common learning across strands in order to address the OE research questions. Our synthesis will be informed by an internal whole-team analysis meeting, five strand-level analyses and PowerPoint slide packs, which will be presented to EEF during an event in September, plus analysis of the OE survey data.

Survey analysis

On receipt of the completed survey responses, initial exploration will compare the survey respondents to the population of schools within the three regions of interest. Any statements made about the survey responses toned to be generalizable to the population the sample was drawn from. We will analyse the samples gained for both schools participating in AF pipeline activities and non-participating schools and compare against the known characteristics of these populations (e.g. phase, proportion of FSM pupils at school, region). This analysis will help us understand if there are specific types of schools that have not responded to the survey, which might introduce some bias into findings. The use of a Chi-Square test will identify any differences between these two groups. If there are significant differences then a decision will be made as to whether the sample of respondents will be weighted so that the sample matches the population on these characteristics. A degree of caution will be required when interpreting any differences between the two groups of schools, as we do not know what the population of participating and non-participating schools actually look like. This will make generalisations difficult and so care will be taken when describing differences to ensure there is no misinterpretation of the data. Response rates, particularly for non-participating schools, are likely to be low and so analysis using school level characteristics will only take place if cell counts are of sufficient size to support such analysis.

Analysis will test the hypothesis that schools taking part in EEF pipeline activities, are more likely than other schools in the region to be:

- a) using research evidence to inform their decision making
- b) aware of (and potentially using) EEF resources and opportunities.

Once the survey design is agreed and clear, we will compare survey responses by the three school-level characteristics identified above, as well any other factors the research teams, the EEF and DfE feel would be useful. A degree of care should be taken with how many characteristics are used to highlight variation in response, as this can become difficult to interpret. In addition to these comparisons we will also compare responses broken down by respondents who have engaged with one or more elements of the EEF pipeline, versus those that have not.

If the dataset is large enough to support a more granular level of analysis, we will compare survey responses by pipeline strand to identify variation in response. Consideration will be made for schools that belong to more than one strand.

In October, following the strand findings workshops in September, we will synthesise the strand findings, combining this analysis with the survey findings to form an overarching evaluation report, with executive summary, which will be submitted in draft form by the end of October 2022 for EEF, DfE and peer review.

The reporting format is still being agreed with the EEF, so it is not possible to be precise, at this stage, about the analysis framework that will underpin the report. However, we anticipate that it will be guided by the following principles.

The analysis framework will:

- prioritise the OE RQs and core strand research questions (see Chapters 4-8 for RQs that are common across all strands)
- seek to identify strand-specific learning through analysis of core strand research questions and strand specific questions
- carefully map data sources to RQs to ensure that there is full thematic triangulation of qualitative and quantitative data (interview data; survey data; MI data).

Reporting will

- focus on the OE RQs and core strand research questions
- present findings relating to the bespoke strand RQs separately (possibly via a short chapter, or annexe, for each strand) including illustrations or vignettes
- summarise common learning across both the core and bespoke RQs to provide recommendations to the EEF and DfE about effective implementation approaches, progress towards intended outcomes, and the viability of the pipeline model itself
- include detail on the relationship of data sources to RQs and analysis (possibly as an appendix).

3.7 OE evaluation timeline

Table 4: OE evaluation timetable

Month	Activity	Who
January - February	Inception meeting Two ToC development meetings	NFER/SIoE/EEF NFER/SIoE/EEF/DfE
March	Develop privacy notice, evaluation information sheet and data-sharing agreements Obtain ethical approval Develop core interview questions (based on common RQs) to be asked across all strand interviews Scan of practice and theory approaches	NFER/SIoE NFER/SIoE/EEF/DfE NFER/SIoE NFER/SIoE
April	Scan of practice and theory approaches continues ToC refinement workshop ToC revision Survey design	NFER/SIoE NFER/SIoE/EEF NFER/SIoE/EEF NFER/SIoE/EEF
May	Survey build and testing	NFER
June	Survey live	NFER
July	Survey close	NFER
August	Survey data cleaning and analysis Synthesis of strand data	NFER NFER/SIOE
Sept	Synthesis of strand data continues Survey interpretation and report drafting Findings workshop	NFER/SIoE NFER/SIoE NFER/SIoE/EEF/DfE
October	Draft report submitted (end of month)	NFER/SIoE
November-February	Report review, peer review and publication	NFER/SIoE/EEF

3.8 Personnel/team

Table 5: OE Evaluation team

Name	Role/responsibility in this strand	Organisation
Suzanne Straw ⁴	Co-principal investigator	NFER
Mark Boylan	Co-principal investigator	SloE
Ben Styles	Specialist Adviser	NFER
Mike Coldwell	Specialist Adviser	SloE
Nick Moore	Project management	SloE
Kelly Kettlewell (Research Associate)	Survey development	NFER

Table 6: OE EEF team

Name	Role
Celeste Cheung	Senior Evaluation Manager
Amy Clark	Senior Programme Manager
Iggy Rhodes	Head of Programmes
Stu Mathers	Head of Dissemination & Impact

⁴ From January to May 2022, the NFER Principal Investigator was Julie Nelson.

4 Early-stage programme development strand

4.1 Description of strand activities

EEF has identified a need to improve its support for the design, development and early formative evaluation of early-stage education programmes. This strand of AF pipeline activity adopts an exploratory approach to addressing this need. It aims to build and develop EEF's expertise and to support further learning, particularly about the potential benefits of external support using two models. The first is the Rapid Cycle and Development Testing model (RCDT) developed by Dartington Service Design Lab ('Dartington'). The second is the Behaviour Change Wheel based on the COM-B model (BCW), developed and used by Centre for Behaviour Change at University College London (UCL) ('CBC').

Strand activity is based around nine Research Schools (RSs), which act as programme developers, each developing a single programme. Dartington is supporting all nine RSs. CBC is providing additional support to two of these RSs, which were identified by EEF as having greater capability to extend their application of learning by using both the Dartington and CBC models and receiving two sets of support input to the development of their programmes. The EEF's Head of Dissemination and Impact and two early pipeline development leads provide most of the support to the RSs for programme development, with an element of Dartington building EEF's capacity to provide support.

The EEF identified gaps in programme availability and invited RSs to apply to develop programmes that addressed one of these areas. The strand is operating on a short time scale. This means that programmes developed must be ready for implementation in schools in the summer term 2022. RSs may be delivering their early programmes outside of their home region. Details of the support available to the RSs for programme development work are provided below.

Table 7: Support provided to Research Schools

Support provider	Summary details of support and contribution
EEF	<ul style="list-style-type: none"> Identify foci areas for programmes to address priorities ('gaps') Identify appropriate support partners and development tools Identify and select RSs and potential programmes Tailored support to each RS lead for programme development and implementation including mentoring and coaching as needed Provide programme implementation documentation and tools Support RS recruitment, including by drawing on support of RS network
Dartington	<ul style="list-style-type: none"> Rapid Cycle and Development Testing model (RCDT) Workshops for all nine RSs to support use of the RCDT and develop programme Theory of Change (ToC) Action learning support focusing on implementation

Support provider	Summary details of support and contribution
	Supporting RSs to conduct early evaluation of programme implementation and design through drop-in clinics on data analysis and evaluation and leading on reflection discussion
CBC	The BCW based on the COM-B model Workshops for two RSs focused on the BCW approach

The programme delivery takes place in the following phases:

Stage 0: Set-up (Oct/Nov 2021)

Stage 1: Assess (Nov 2021- Feb 2022)

Stage 2: Design (Feb 2022 – Apr 2022)

Stage 3: Deliver & Observe (Apr 2022 – Jul 2022)

Stage 4: Analyse & Learn (Jul 2022 – Aug 2022)

Step 5: Pause & Decide (Sept 2022 – Oct 2022)

Workshop input by Dartington and CBC takes place during Stage 2.

Summary of the workshop inputs:

Dartington:

- Stage 1 - Developing a Theory of Change, programme blueprint, and evaluation plan
- Stage 2 - no workshops, some support provided for QA of RS outputs above
- Stage 3 - Implementation support
- Stage 4 - Data analysis support
- Stage 5 - no workshops, but Dartington will have a structured/reflective discussion with each RS.

Workshop inputs by CBC took place in Stage 2 and focused on applying Behaviour Change Wheel tools to early-stage education programme development.

4.2 Evaluation objectives

Objective 1: refine the ToC to set out the underlying assumptions, mechanisms and context underpinning the inputs, outputs and anticipated outcomes of the strand (see draft ToC provided by EEF in Appendix A)

Objective 2: explore implementation, key enablers and barriers, and progress of this new model – through case studies of nine Research Schools (RSs)/developers, each including interviews with RS leads, support partners (Dartington and CBC), and EEF staff.

Objective 3: review programme and support materials and MI data.

Objective 4: undertake both inductive and deductive analysis (including triangulation and comparison of different programmes) to evaluate the development programmes strand against objectives and in relation to research questions.

Objective 5: identify the extent to which the strand has met its objectives and, to support future developments, revisit and refine the ToC.

Objective 6: summarise formative findings for presentation.

The focus for this strand evaluation is on the support provided for the design and development of early-stage programmes. Participant school engagement in programme development activities and school-level outcomes are out of scope for this strand of the evaluation.

4.3 Research questions (RQs)

Core to all strands:

RQ1.1 What was delivered in the strand during the timescale of the AF pipeline? Was this in line with expectations? Was the strand delivered with fidelity to its intentions?

RQ 1.2 To what extent did this strand meet its key objectives, as relevant to the evaluation?

RQ 1.3 What were the mechanisms, facilitators, barriers, and contextual issues influencing the promise, feasibility, and scalability of this approach?

Bespoke to this strand:

RQ 1.4 To what extent did research schools believe the programme designed under this strand addressed practice/evidence gaps?

RQ 1.5 What were the benefits and limitations of the support and ToC planning models, tools and activities used by Dartington and CBC to support programme development?

RQ 1.6 What was effective and what could be improved in the EEF team's support for RSs/developers and their work with support partners?

RQ 1.7 What lessons are there for EEF and others supporting programme development in the future to increase the likelihood of programmes proceeding to pilot stage?

4.4 Research design

4.4.1 Summary

The research design will primarily take the form of a process evaluation, examining how the early-stage programme development support was conducted. It will focus specifically on the promise, feasibility and scalability of the support provided through this strand.

Prior to commencement of data collection, the research team reviewed an initial strand ToC that had been prepared by the EEF strand team (this initial ToC can be found at Appendix A). The researchers then held an inception meeting with the EEF strand team (see Table 11) to gain an in-depth understanding of the strand's activities and processes, and the outcomes anticipated in the longer term (which are outside the scope of this evaluation). The information gained from the meeting was used to begin a process of ToC refinement, to identify key strand RQs, and to understand the type of MI data or reports being collected by the EEF team, as well as the process for sharing this data.

4.4.2 Research instruments

Following this inception meeting, the research team is now developing a series of semi-structured interview schedules, covering a number of topics to align with the strand RQs. Each interview schedule will contain a small number of 'core interview questions'. These will align with the 'core RQs' outlined above and will be applied across all interview schedules across all five strands. Additionally, the schedules will contain bespoke strand-specific questions to reflect the themes covered by the bespoke RQs above.

For this strand, seven separate interview schedules will be developed for four interviewee types (EEF strand leads – initial and follow-up schedules), Dartington (initial and follow-up schedules), CBC (a single schedule) and RS project leads (initial and follow-up schedules). All interview schedules, with the exception of those for use with the EEF team, will be shared with EEF and DfE for comment before they are used for data collection. We will share agenda/discussion themes with the EEF team in advance of interviews with them.

4.4.3 Data collection

Interviews with EEF strand leads

We will conduct interviews with EEF strand leads on two occasions.

- The first interview will be conducted in April and will be a one-hour semi-structured group discussion with the three strand leads (the Head of Dissemination and Impact and the two early pipeline leads). This interview will aim to gain more insight into the design of the strand, its aims and intended outcomes and reflections on progress so far (RQs 1.1, 1.5 and 1.6).
- The second interviews will be conducted in June and will comprise two individual 75 to 90 minute discussions with each early pipeline lead. Each interview will focus in detail on the progress and experiences of the specific RSs and programmes that each early pipeline lead is working with. This interview length will allow approximately 15 minutes for questions and discussion about each RS, as well as allowing time to collect the early pipeline leads' reflections on progress and lessons learned across the strand as a whole (RQs 1.1, 1.2, 1.3, 1.5, 1.6 and 1.7).

These interviews will supplement the initial ToC discussion meeting (which was held in February) and ongoing ToC development work. The ToC will be revisited and refined as necessary in the summer towards the end of the evaluation.

Interviews with support partners

We will conduct 45-60 minute interviews with Dartington on two occasions. These interviews will be conducted as one-to-one or group discussions, as appropriate to the make up of the Dartington team.

The first interview will be conducted in April or May (depending on availability around the school Easter holidays) and will collect data on: the RCDT process and how it is being used to generate robust ToCs and robust ToC steps; Dartington's overall support approach; and their initial training support (RQs 1.1, 1.5 and 1.6).

The second interview will be conducted in July and will focus on Dartington's on-going support and action learning sets as well as reflections on the process of providing support for RSs and working with the EEF (RQs 1.1, 1.2, 1.3, 1.5, 1.6 and 1.7).

We will conduct a single interview with CBC. This will be an interview or group discussion of 45-60 minutes, as appropriate to the make up of the CBC team. It will focus on CBC's initial training support for the two RSs that are receiving CBC, in addition to Dartington, support. This interview will be conducted in May (RQs 1.1, 1.5 and 1.6).

Interviews with RS project leads

We will conduct 45-60 minute one-to-one interviews with the project lead from each of the nine RS on two occasions (18 interviews in total). Interviews with the two RSs receiving support from both CBC and Dartington are likely to be 60 minutes in duration

- The first nine interviews will be conducted in April or May (depending on availability around the school Easter holidays) and will focus on the Research School project leads' views on the identification of gap areas, and their experience of programme support and how this has influenced programme development (RQs 1.1, 1.4, 1.5 and 1.6).
- The second nine follow-up interviews (with the same individuals) will be conducted in July allowing an opportunity for reflection on the ongoing support, and on the experience of developing early-stage (RQs 1.1, 1.2, 1.3, 1.5, 1.6 and 1.7).

Review of support materials and tools, and early-stage development programme materials

We will collect and review support materials and tools produced for developers by Dartington, CBC and EEF to support analysis of interview data. We also collect and review programme support materials developed by RSs to analyse in what ways the project support has influenced programme design (RQs 1.1, 1.2, 1.3, 1.5 and 1.7). For detail of analysis and reporting approach, see Section 4.5.

Review of EEF MI data/progress reports and satisfaction data

We will collate MI and satisfaction data collected by EEF as well as data recorded in monthly progress reports submitted by programmes during the period of support. This data will relate to quantitative outputs such as deliverables completed for each programme, attendance at workshops and whether scheduled work is on/off track. MI used will also include qualitative data such as reflections on activities conducted and challenges/lessons. We are in the process of mapping out what MI data is available for each strand and the extent to which it will need to be analysed by NFER/SIoE, but we envisage that this data will provide insights into the extent to which the strand objectives were achieved, and capture challenges and barriers experienced throughout the project (RQ 1.1-1.3). For detail of reporting and analysis approach, see Section 4.5.

RQ1.1 What was delivered in the strand during the timescale of the AF pipeline? Was this in line with expectations? Was the strand delivered with fidelity to its intentions?

RQ1.2 To what extent did this strand meet its key objectives, as relevant to the evaluation?

RQ1.3 What were the mechanisms, facilitators, barriers, and contextual issues influencing the promise, feasibility, and scalability of this approach?

RQ1.4 To what extent did research schools believe the programme designed under this strand addressed practice/evidence gaps?

RQ1.5 What were the benefits and limitations of the support and ToC planning models, tools and activities used by Dartington and CBC to support programme development?

RQ1.6 What was effective and what could be improved in the EEF team’s support for RSs/developers and their work with support partners?

RQ1.7 What lessons are there for EEF and others supporting programme development in the future to increase the likelihood of programmes proceeding to pilot stage?

Table 8: Data-collection methods mapped to RQs

	Interviews with EEF team	Interviews with early pipeline leads	Interviews with support partners	Review of materials	Interviews with RS leads	Review of MI data
RQ1.1	✓	✓	✓	✓	✓	✓
RQ1.2	✓	✓	✓	✓	✓	✓
RQ1.3	✓	✓	✓	✓	✓	✓
RQ1.4					✓	
RQ1.5	✓	✓	✓	✓	✓	
RQ1.6	✓	✓	✓		✓	
RQ1.7	✓	✓	✓	✓	✓	✓

4.5 Analysis and reporting

We will audio record all of our interviews (following participant agreement) and collate data thematically. We will initially use a deductive analysis approach by devising a broad set of codes aligned to the research aims and questions. However, we will be driven by the data, adding key sub-codes arising from the data which were not present in the initial framework to ensure emergent findings are identified and to add detail to the analysis, including accounting for variation across RS/early programmes.

The strand analysis will include consideration of the two RSs receiving additional CBC support as a pair within the overall set of nine RS. This will include analysis of support and training materials and will synthesise findings from interviews with the EEF team.

For those research questions addressed by multiple data sources will triangulate EEF, RS and support partner interview data, along with the MI and documentary analysis – as appropriate to the research questions. Further, we will compare and contrast views,

experiences and practices for each early programme to provide a comprehensive assessment of the implementation of the strand. We will produce a slide pack presenting the key findings of the evaluation.

4.6 Timeline

Table 9: evaluation timetable

Month	Activity	Who
March	Preparation of recruitment documents Refinement of logic model/ToC Design of research instruments	NFER/SIoE NFER/SIoE/EEF NFER/SIoE
April-May	Interview with EEF programme managers (early pipeline development leads and HoD) (overall strand focus and discussion re collection of MI data) Interviews with Dartington and CBC Interviews with RS leads Collection and analysis of programme documentation	NFER/SIoE/EEF NFER/SIOE/EEF /Dartington/CBC programme managers NFER/SIoE/Research schools NFER/SIoE
June	Interviews with EEF early pipeline development leads – RS focus	NFER/SIoE/EEF
July	Interviews with Dartington Interviews with RS leads Collection and analysis of MI and satisfaction data	NFER/SIoE/Dartington NFER/SIoE/Research schools NFER/SIoE/EEF
August	Data analysis	NFER/SIoE
September	Slide pack of presentation of findings	NFER/SIoE/EEF/DfE

4.7 Personnel/team

Table 10: Evaluation team

Name	Role/responsibility in this strand	Organisation
Juliet Sizmur	Strand Lead	NFER
Mark Boylan	Strand Co-Lead	SIOE
Caroline Stanford (Research Associate)	Data collection and analysis	NFER
Nick Moore	Data collection and analysis	SIOE

Table 11: EEF team

Name	Role	EEF department
Mia Travers-Hayward	Early Pipeline development lead	Dissemination and Impact
Sarah Dobie	Early Pipeline development lead	Dissemination and Impact
Stu Mathers	Head of Dissemination and Impact	Dissemination and Impact

5 Trials strand

5.1 Description of strand activities

This strand focuses on the four efficacy and two effectiveness evaluations of six programmes. The strand activities are similar to existing EEF trials. However, as noted in Section 1.1, this strand builds on, but is distinctive from, previous activity in its focused recruitment to trials in a small number of regions around identified areas of need.

The two effectiveness trials (see Table 12) have been funded using a mix of EEF and DfE AF funding. Both evaluations were underway before the AF and so, unlike the efficacy trials, have recruited nationally. Both evaluations are underway, with one due to complete implementation at the end of 2021/22 (Improving Working Memory) and the other at the end of 2022/23 (Children's University).

The four programmes were selected for efficacy trials (see Table 13), and new evaluators commissioned for the efficacy re-trials, with set-up and IDEA meetings occurring before the start of the strand evaluation. During the AF pipeline evaluation period the following activities will be taking place:

- school recruitment
- publication of re-trial protocols for two of the new evaluations (English Mastery & 5Rs)
- randomisation (before the end of the 2021/22 academic year).

Thus, strand outputs during the evaluation period will be the production of protocols for AF re-trials and recruitment of schools. Following this evaluation, the implementation and evaluation of the four programmes will take place.

The strand has the following common objectives for both efficacy and effectiveness trials -

- Recruit between 300 and 400 schools.
- Increased awareness of EEF trials in schools in the three regions
- More pupils receiving evidence-informed programmes in the AF three regions (longer term objective beyond the initial AF pipeline funding)

The efficacy trials aim to -

- Provide the opportunity, for some projects, to refine their programme to enable further scale in the future e.g. through EEF effectiveness trials
- Give concurrent support for programmes to scale further whilst maintaining quality (capability building)
- Increase awareness among delivery teams and the EEF about important implementation factors for scaling under different contexts

In addition, the effectiveness trials aim to -

- Commission delivery and evaluations of up to two programmes that show evidence of raising pupil attainment from previous EEF trials
- Produce evaluation reports with new evidence on the effectiveness of the programmes and lessons learned
- Increase availability of programmes ready for scale or further refinement.

Table 12: The two effectiveness trials

Project	Focus/features	Timeline details	Previous Evaluation details
Improving Working Memory+ Arithmetic (IWM+A)	Designed to improve the numeracy skills of 7-8 year-old children who show low attainment in arithmetic at the end of KS1. The intervention seeks to improve pupils' working memory, as well as the way they organise information about numbers and operations in their long-term memory.	Project sets-up & recruitment (national) took place prior to AF funding. Delivery will be completed but not reported during evaluation period.	Efficacy (+3 months; 4 pads; 1 £)
Children's University	Children's University centres support schools to provide a range of learning opportunities, such as after-school clubs, projects and enrichment activities, and visits to destinations such as libraries, sports clubs, historic centres, museums, or anywhere that offers structured learning activities for children.	Project sets-up & recruitment (national) took place prior to AF funding. Delivery will be ongoing until Summer 2023 and not reported during evaluation period.	Efficacy (+2 months; 3 pads; 1 £)

Table 13: The four efficacy trials

Project	Focus/features	Timeline details	Previous Evaluation details
Learning Language & Loving It (LLLI) -	Provides training for Early Years practitioners to promote language and early literacy, including practical strategies for helping all children in the classroom build language and social skills. These strategies include encouraging children to initiate and engage in conversational interaction, and modelling more sophisticated language to expand the child’s oral language skills.	Project sets-up and research design completed, recruitment of nursery settings ongoing, protocol summer due in April, delivery from Sept 22	No IE due to Covid, IPE due summer 2022. SAPs & Protocols.
Pupil Assisted Learning Strategies (PALS)	Peer Assisted Learning Strategies (PALS-UK) is a whole-class, structured, paired-reading intervention. Pupils work together in pairs and take turns to act as coach and reader. The aim is to improve reading comprehension and reading fluency.	Project set-up and research design completed, recruitment in primary schools ongoing, protocol due in April, delivery from Sept 21	No IE due to Covid, IPE report published in Feb 225. SAPs & Protocols.
English Mastery	Trains and provides resources to English teachers to deliver a knowledge-rich Key Stage 3 English curriculum consisting of three distinct elements: Literacy Heritage; Mastery Writing and Reading for Pleasure. 2-year programme.	Project set-up and research design completed, recruitment in secondary schools ongoing, protocol due in April, delivery from Sept 21	No IE due to Covid, IPE report due Spring 2022. SAPs & Protocols.

⁵ PALS-UK evaluation report available from: Outcome data from the initial trial found a negative impact on reading of -0.07 overall and a positive impact of 0.05 for pupils eligible for FSM, however, there is considerable uncertainty about the reliability of these results due to delayed post-testing and high attrition (1 padlock). There is a compelling case for an evaluation to identify a programmatic vehicle for getting the benefits of peer tutoring into classrooms. The available evidence for PALS, and the wider evidence on peer tutoring, suggests that it is particularly beneficial for pupils with low prior attainment.

Project	Focus/features	Timeline details	Previous Evaluation details
		Randomising before ITT sample identified.	
The 5Rs Approach to GCSE Maths Resits	A manualised intervention consisting of both a set lesson structure and lesson content, and three short initial diagnostic tests, which are used to determine weaknesses in nine underpinning mathematical skills.	Project set-up and research design completed, recruitment in post-16 colleges, protocol due in April, delivery from Sept 21 Randomising before ITT sample identified.	Planned IE did not take place ⁶ (just Nov 2019 resits prior to Covid). Report published 2021. SAPs & Protocols.

⁶ The cancellation of exams meant the planned end of year attainment data was not available.

5.2 Evaluation objectives

Objective 1: Set out the underlying assumptions, mechanisms and context underpinning the inputs, outputs and anticipated outcomes of the strand, further develop the theory of change.

Objective 2: Undertake light-touch evaluation through interview data collected from developers, evaluators and EEF to explore implementation, key enablers and barriers, and progress within the strand/model.

Objective 4: Explore progress towards target outputs/outcomes in relation to recruitment, undertake analysis of relevant internal MI data.

Objective 6: Undertake both inductive and deductive analysis (including triangulation and comparison of programmes) to evaluate the development programme against objectives and in relation to research questions

Objective 7: To identify the extent to which the strand has been effective in meeting its aims and to support future developments, revisit and refine the ToC.

Objective 8: Summarise formative findings for presentation.

As noted above, given the relatively short overlap between the strand evaluation time period and the trial time period, the focus of the strand evaluation for efficacy trials is on the early stages of trial design, recruitment and implementation. For effectiveness trials that are currently delivering interventions or collecting outcome data, the focus is more likely to cover any issues experienced at these stages and on participant attrition.

5.3 Research questions (RQs)

Core to all strands:

RQ2.1 What was delivered in the strand during the timescale of the AF pipeline? Was this in line with expectations? Was the strand delivered with fidelity to its intentions?

RQ2.2 To what extent did this strand meet its key objectives for efficacy and effectiveness trials?

RQ2.3 What were the mechanisms, facilitators, barriers, and contextual issues influencing the promise, feasibility, and scalability of this approach?

Bespoke to this strand:

RQ2.4 How effective was recruitment overall and across the six programmes? What aided and hindered recruitment? How do participating settings compare with settings in the AF regions? What insights, if any, do developers have about motivations for engagement?

5.4 Research design

5.4.1 Summary

Given the timescale for the strand evaluation, an IPE evaluation is appropriate, focused on trial design and recruitment to trials. In addition to the development of ToC, the evaluation will have two key methods:

1. Review of programme and trial documentation. MI data relating to school recruitment will be obtained and analysed with respect to recruitment targets
2. Interviews with:
 - a. EEF staff (programme & evaluation managers; regional leads)
 - b. project teams
 - c. project evaluators.

The four developers attached to the efficacy trials are in the rare position of being able to reflect on past experiences in terms of recruitment, ToC and implementation and how this was developed for the new efficacy trial. The two developers attached to the effectiveness trials can also reflect on changes to their approaches to recruitment, ToC and implementation between the efficacy and effectiveness stages.

Prior to commencement of data collection, the research team reviewed an initial strand ToC that had been prepared by the EEF strand team (this initial ToC can be found at Appendix A). The information gained from the meeting will be used to refine the strand ToC and key RQs and to understand the type of MI data or reports being collected and the process for sharing the data, as well as designing the research instruments. During this initial work we will also collect and review programme and trial documentation, including protocols and for the two effectiveness trials, documentation from previous trials.

5.4.2 Research instruments

Following the inception meeting, the research team is now developing a series of semi-structured interview schedules, covering topics to align with the strand RQs. Each interview schedule will contain a small number of 'core interview questions'. These will align with the 'core RQs' outlined above and will be applied across all interview schedules across all five strands. Additionally, the schedules will contain bespoke strand-specific questions to reflect the themes covered by the bespoke RQs above.

For this strand, three separate interview schedules will be developed for three interviewee types (EEF programme and evaluation managers, programme developers, evaluators). All interview schedules, with the exception of those for use with the EEF team, will be shared with EEF and DfE for comment before they are used for data collection. We will share agenda/discussion themes with the EEF team in advance of interviews with them.

Interview schedules with a semi-structured framework of the themes and questions will be developed following the inception meeting with the EEF team. Three separate interview schedules will be developed, one for each interview type. The interview schedules for developers and evaluators will be shared and agreed with the EEF and DfE in advance.

5.4.3 Data collection

Interviews with EEF programme and evaluation managers

We will conduct one-hour paired interviews with the EEF programme and evaluation manager for each trial (four efficacy trials and two effectiveness trials) in April and early May (RQ2.1-5).

Interviews with programme developers

We will conduct interviews with the representatives of each programme teams, six group or individual interviews in total in May (RQ2.1-3).

Interviews with evaluators

We will conduct one-hour group interviews with representatives from the evaluation team for each trial, six interviews in total in May. We anticipate including the impact and IPE evaluation leads from the evaluation team in this interview (RQ2.1-4).

Focus group interview on recruitment with EEF staff including regional recruitment leads

We will conduct a semi-structured focus group with EEF staff supporting recruitment including regional recruitment leads in late June. The discussion will aim to explore the overall strategy for recruitment that has been part of the AF pipeline and the perceived success of this. We will also explore the role and approach they have taken in supporting projects in their efforts to recruit project participants, and their views of how this has worked. EEF will identify an appropriate group to take part. This will cover both the trials and scaling/scale up strands RQ2.2 ,4.

Review MI data

We will collate and review MI data and analyse monthly progress reports on recruitment focusing on targets and how participating schools compare with other schools within the regions. This data will provide insights into the extent to which of the strand objectives were achieved and capture challenges and barriers experienced throughout the project (RQ2.1-3). Progress reports will be used to gain qualitative insight into recruitment and implementation of the trials and identify common issues that have been identified by programme teams (RQ2.2,4).

RQ2.1 What was delivered in the strand during the timescale of the AF pipeline? Was this in line with expectations? Was the strand delivered with fidelity to its intentions?

RQ2.2 To what extent did this strand meet its key objectives?

RQ2.3 What were the mechanisms, facilitators, barriers, and contextual issues influencing the promise, feasibility, and scalability of this approach?

Bespoke to this strand:

RQ2.4 How effective was recruitment overall and across the six programmes? What aided and hindered recruitment of schools? How do participating schools compare with primary/secondary schools in the AF regions? What insights, if any, do developers have about motivations for school engagement?

Table 14: Data-collection methods mapped to RQs

	Interviews with EEF Programme & Evaluation Managers	Interviews with Evaluators	Interviews with Programme Developers	MI data review
RQ2.1	✓	✓	✓	
RQ2.2	✓	✓	✓	✓
RQ2.3	✓	✓	✓	
RQ2.4	✓	✓	✓	✓

5.5 Analysis and reporting

We will record all of our interviews (following participant agreement) and collate data thematically. We will initially use a deductive analysis approach by devising a broad set of codes aligned to the research aims and questions. However, we will be driven by the data, adding key sub-codes arising from the data which were not present in the initial framework, to ensure emergent findings are identified and to add detail to the analysis, including accounting for variation in performance across trials.

For those research questions addressed by multiple data sources, we will triangulate EEF, developer and evaluator interview data along with the MI and documentary analysis – as appropriate to the research questions. Further, we will compare and contrast views, experiences and practices for each programme to provide a comprehensive assessment of the implementation and early outcomes of the strand. We will produce a slide pack presenting the key findings of the evaluation.

5.6 Timeline

Table 15: evaluation timetable

Month	Activity	Who
February	Inception, ToC, begins, study plan design	SloE
March	Publication reviews; implementation; IPE & impact designs	SloE
	Preparation of recruitment documents	SloE
	Design research tools, plan interview schedules	SloE
	Interviewers briefed	SloE

Month	Activity	Who
April/May	EEF programme and evaluation managers interviews Programme developer and evaluator interviews	SloE
June/July/August	Interview analysis Recruitment & MI analyses	SloE
Sept	Finalise slide pack	SloE
Sept	Slide pack of presentation of findings	SloE

5.7 Personnel/team

Table 16: Evaluation team

Name	Role/responsibility in this strand	Organisation
Josephine Booth	Strand Lead ⁷	SloE
Eleanor Byrne	Researcher	SloE

Table 17: EEF team

Name	Role	EEF department
Iggy Rhodes	Head of Programmes	Programmes
Amy Clark	Senior Programme Manager	Programmes
Celeste Cheung	Senior Evaluation Manager	Evaluation

⁷ During the project inception period, Strand 2 was led by Sean Demack.

6. Scaling strand

6.1 Description of strand activities

The EEF has commissioned delivery of four programmes that have shown evidence of raising pupil attainment from previous EEF trials.

For three of these, EEF is providing support to scale the programme through funding and support in recruitment of approximately 100 schools to each. These are:

- 1st Class@Number
- Reciprocal Reading
- Switch-on Reading.

One programme is undertaken more fully at scale – ‘a scale up’ project: Embedding Formative Assessment (EFA).

The aims and objectives for the strand are:

- for some projects, there is the opportunity to refine their programme to enable further scale in the future e.g. through EEF effectiveness trials
- there is concurrent support for programmes to scale further whilst maintaining quality (capability building)
- there is increased awareness among delivery teams and the EEF about important implementation factors for scaling under different contexts.

The scaling stage is somewhat exploratory: to support the delivery at scale ahead of an effectiveness evaluation, enabling delivery teams to work through logistical issues, including those related to delivering in the hybrid format favoured by schools due to the pandemic. . All four scaling projects are supported by capability building with the aim of recruiting 190 primary (90 for 1stClass@Number; 100 for Reciprocal Reading) and 100 secondary (Switch-on Reading) schools and 150 secondary schools for the scale up project (EFA). Some formative evaluation of the capability building work is embedded to indicate the effectiveness of the capability building work, however, no programme impact will not be independently evaluated for any of the scaling programmes. For EFA, an implementation and process evaluation has been commissioned to understand the scaling mechanisms.

One scale-up project is included in this strand, which has sufficient promise at effectiveness stage to be rolled out to up to 150 new secondary schools this year.

6.2 The programmes and previous trial outcomes

Table 18: Overview of the three programmes

Project	Focus/features	Recruitment aims/date
1 st Class@Number	KS1 Maths, Y2 6-7, 2 years, targeted small group TA	scaling aiming for 90 primary schools (by April 2022)
Reciprocal Reading	KS2 Reading, Y4, Y5 & Y6 8-11, 1 year, teacher resources & PD, pupil training – targeted peer assisted	aiming for 100 primary schools (by July 2022)
Switch-on Reading	KS3 Maths, Y7 11-12, 10 week, targeted 1to1 TA	aiming for 100 secondary schools (by May 2022).
EFA	GCSE attainment 8, Y10-Y11 14-16, 2 years, whole school teacher PD	150 schools

1stClass@Number programme

A previous effectiveness trial found a positive impact for the 1st class number programme of +0.18 sds for a ‘Quantitative Reasoning’ primary outcome overall but a null effect for the FSM subsample. However, no impact was observed for the KS1 maths secondary outcome and, for the FSM subsample, a sizable (but not significant) negative impact of -0.25 sds was observed. This first effectiveness trial was not powered for the FSM subsample. Alongside the scaling work, this project is part of the capability building strand which will investigate possible pathways through which it could have a different impact on FSM-eligible pupils and how those effects could be mitigated. Following this analysis and other capability building work, the EEF will make a decision about a possible effectiveness re-trial.

Switch-on Reading

EEF has funded two evaluations of the switch-on reading programme; the first efficacy trial used a non-traditional gain-scores approach but found a relatively large effect size of +0.24 sds overall (+0.39 sds lower attainers; +0.36 FSM; +0.31 SEN). However, the second effectiveness trial found a null impact – but between these first two trials, the pupil cohort changed from KS3 (Y7-8, Efficacy Trial) to KS2 (Y3-Y4, Effectiveness trial). As with 1st class above, these mixed results are why the switch-on reading programme is included in the scaling group, with the aim of an effectiveness re-trial with a Y7 & Y8 KS3 pupil cohort (assuming scaling goes well).

Reciprocal Reading

More confidence was expressed that the third scaling project (targeted version of reciprocal reading) will progress to the effectiveness stage. In a previous efficacy trial the universal version of the programme showed no evidence of impact but the targeted version found a statistically significant impact for reading (+0.14 sds) and reading comprehension (+0.18 sds) overall. The targeted version of the programme also resulted in a significant positive impact for the FSM subsample for reading comprehension. Delivery during AF is of the targeted version of the programme. The grantee is receiving capability-building support to develop a train the trainer model, to prepare for effectiveness trial.

The greater consistency of findings from the efficacy trial (compared with the other two scaling projects) suggests more confidence that this programme will progress to effectiveness stage. Delivery during AF is of the targeted version of the programme. The grantee is receiving capability-building support to develop a train the trainer model, to prepare for effectiveness trial.

EFA

The single scale-up project (KS4 EFA) was selected using a balance of impact (as measured by effect size for attainment 8 outcome). The previous effectiveness trial found a modest positive impact (+0.10 sds overall; +0.07 sds for FSM). However, the cost of this programme is low and it was cost-effectiveness that was key for selecting this programme for scale up.

6.3 Evaluation objectives

Objective 1: Develop and further refine the ToC for this strand: set out the underlying assumptions, mechanisms and context underpinning the inputs, outputs and anticipated outcomes of the strand

Objective 2: Undertake light-touch evaluation through interview data collected from developers, evaluator (EFA programme only) and EEF to explore implementation, key enablers and barriers, and progress within the strand.

Objective 3: Explore progress towards target outputs/outcomes in relation to recruitment, undertake analysis of relevant internal MI data.

Objective 4: Undertake both inductive and deductive analysis (including triangulation and comparison across programmes) to evaluate the progress of the four programmes against strand objectives and in relation to research questions.

Objective 5: summarise formative findings for presentation.

As noted above, given the relatively short overlap between the strand evaluation time period and the trial time period, the focus of the strand evaluation is on recruitment and implementation.

6.4 Research questions (RQs)

Core to all strands:

RQ3.1 What was delivered in the strand during the timescale of the AF pipeline? Was this in line with expectations? Was the strand delivered with fidelity to its intentions?

RQ3.2 To what extent did this strand meet EEF's key objectives?

RQ3.3 What were the mechanisms, facilitators, barriers and contextual issues influencing the promise, feasibility and scalability of this approach?

Bespoke to this strand:

RQ3.4 How effective was recruitment overall and across the four scale-up projects? What were the barriers and facilitators for recruitment of schools to the scale up trials? What insights, if any, do developers have about motivations for school engagement?

RQ3.5 What processes support or are barriers to evaluation providing sufficient information to support the decision to recommend and/or support a scale up programme to be promoted in the future

RQ3.6 How could the EEF improve the process of programme selection and evaluation design for scale up projects to increase the likelihood of impact on attainment outcomes for disadvantaged children?

6.5 Research design

6.5.1 Summary

Given the timescale for the strand evaluation, an IPE is appropriate; with a focus on the effectiveness of support for trial design progress and school recruitment. This will be carried out with a mix of analysis of MI data and existing programme documentation, and primary data collection in the form of interviews and focus groups with EEF staff, programme teams and the evaluator for Embedding Formative Assessment.

Prior to commencement of data collection, the research team reviewed an initial strand ToC that had been prepared by the EEF strand team (this initial ToC can be found at Appendix A). The information gained from the meeting was used to refine the strand theory of change, develop RQs and to understand the type of MI data being collected, as well as designing the research instruments.

6.5.2 Research instruments

Following this inception meeting, the research team is now developing a series of semi-structured interview schedules, covering topics to align with the strand RQs. Each interview schedule will contain a small number of 'core interview questions'. These will align with the 'core RQs' outlined above and will be applied across all interview schedules across all five strands. Additionally, the schedules will contain bespoke strand-specific questions to reflect the themes covered by the bespoke RQs above.

For this strand, four separate interview schedules will be developed for the three interviewee types (EEF staff, programme developers, evaluators) and a focus group with, EEF recruitment staff, (as noted in the previous section this will cover recruitment issues for trials

and scaling/scale up strands). All interview schedules, with the exception of those for use with the EEF team, will be shared with EEF and DfE for comment before they are used for data collection. We will share agenda/discussion themes with the EEF team in advance of interviews with them.

6.5.3 Data collection

Review of programme and previous trial documentation

We will review programme and previous trial documentation including protocols and, where available, reports. The review of documentation will provide context to inform interviews and for analysis (RQ3.1-6).

Interview with EEF staff for each programme

We will conduct a semi-structured interview with EEF staff related to each of the four programmes, a total of four interviews in April and May. Because of potential cross-over areas of interest with the evaluation of the capability strand, the interviews will be transcribed for ease of data sharing across strands (RQ3.1-3, 5-6).

Interviews with programme developers

We will conduct a semi-structured interview with the representatives of the programme teams for each of the four projects in this strand, a total of four interviews in May. This will enable us to gather in-depth information around their project's aim and outcomes, previous trials and where they feel they need support. We will also ask questions around the nature of the support that EEF has offered, any perceived impacts of this, for example in terms of any changes the intervention planned, recruitment strategies, approaches to fidelity, confidence in taking their project forward etc. and why/how these have come about. We will also explore areas where they feel that their support needs have not been met, and what might be done to address this issue. Because of potential cross-over areas of interest with the evaluation of the capability strand, interviews will be transcribed for ease of data sharing across strands (RQ3.1-3, 56).

Interview with evaluator (EFA only)

We will conduct a semi-structured interview with representatives from the evaluation team for the Embedding Formative Assessment programme in May. This will gather their perceptions of the support offered by EEF for this project, how it may have helped the project move towards a situation where a scale up is more likely in terms of the project's design and implementation and the project team's understanding of what is needed. As a result of potential cross-over areas of interest with the evaluation of the capability strand, interviews will be transcribed for ease of data sharing across strands (RQ3.1-6).

Focus group interview on recruitment with EEF staff including regional recruitment leads

As per 5.4.3, we will conduct a semi-structured focus group with EEF staff supporting recruitment including regional recruitment leads in June. (RQ3.1-3,5,6).

Review of MI data

We will collate the MI data submitted by programmes during the period of support. This data will provide insights into the extent to which the strand objectives were achieved and capture

challenges and barriers experienced throughout the project (RQ2.1-3). For all four programmes we will review MI data on recruitment including EEF analysis of patterns in recruitment in relationship to school characteristics in the regions (RQ3.1-5).

RQ3.1 What was delivered in the strand during the timescale of the AF pipeline? Was this in line with expectations? Was the strand delivered with fidelity to its intentions?

RQ3.2 To what extent did this strand meet EEF’s key objectives?

RQ3.3 What were the mechanisms, facilitators, barriers and contextual issues influencing the promise, feasibility and scalability of this approach?

RQ3.4 How effective was recruitment overall and across the four scale-up projects? What were the barriers and facilitators for recruitment of schools to the scale up trials? What insights, if any, do developers have about motivations for school engagement?

RQ3.5 What processes support or are barriers to evaluation providing sufficient information to support the decision to recommend and/or support a scale up programme to be promoted in the future

RQ3.6 How could the EEF improve the process of programme selection and evaluation design for scale up projects to increase the likelihood of impact on attainment outcomes for disadvantaged children?

Table 19: Data-collection methods mapped to RQs

	Review of documents	Interviews with EEF staff	Interviews with developers	Interview with evaluator (EFA only)	Focus group with recruitment leads	Review of MI data
RQ3.1	✓	✓	✓	✓	✓	✓
RQ3.2	✓	✓	✓	✓	✓	✓
RQ3.3	✓	✓	✓	✓	✓	✓
RQ3.4		✓	✓	✓	✓	✓
RQ3.5	✓	✓		✓		
RQ3.6	✓	✓	✓	✓	✓	

6.6 Analysis and reporting

We will record all of our interviews (following participant agreement) and collate data thematically. We will initially use a deductive analysis approach by devising a broad set of codes aligned to the research aims and questions. However, we will be driven by the data,

adding key sub-codes arising from the data which were not present in the initial framework, to ensure emergent findings are identified and to add detail to the analysis. This analysis could include accounting for variation in recruitment across trials and identifying implementation enablers for scaling activities.

For those research questions addressed by multiple data sources, we will triangulate EEF, developer and evaluator interview data along with the MI and documentary analysis – as appropriate to the research questions. Further, we will compare and contrast views, experiences and practices for each programme in order to provide a comprehensive assessment of the implementation and early outcomes of the strand. We will produce a slide pack presenting the key findings of the evaluation.

6.7 Timeline

Table 20: evaluation timetable

Month	Activity	Who
Feb	Inception and TOC activity	SloE
March	Design research tools, plan interview schedule	SloE
March	Preparation of recruitment documents	SloE
April-July	Review of MI data	SloE
May	EEF staff interviews Review of programme documentation Programme developers and Evaluator interviews	SloE
June	EEF recruitment team focus group Analysis Analysis of MI and other existing programme data	SloE
July	Analysis	SloE
August	Analysis continues if needed	SloE
Sept	Slide pack of presentation of findings	SloE

6.8 Personnel/team

Table 21: Evaluation team

Name	Role/responsibility in this strand	Organisation
Josephine Booth	Strand Lead	SloE
Eleanor Byrne	Interviews & analysis	SloE

Table 22: EEF team

Name	Role	EEF department
Iggy Rhodes	Head of Programmes	Programmes
Amy Clark	Senior Programme Manager	Programmes
Celeste Cheung	Senior Evaluation Manager	Evaluation

7 Capability Building strand

7.1 Description of strand activities

This strand aims to help programmes develop the type of infrastructure capabilities needed for large-scale delivery of an intervention. Programmes that show signs of effectiveness at an early stage are provided tailored support considered likely to strengthen their capacity to deliver at a wider level. This capability building support aims to support both the supply and demand side of EEF's pipeline of quality programmes to schools: improving supply by ensuring programmes have the infrastructure to deliver effectively and at scale, and improving demand by making programmes more accessible to schools by improved marketing and communications.

Eight programmes that had all been the subject of previous EEF trials have been selected to receive capability-building support under AF. Through their knowledge of these programmes, including lessons from previous trial results, EEF identified a capability issue that could be effectively addressed with support and determined how this work could be best delivered. Depending on the type of work required, programmes were given direct funding and/or assigned a capability building partner to work with. Three partners were selected via an open tender to provide this capability partner support: the Behavioural Insights Team, the Centre for Evidence and Implementation, and Pedroza Communications.

Five of the eight programmes received a direct grant for work and similarly seven of the eight programmes received support from one or more of the partner organisations, which includes short-term evaluation by an external partner. A work plan was created for each programme with planned activities and objectives for the six month (February – July 2022) period of support.

The aim is that the objectives outlined in each programme's work plan are completed by August 2022. The completion of this work is intended to achieve the following outcomes:

- Improve the skills and understanding of factors relevant to delivering at scale for programme teams.
- Develop programme structure or delivery model to improve capacity to deliver at scale while maintaining programme effectiveness.
- Improve programme content to increase likelihood of programme having a positive impact on pupils.

7.2 Evaluation objectives

Objective 1: Develop and further refine the ToC for this strand.

Objective 2: Explore how the capability ‘deficit’ of a deliverer and/or programme and its potential solution was identified.

Objective 3: Examine whether support has addressed specified capability ‘deficit’ and made programme better equipped to delivery at scale and/or reach schools.

Objective 4: Examine extent to which programme teams have developed relevant learning and have embedded this into routine practices.

Objective 5: Explore how EEF provided support for capability building partners and programme teams.

Objective 6: Summarise formative findings for presentation.

7.3 Research questions (RQs)

Core to all strands

RQ4.1: What was delivered in the strand during the timescale of the AF pipeline? Was this in line with expectations? Was the strand delivered with fidelity to its intentions?

RQ4.2: To what extent did the strand meet its key objectives?

RQ4.3: What were the mechanisms, facilitators, barriers and contextual issues influencing the promise, feasibility and scalability of this approach?

Bespoke to this strand

RQ4.4: How effective were the diagnostic and problem-solving aspects of the capability support in terms of timeliness, quality, and level of support?

RQ4.5: To what extent have programme teams developed their learning, understanding, skills and organisational processes to deliver at scale because of capability-building support?

RQ4.6: To what extent have programme teams embedded this learning in their routine practices (e.g. within-organisation dissemination; cascaded learning)?

RQ4.7: To what extent are programmes better equipped to deliver at scale and/or reach schools as a result of capability-building support?

RQ4.8: Is this a feasible and sustainable model of support for the EEF to implement in the future?

RQ4.9: What was effective and what could be improved in the EEF team's support for capability building partners and programme teams?

7.4 Research design

7.4.1 Summary

The research design will primarily take the form of an IPE evaluation examining how the capability development was conducted and the extent to which the intended outcomes were achieved as a result of this work.

Prior to commencement of data collection, the research team reviewed an initial strand ToC that had been prepared by the EEF strand team (this initial ToC can be found at Appendix A). The information gained from the meeting are used to refine the strand theory of change and key RQs and to understand the type of MI data or reports being collected and the process for sharing the data, as well as designing the research instruments.

7.4.2 Research instruments

Following the inception meeting, the research team is now developing a series of semi-structured interview schedules, covering a number of topics to align with the strand RQs. Each interview schedule will contain a small number of 'core interview questions'. These will align with the 'core RQs' outlined above and will be applied across all interview schedules across all five strands. Additionally, the schedules will contain bespoke strand-specific questions to reflect the themes covered by the bespoke RQs above.

For this strand, five separate interview schedules will be developed for the three interviewee types (EEF team, Programme team, Partner organisations). All interview schedules, with the exception of those for use with the EEF team, will be shared with EEF and DfE for comment before they are used for data collection. We will share agenda/discussion themes with the EEF team in advance of interviews with them.

7.4.3 Data collection

Interviews with EEF team members

We will conduct a one-hour semi-structured interview with the EEF team members involved in the strand at the start of the evaluation in April (round 1) and also at the end of strand activity in July (round 2). These interviews will aim to gain insight on the design of the intervention, the strand's aims and intended outcomes and then reflections on outcomes and lessons learned.

The interviews will aim to:

- Understand whether the strand had met its objectives from the EEF's perspectives (RQ4.1-2) and what were the factors influencing its delivery (RQ4.3).
- Discuss the type of skills that programmes teams have developed as a result of this strand's work, and whether these are embedded within their practice (RQ4.5-6).
- Understand whether there are early indicators for programmes to be better equipped to deliver at scale (RQ4.7).
- Discuss the background to a focus on capability building for programme development at EEF; how did the idea arise, advice received in developing the idea, what capability support was provided prior to the AF project? (context only)
- Discuss the key lessons learned for the strand (RQ4.8-9).

- Early decision making for the AF project; how were programmes selected for support; how were issues/solutions to capability problems identified, selection and early work with partner organisations? (RQ4.4)
- The approach to EEF support and monitoring of ongoing capability building; how it works, systems etc.? (RQ4.9)
- Progress to date, including perceptions of what is working well or otherwise. (RQ4.2)
- Is this model feasible and sustainable for EEF to implement, and discuss potential improvements and plans for future. (RQ4.3, RQ4.8)

Following these interviews we will produce a revised version of the ToC, including a logic model depiction, for comment. The ToC will be revisited and refined as necessary in the summer towards the end of the evaluation.

Interviews with Programme teams

We will conduct a one-hour semi-structured interview with a senior staff member from each of the eight programmes at the start of their programme receiving support (round 1) and then again at the end of the period of support (round 2). Round 1 will take place in April when work plans have been completed and the programmes are recruiting. Round 2 will take place at the end of July when the programme delivery is completed.

The two interviews will aim to achieve the following:

- insight into the programme's stage of development and their understanding of the capability 'deficits' that could be addressed (round 1) (RQ4.4)
- insight into how capability issues to be addressed were identified and how solutions to these issues were decided (round 1) (RQ4.9)
- insight into how the project was set up and work plan created (round 1) (RQ4.9)
- explore work undertaken, relationship with partner organisation (if applicable) and EEF team (round 2) (RQ4.9) explore whether/how support has led to within-organisation learning, embedding of new practices etc. (round 2) (RQ4.5-6)
- explore if programme teams think they are better equipped to deliver their programme at scale. (RQ4.7)
- explore whether support has produced desired changes and/or unintended changes, and consider whether these changes are likely to be sustained beyond duration of the project (round 2). (RQ4.6)
- insight into how effectively the EEF team facilitated the capability building process (round 2). (RQ4.9)

Interviews with Partner organisations

We will conduct a one-hour semi-structured interview with a senior staff member from each of the three partner organisations at the end of the period of support. The interviews will aim to achieve the following:

- explore how partners worked with programme teams to deliver capability building support and the support given by the EEF to both programmes teams and capability partners (RQ4.9)
- insight into perceptions of the feasibility of the strand activities and approach.(RQ4.8)

- Perceptions of whether outcomes were achieved, learning and evidence of changes implemented within programme (RQ4.1-2)
- insight into whether changes achieved are likely to be sustainable for programmes (RQ4.8)
- insight into relationship with EEF and process of determining issues to address/solution to problems (RQ4.9)
- explore whether/how the overall capability building process with partners can be improved. (RQ4.9)

MI data/progress reports

We will collate the MI data recorded and analysed in monthly progress reports submitted by programmes during the period of support. This data will relate to quantitative outputs such as deliverables completed for each programme, attendance at workshops and whether scheduled work is on/off track. MI used will also include qualitative data such as reflections on activities conducted and challenges/lessons. This data will provide insights into the extent to which the strand objectives were achieved and capture challenges and barriers experienced throughout the project (RQ4.1-3). For example, progress reports provide detailed information on work completed against objectives and commentary on factors that enable/hinder progress. We expect to use MI data and information from the progress reports to inform the bespoke questions (RQ4.4-9).

RQ4.1 What was delivered in the strand during the timescale of the AF pipeline? Was this in line with expectations? Was the strand delivered with fidelity to its intentions?

RQ4.2 To what extent did the strand meet its key objectives?

RQ4.3 What were the mechanisms, facilitators, barriers and contextual issues influencing the promise, feasibility and scalability of this approach?

RQ4.4 How effective were the diagnostic and problem-solving aspects of the capability support in terms of timeliness, quality, and level of support?

RQ4.5 To what extent have programme teams developed their learning, understanding, skills and organisational processes to deliver at scale because of capability-building support?

RQ4.6 To what extent have programme teams embedded this learning in their routine practices (e.g. within-organisation dissemination; cascaded learning)?

RQ4.7 To what extent are programmes better equipped to deliver at scale and/or reach schools as a result of capability-building support?

RQ4.8 Is this a feasible and sustainable model of support for the EEF to implement in the future?

RQ4.9 What was effective and what could be improved in the EEF team's support for capability building partners and programme teams?

Table 23: Data collection methods mapped to RQs

	Interviews with EEF team	Interviews with Programme teams	Interviews with Partner organisations	Collation of MI data/progress reports
RQ4.1	✓		✓	✓
RQ4.2	✓		✓	✓
RQ4.3	✓			✓
RQ4.4		✓		✓
RQ4.5	✓	✓		✓
RQ4.6	✓	✓		✓
RQ4.7	✓	✓		✓
RQ4.8	✓		✓	✓
RQ4.9	✓	✓	✓	✓

7.5 Analysis and reporting

We will record all of our interviews (following participant agreement) and collate data thematically. We will initially use a deductive analysis approach by devising a broad set of codes aligned to the research aims and questions. However, we will be driven by the data, adding key sub-codes arising from the data which were not present in the initial framework. This analysis could include identifying common enablers/barriers to implementation and factors correlated to practice change within delivery organisations.

For those research questions addressed by multiple data sources, we will triangulate EEF, programme team and support partner interview data along with the MI analysis – as appropriate to the research questions. Further, we will compare and contrast views, experiences and practices for each programme in order to provide a comprehensive assessment of the implementation and early outcomes of the strand. We will produce a slide pack presenting the key findings of the evaluation.

7.6 Timeline

Table 24: evaluation timetable

Month	Activity	Who
March	Instrument design, finalisation of logic model/ToC	NFER
March	Preparation of recruitment documents	NFER
April	Interview with EEF team (round 1)	NFER/EEF team
April	Programme team interviews (round 1)	NFER/Programme teams
April-July	Collation of MI data/progress reports	NFER
June	Programme team interviews (round 2)	NFER/Programme teams
July	Partner organisation interviews	NFER/Partner organisations
July	Interview with EEF team (round 2)	NFER/EEF team
August	Data analysis	NFER
September	Slide pack of presentation of findings	NFER/EEF/DfE

7.7 Personnel/team

Table 25: Evaluation team

Name	Role/responsibility in this strand	Organisation
Roland Marden	Strand Lead	NFER
Caroline Stanford (Research Associate)	Data collection and analysis	NFER

Table 26: EEF team

Name	Role	EEF department
Ben Simuyandi	Programme Manager	Programmes
Amy Clark	Senior Programme Manager	Programmes
Rachael Morris	Evaluation Manager	Evaluation

8 School Support strand

8.1 Description of strand activities

This strand aims to support leaders in these schools to make more use of evidence to support decision making and school improvement planning, targeting schools in the three AF regions which have not previously engaged with the EEF. The strand activities include two twilight sessions for senior school leaders providing an introduction on how to critically assess issues and identify solutions. Activities also include a six module training programme which focus on evidence-based decision making within the context of Covid-19 educational recovery. As part of the sessions, they are introduced to EEF's resources and training, with the aim of them engaging further with these in the future.

There are 17 Research Schools and Associate Research Schools in the three target regions, and all are involved in delivering this strand. This includes three of the 17 Research Schools acting as recruitment hubs, one in each of the three regions. Ten of the research schools are delivering twilights (four delivering as one consortium, two delivering as one consortium and four delivering individually)) and all 17 of the Research Schools are delivering the training programme. Research Schools attended 'train the trainer' sessions before delivery began.

The aim is to recruit schools that have not previously engaged with EEF with a target to reach 450-650 schools in total, with 450-650 engaging with the twilights and around 400 engaging with the training programme (with some overlap). As a minimum, the strand aims to engage 400 schools in at least one twilight session or one module of the training programme.

Recruitment began at the end of November 2021 with twilight delivery starting in January 2022 and the training programme from March 2022. The strand will be completed by the end of the 2021/22 academic year.

8.2 Evaluation objectives

Objective 1: Develop and refine strand ToC

Objective 2: Explore learning and early outcomes from the delivery of this strand, including what has worked well and less well in terms of the recruitment hubs and the delivery of the twilights and training programme

Objective 3: Undertake analysis of MI data to explore progress to target outputs

Objective 4: Undertake interviews with strand participants to understand likely outcomes e.g. likeliness to engage with EEF resources, training or programmes or to use evidence in school decision making.

Objective 5: summarise formative findings for presentation.

The evaluation will place more emphasis on processes than outcomes, with the aim of uncovering any potential learning for EEF in relation to its wider regional school support offer. However, we will explore early indications of participant outcomes to augment findings relating to effectiveness of implementation.

8.3 Research questions (RQs)

Core to all strands

RQ5.1: What was delivered in the strand during the timescale of the AF pipeline? Was this in line with expectations? Was the strand delivered with fidelity to its intentions?

RQ5.2: To what extent did the strand meet its key objectives?

RQ5.3: What were the mechanisms, facilitators, barriers and contextual issues influencing the promise and feasibility⁸ of this approach to engaging schools which do not usually engage with the EEF?

Bespoke to this strand

RQ5.4: Did the strand meet its recruitment and attendance targets? On average, how many twilight and/or training sessions did schools engage with? Were there any common characteristics of the less and more engaged schools?

RQ5.5: What worked well and less well in recruitment and training delivery? Is there any learning from the model, or elements of the model, for EEF's future work with non-engaged schools?

RQ5.6: How does this model compare to EEF's usual regional support model? What are the benefits and disadvantages of this support model compared to EEF's usual regional school support model?

RQ5.7: How useful did schools find the signposting to EEF programmes or resources provided through the twilight sessions and training programme?

RQ5.8: Are there any indications of early participant outcomes (e.g. awareness, knowledge, confidence or motivations to engage with EEF's resources, training or programmes or to use evidence in decision making)?

RQ5.9: What, if anything, will schools do differently as a result of attending the twilights/training?

8.4 Research design

8.4.1 Summary

The research design will primarily take the form of an IPE evaluation with a light-touch exploration of early outcomes.

Prior to commencement of data collection, the research team reviewed an initial strand ToC that had been prepared by the EEF strand team (this initial ToC can be found at Appendix A). The information gained from the meeting was used to refine the strand theory of change,

⁸ We have not included scalability, as we understand that EEF do not wish to scale this approach beyond the immediate AF regions.

develop RQs and to understand the MI data being collected, as well as designing the research instruments.

8.4.2 Research instruments

Following this inception meeting, the research team is now developing a series of semi-structured interview schedules, covering a number of topics to align with the strand RQs. Each interview schedule will contain a small number of 'core interview questions'. These will align with the 'core RQs' outlined above and will be applied across all interview schedules across all five strands. Additionally, the schedules will contain bespoke strand-specific questions to reflect the themes covered by the bespoke RQs above.

For this strand, three separate interview schedules will be developed for the three interviewee types (recruitment hubs, Research Schools and school leader Focus groups). All interview schedules, with the exception of those for use with the EEF team, will be shared with EEF and DfE for comment before they are used for data collection. We will share agenda/discussion themes with the EEF team in advance of interviews with them.

8.4.3 Data collection

Interviews with EEF team members

We will conduct a one-hour semi-structured interview with the EEF team members involved in the strand at the start of the evaluation in April (round 1) and also at the end of strand activity in July (round 2). These interviews will aim to gain insight on the design of the intervention, the strand's aims and intended outcomes and then reflections on outcomes and lessons learned.

Interviews with recruitment hubs

In May, we will conduct interviews of around 45 minutes with the project managers within the recruitment hubs to explore their perceptions of how their role in the strand has been working, including what has worked well or otherwise, any challenges faced and if/how they have been overcome. This will include exploring their role in recruiting schools to the strand and the different training sessions, and in liaising with the RSs (RQ5.1-6).

Interviews with RSs

In June, we will interview RSs to explore how their delivery of the twilight sessions and training programme has gone and their views on the role of the recruitment hubs in engaging schools. We propose to conduct interviews of around 45 minutes with a key staff member within six of the ten RSs involved in the twilights.

We will interview RSs delivering training using purposive sampling RSs which have had different experiences in terms of the amount of engagement from schools they have experienced: three schools which have had high levels of engagement and three schools with lower levels of engagement. (RQ5.1-6).

Focus groups with school leaders

Between June and July, we will undertake three short (around 30 minutes) regional focus groups, with around four participants in each, at the end of one of the training programme modules. This will lead to feedback from around 12 participants. Consulting with schools

whilst they are already undertaking programme activity will be a less burdensome and a more time- and cost-effective way of bringing together focus groups and engaging senior leaders in the evaluation than contacting them after programme completion.

The aim of this task will be to gain an external, independent perspective on the benefits of this type of training and the early outcomes that it can lead to. We will explore changes in participants' awareness, knowledge, confidence and motivations in using evidence in decision-making. Also, we will discuss actions they may have taken or plan to take next academic year as result of attending the training programme. And if they report no action or plans, we will investigate the barriers for changing practice. (RQ5.1-3, 5, 7-9)

Analysis of MI data

We understand that a range of internal MI data on recruitment and attendance is being collected on the train-the-trainer, twilight and training programmes by the EEF team. We plan to analyse this data as part of the evaluation. We will also want to draw on participant feedback data that is collected at the end of the twilight sessions and is planned to be collected during the training programme. This data will include participant satisfaction and motivations for signing up for the session/training programme. We will use this data will to investigate how participants engaged with the offer and demonstrate whether intentions of participants were aligned with the purpose of the sessions/training programme.

We understand that there is a possibility that a follow-up email will be sent to participants asking about their plans following the training (in using evidence in decision making and drawing on EEF resources and training) and we would also be keen to draw on this data if it is collected. Analysis of this data will help us evaluate the medium term outcomes of the strand, in particular whether participants are likely to translate learning into improved practice in schools. (RQ5.1-5, 8-9).

MI data/progress reports

We will collate the MI data recorded and analysed in monthly progress reports submitted by programmes during the period of support. This data will relate to quantitative outputs such as deliverables completed for each programme, attendance at workshops and whether scheduled work is on/off track. MI used will also include qualitative data such as reflections on activities conducted and challenges/lessons. This data will provide insights into the extent to which the strand objectives were achieved and capture challenges and barriers experienced throughout the project (RQ5.1-3). For example, progress reports provide detailed information on work completed against objectives and commentary on factors that enable/hinder progress. We expect to use MI data and information from the progress reports to inform the bespoke questions (RQ5.4-9)

RQ5.1: What was delivered in the strand during the timescale of the AF pipeline? Was this in line with expectations? Was the strand delivered with fidelity to its intentions?

RQ5.2: To what extent did the strand meet its key objectives?

RQ5.3: What were the mechanisms, facilitators, barriers and contextual issues influencing

the promise and feasibility⁹ of this approach to engaging schools which do not usually engage with the EEF?

RQ5.4: Did the strand meet its recruitment and attendance targets? On average, how many twilight and/or training sessions did schools engage with? Were there any common characteristics of the less and more engaged schools?

RQ5.5: What worked well and less well in recruitment and training delivery? Is there any learning from the model, or elements of the model, for EEF’s future work with non-engaged schools?

RQ5.6: How does this model compare to EEF’s usual regional support model? What are the benefits and disadvantages of this support model compared to EEF’s usual regional school support model?

RQ5.7: How useful did schools find the signposting to EEF programmes or resources provided through the twilight sessions and training programme?

RQ5.8: Are there any indications of early participant outcomes (e.g. awareness, knowledge, confidence or motivations to engage with EEF’s resources, training or programmes or to use evidence in decision making)?

RQ5.9: What, if anything, will schools do differently as a result of attending the twilights/training?

Table 27: Data-collection methods mapped to RQs

	Interviews with recruitment hubs	Interviews with Research Schools	Focus groups with school leaders	Analysis of MI data
RQ5.1	✓	✓	✓	✓
RQ5.2	✓	✓	✓	✓
RQ5.3	✓	✓	✓	✓
RQ5.4	✓	✓		✓
RQ5.5	✓	✓	✓	✓
RQ5.6	✓	✓		
RQ5.7			✓	✓
RQ5.8			✓	✓
RQ5.9			✓	✓

⁹ We have not included scalability as we understand that EEF do not wish to scale this approach beyond the immediate AF regions.

8.5 Analysis and reporting

We will record all of our interviews and focus groups (following participant agreement) and collate data thematically. We will initially use a deductive analysis approach by devising a broad set of codes aligned to the research aims and questions. However, we will be driven by the data, adding key sub-codes arising from the data which were not present in the initial framework to ensure emergent findings are identified and to add detail to the analysis. This analysis could include accounting for variation in engagement across regions and identifying factors associated with effective learning in sessions.

For those research questions addressed by multiple data sources, we will triangulate EEF, recruitment hub and RS interview data along with the MI analysis – as appropriate to the research questions to provide a comprehensive assessment of the implementation and early outcomes of the strand. We will produce a slide pack presenting the key findings of the evaluation.

8.6 Timeline

Table 28: Evaluation timetable

Month	Activity	Who
March	Inception meeting and interview with EEF programme managers	NFER/EEF programme managers
April	Finalisation of ToC Design of research instruments	NFER/EEF programme managers
April-July	Review of MI data	NFER
April	Preparation of recruitment documents	NFER
April	EEF team interview	NFER
June	Interviews with recruitment hubs	NFER/recruitment hubs
June	Interviews with Research Schools Focus groups with school leaders	NFER/recruitment hubs/schools
July	Focus groups with school leaders	NFER/schools
August	Data analysis	NFER
September	Slide pack of presentation of findings	NFER/EEF/DfE

8.7 Personnel/team

Table 29: Evaluation team

Name	Role/responsibility in this strand	Organisation
Roland Marden	Strand Lead	NFER
Kelly Kettlewell (Research Associate)	Data collection and analysis	NFER

Table 30: EEF team

Name	Role	EEF department
Emma Price	Programme Manager	Dissemination and Impact
Stu Mathers	Head of Department	Dissemination and Impact

9 Risks

Below is a table outlining the risks associated with this study, our assessment of their likelihood of occurring and the countermeasures in place to mitigate them.

Table 31: Risk assessment and contingencies

Risk	Strands/OE/All	Assessment	Countermeasures
Low participant response rates (interviews)	All	Likelihood: low Impact: high	Direct, tailored participant contact through recruitment plan; raise awareness of expectations and value of contribution; targeted/short research instruments to minimise burden
Timescale slippage including delays in delivery, or production of MI	All	Likelihood: Moderate Impact: high	EEF keep NFER/SIoE informed of changes in delivery expectations; and ensure there is adequate planning on data sharing process and communicate clearly and early with EEF team about what data is available or should be shared to ensure quality MI data is available when needed; NFER/SIoE apply rigorous project planning processes and have open/collaborative discussions with EEF to resolve issues
Conflicts in NFER/ SloE collaboration	All	Likelihood: low Impact: moderate	Build on/apply effective processes developed through successful current/past collaborations; assign clear roles and responsibilities; organisations' quality advisors will support early resolution of issues arising
Data security compromised	All	Likelihood: low Impact: moderate	Steps taken throughout to ensure personal data is protected, in accordance with NFER/SIoE data security policies
Changes to the evaluation team due to sickness, absence or staff turnover	All	Likelihood: moderate Impact: moderate	Both organisations have large staff pools to replace team members. Clear and accurate project documentation will support continuity

11 References

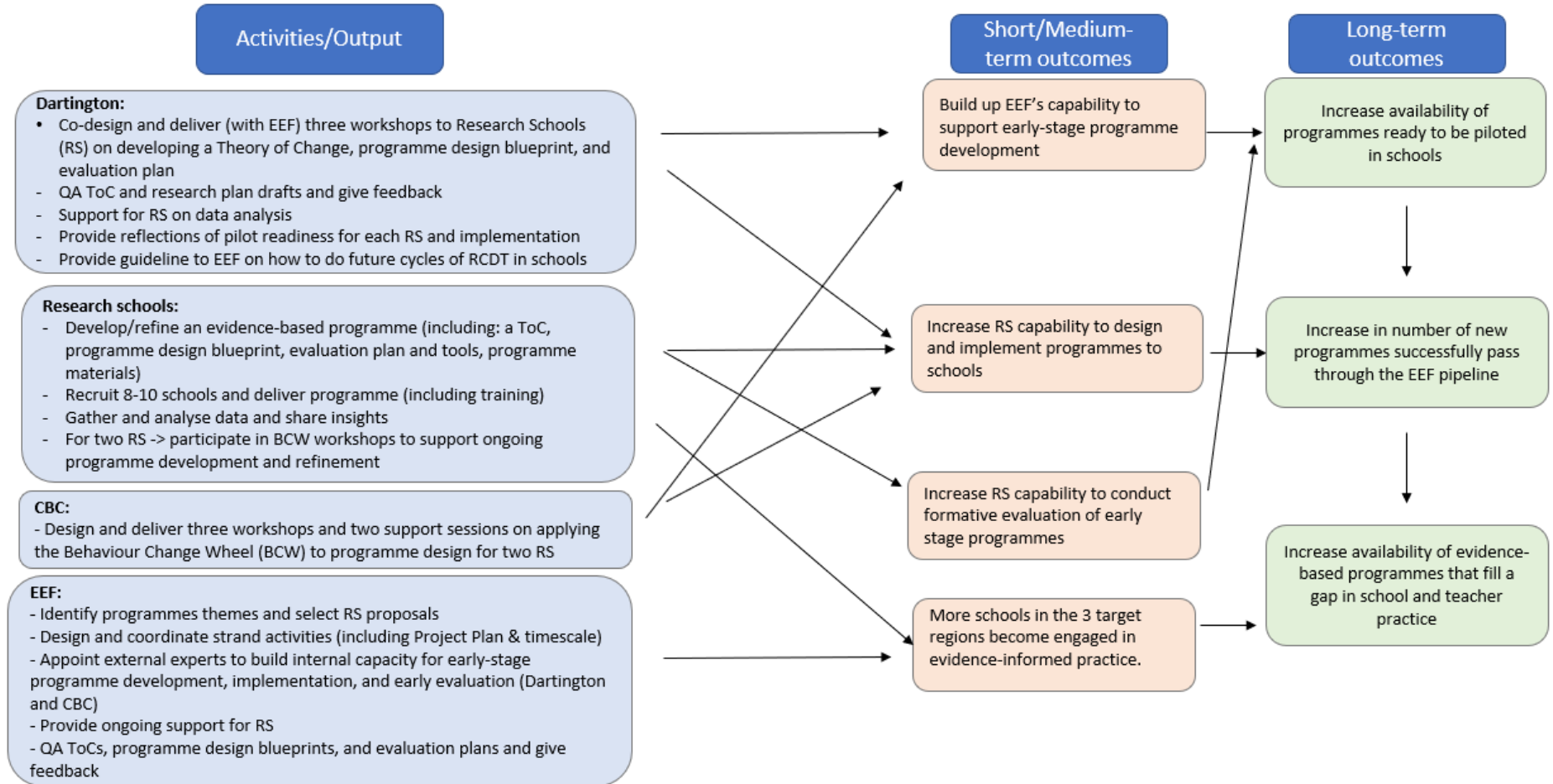
Asmussen, K., Brims, L. and McBride, T. (2019). *Ten Steps for Evaluation Success* [online]. Available: <https://www.eif.org.uk/files/pdf/10-steps-for-evaluation-success.pdf> [16 December, 2021].

Coldwell, M., Maxwell, B, and Sharples, J. (2021). Developing a Systems-Based Approach to Research-Use in Education: the Case of the EEF and the Deployment of Teaching Assistants. Review of Education. Unpublished report.

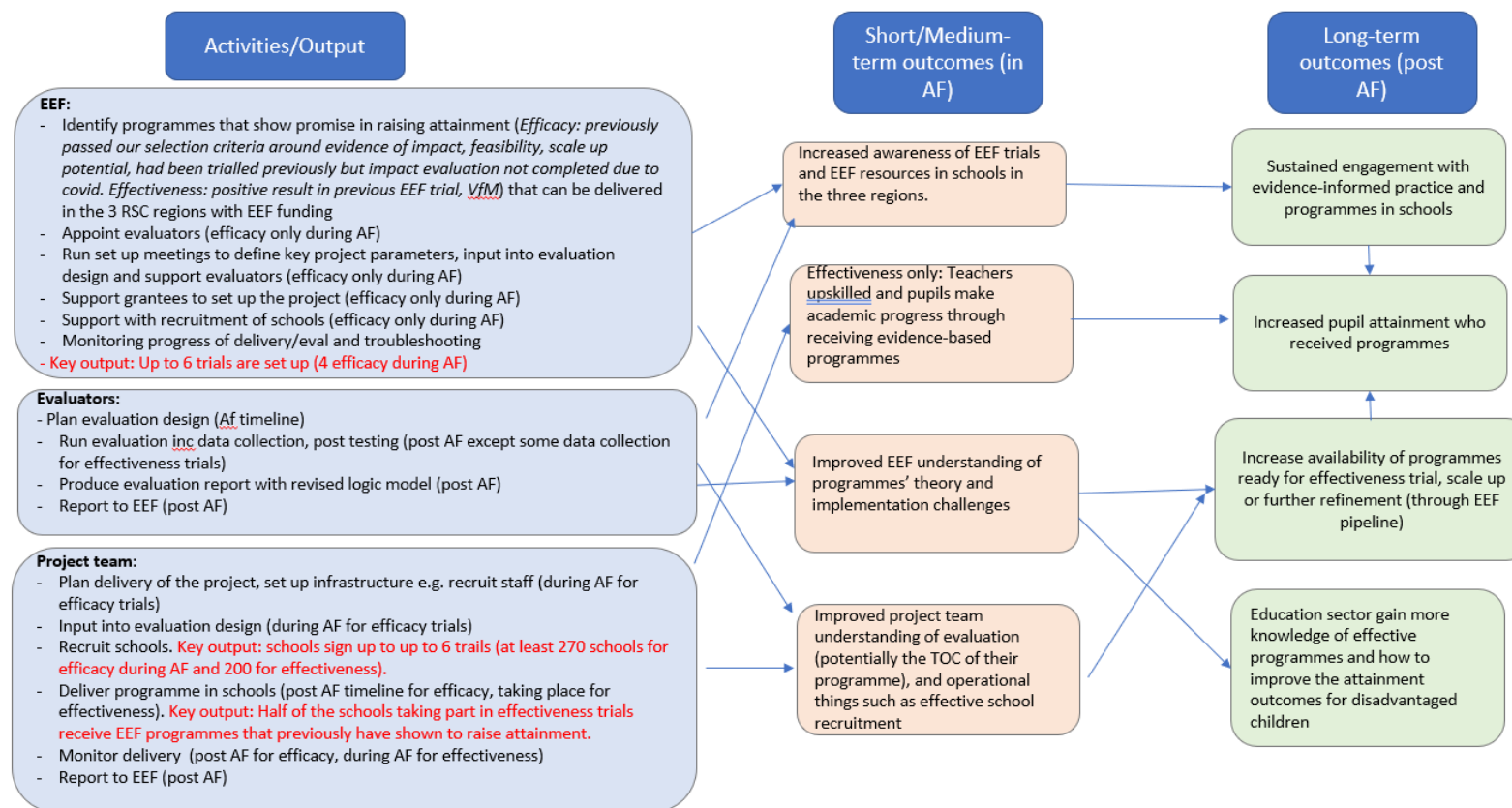
Gough, D., Maidment C. and Sharples, J. (2018). *UK What Works Centres: Aims, Methods and Contexts* [online]. Available: <https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3731> [16 December, 2021].

Martin, G. and William, O. (2019). 'Evidence and service delivery.' In: A. Boaz, H, Davies, A. Fraser and S. Nutley (Eds) *What Works Now? Evidence-informed Policy and Practice*. London: Palgrave.

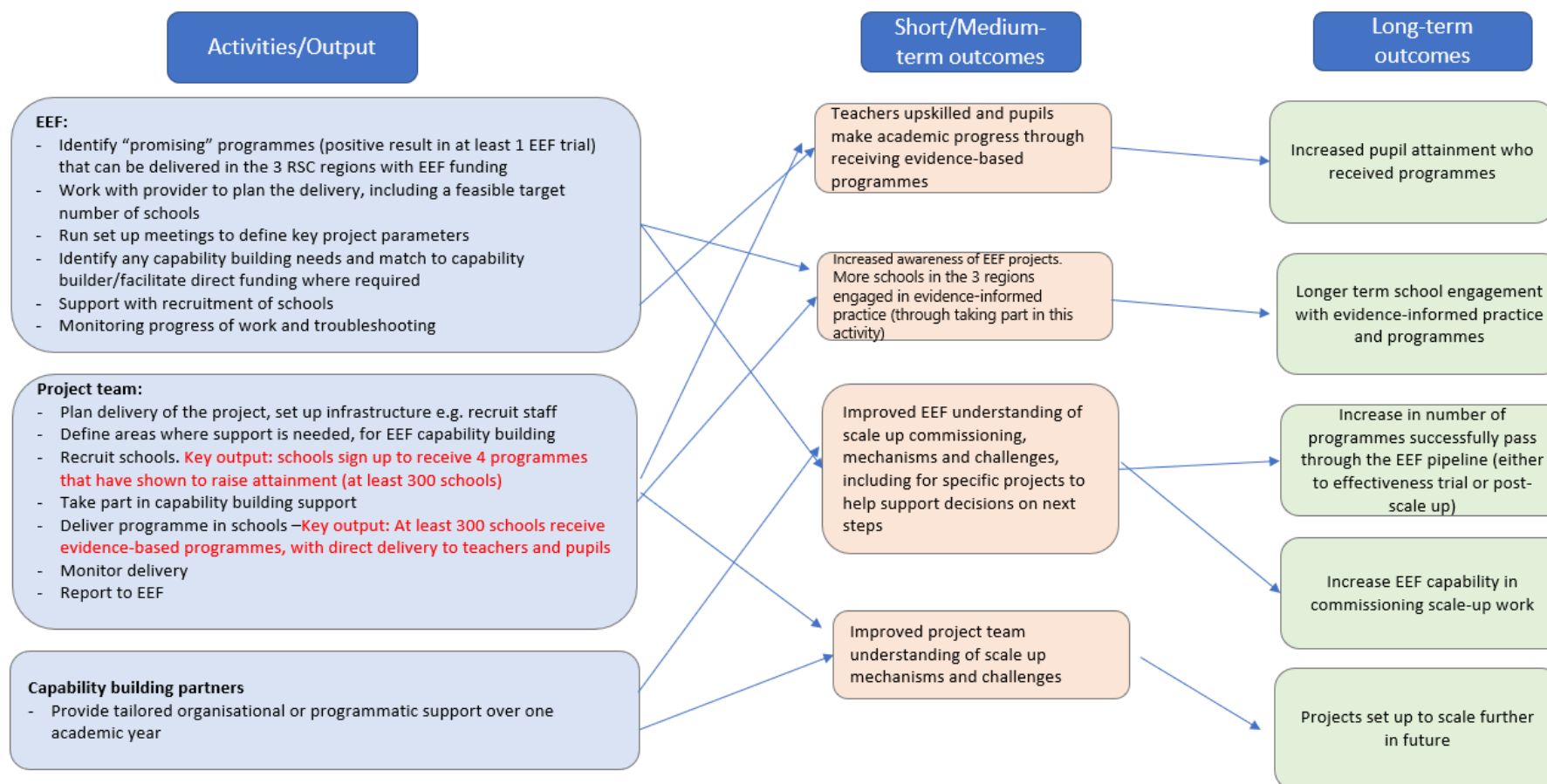
Appendix A – First ToC draft for Early-stage programme development strand



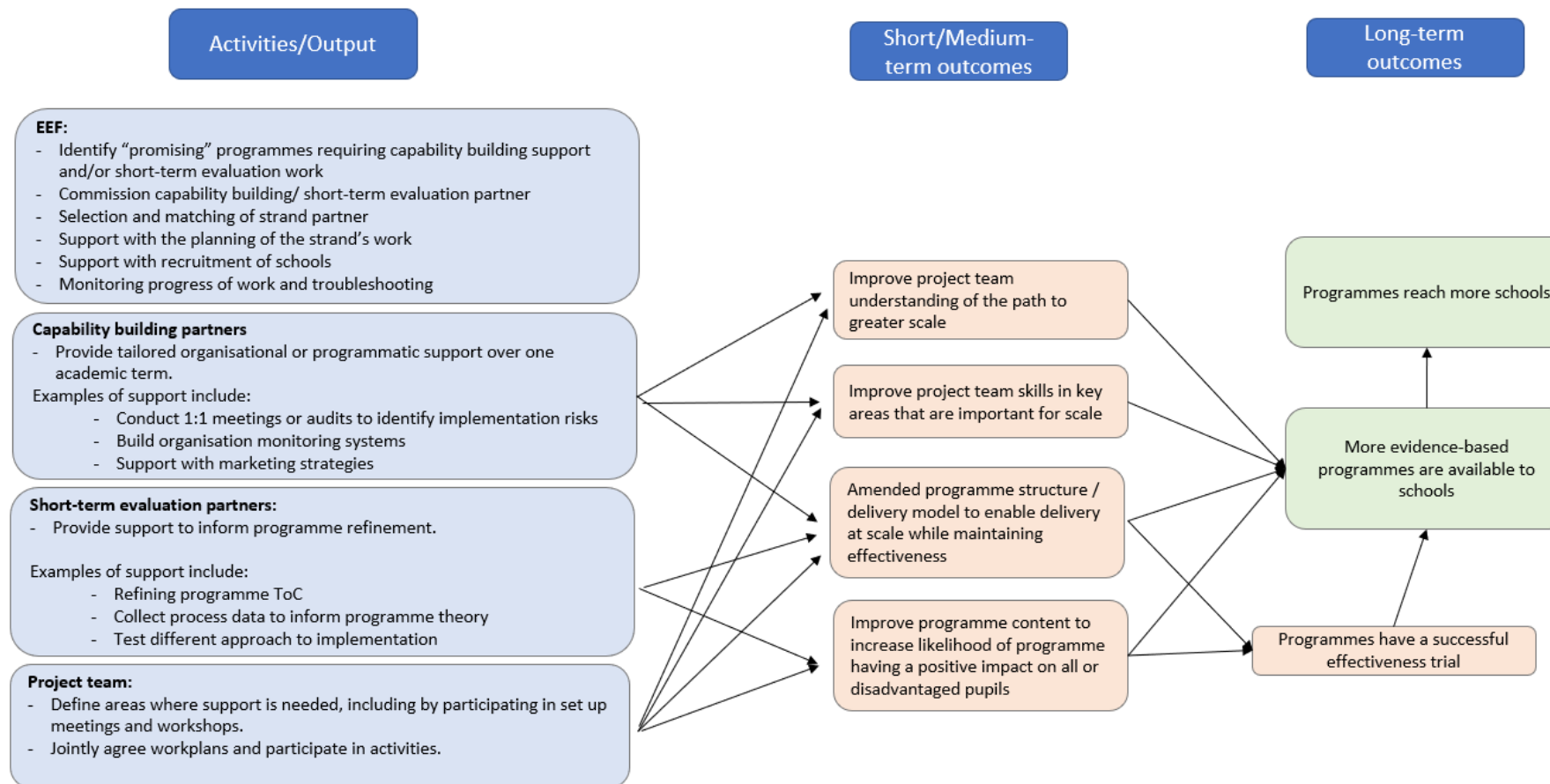
Appendix B – First ToC draft for Trials strand



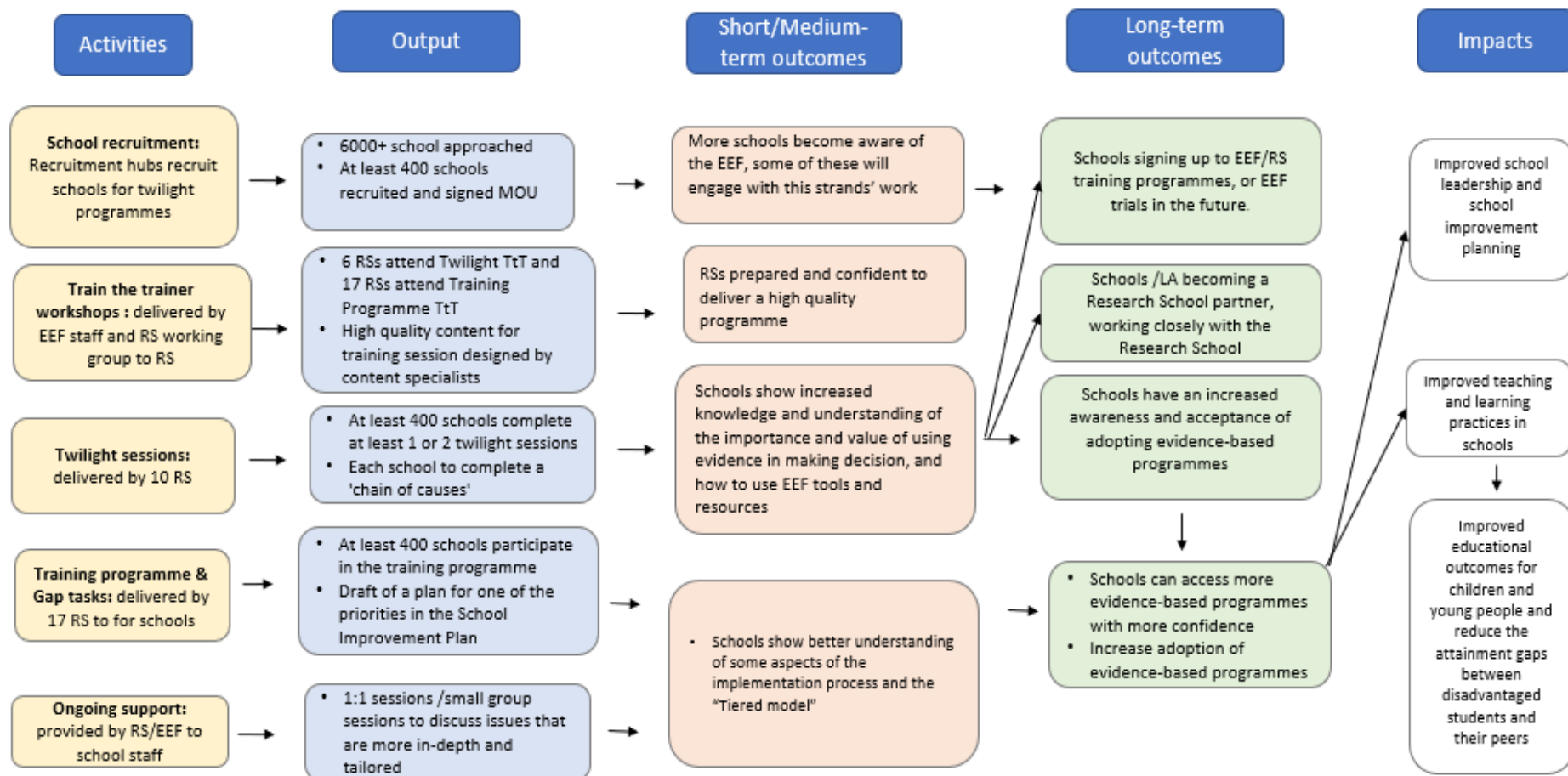
Appendix C – First ToC draft for Scaling strand



Appendix D – First ToC draft for Capability Building strand



Appendix E – First ToC draft for School Support strand



Evidence for excellence in education

Confidential

© National Foundation for Educational Research 2022

All rights reserved. No part of this document may be reproduced or transmitted in any form or by any means, electronic, mechanical, photocopying, or otherwise, without prior written permission of NFER.

The Mere, Upton Park, Slough, Berks SL1 2DQ
T: +44 (0)1753 574123 • F: +44 (0)1753 691632 • enquiries@nfer.ac.uk

www.nfer.ac.uk

NFER ref. EEPE

