Pilot Study Plan
NIESR and IES
Anneka Dawson and Lucy Stokes

Education Endowment Foundation

PROJECT TITLE	Early Years Toolbox
DEVELOPER (INSTITUTION)	Action for Children and Professor Ted Melhuish
EVALUATOR (INSTITUTION)	National Institute of Economic and Social Research (NIESR) and Institute for Employment Studies (IES)
PRINCIPAL INVESTIGATOR(S)	Anneka Dawson (IES) and Lucy Stokes (NIESR)
STUDY PLAN AUTHOR(S)	Anneka Dawson, Clare Huxley, Nathan Hudson-Sharp, Chiara Manzoni, Johnny Runge, Lucy Stokes, Helena Takala, Ceri Williams
PUPIL AGE RANGE AND KEY STAGE	Nursery (3 and 4 year olds)
NUMBER OF SCHOOLS	27 settings
NUMBER OF PUPILS	737

Study plan version history

VERSION	DATE	REASON FOR REVISION
1.2 [*latest*]		
1.1		
1.0 [*original*]		[*leave blank for the original version*]

Intervention

The Early Years Toolbox (EYT) is a series of apps designed to measure young children's emerging cognitive, self-regulatory, language and social development. Each measure is a brief, game-like assessment that has been developed for the iPad. Developed by leading early years researchers, Professor Edward Melhuish (University of Oxford) and Dr Steven Howard (University of Wollongong, Australia), the EYT is designed to capture abilities that research has shown to be most predictive of later academic, social, emotional, cognitive and life outcomes. The EYT is currently used across five continents (Howard, Melhuish & Chadwick, 2018) in a range of settings for both education and research. The EYT is a simple to use, quick set of measures and has shown promising results in the Australian context, (Howard & Melhuish, 2015) with good internal consistency and convergent validity with other early years' measures.

Early years practitioners will be trained in the use of the EYT, by Professor Ted Melhuish and Action for Children (AfC). By providing staff with information about children's emerging abilities, the idea is that practitioners will be better placed to ensure children's experiences in the settings are tailored to their individual development needs.

Twenty-seven early years settings are participating in the study. Key staff from the participating settings attend an initial regional training day in the first half term of the school year, with staff then expected to cascade the training to other staff in their setting working with three year olds. The training covers how to use the three apps which are the focus of this project: one that practitioners complete with children which examines numeracy skills ('EYT Early Numeracy'), one that practitioners complete with children which examines vocabulary('Expressive Vocabulary') and one that practitioners use themselves to assess children's self- regulation and social development ('Child Self-Regulation and Behaviour Questionnaire'). The EYT also includes several other apps but these are not being evaluated in this project. The training gives a background to the project and development of the apps and then runs through how to use each of the apps. The practitioners then get a chance to practice using the apps with some children from the setting they are being trained in and ask any questions.

Each setting then receives a follow-up training visit from the Action for Children team of approximately two hours before Christmas. The follow- up session will give practitioners the chance to discuss how cascading has gone, for the trainer to observe practitioners carrying out the apps with children and then for practitioners to ask any questions they have. After this, settings will be visited approximately every three months (or term) by a member of the Action for Children team, with further support provided through regular phone calls if they are required, plus access to a helpline via email and telephone. A forum has also been established in order to provide further support and to give the opportunity for participants to share learning and progress. AfC will also provide practitioners with a digital pack of additional evidence based materials on numeracy, language development, self-regulation and social development taken from the research literature and a short summary document with suggestions for the different areas and further reading. Nurseries will be expected to use the app with all three year old children in the setting, at least once per term. A consolidation event will be held at the end of the pilot.

Research questions

The aim of the evaluation is to examine whether the pilot programme shows evidence of promise, is feasible for providers to deliver, and can be replicated, scaled and tested as part of a larger efficacy trial. This pilot evaluation will seek to answer the following, more detailed research questions in each of these areas in order to direct the collection of evidence and show whether these conditions have been met:

- **Whether there is evidence to support the theory of change**
  - *Is there evidence of an improvement in practitioners' practice and behaviour to better support children's early development?*

- *Is there evidence of an improvement in practitioners' understanding of children's competencies and skills?*
- *Do participants perceive a change in work practices from engaging in the intervention?*
- *Do participants believe that the intervention has had (or will have) benefits to children compared to previous practices?*
- *Are there any unintended or negative effects of the intervention?*
- *Are practitioners more confident that they are correctly targeting children in need of support?*

- **Whether the approach is feasible to deliver**
  - *Is the intervention feasible for practitioners to engage in and implement?*
  - *Can the intervention be delivered within routine practice?*
  - *What are the facilitators/barriers to the programme?*
  - *What level of support is needed to enable practitioners to effectively interpret assessment data?*
  - *What level of support is needed to enable practitioners to effectively change their practice based on this data?*
  - *What is the optimal / most feasible frequency and timing of the assessments within the school term?*

- **Whether the programme is ready to be evaluated in a trial**
  - *Does the training and resources constitute a standard intervention that can be replicated and scaled up?*
  - *What, if anything, should be updated or changed for a trial evaluation (e.g. theory, materials, procedures)?*
  - *What will the cost of the intervention be to settings (both direct costs of training and indirect costs e.g. of interventions bought as a result of the assessments)?*

Methods

Recruitment

Settings were recruited to the pilot through Action for Children. The participating settings include some Action for Children nurseries but also other settings in three regions- the South (covering Barnett, Hammersmith, Fulham, Essex, Newham, Bexley), North- East (covering North Hull, Gateshead, Northumberland, Stockton) and the midlands (Birmingham, Derby, Telford and Wrekin).

All 3 year old children at the settings are eligible to participate in the study.[1]

Settings were asked to sign a Memorandum of Understanding, setting out the requirements of the study and what it entails. Settings were also provided with an information letter for sending to parents/guardians, explaining what the study involves and how their child's data will be used. This gives parents clear instructions as to how their child's data can be withdrawn from the evaluation. If a withdrawal request is received, practitioners will not use the app with the child. A privacy notice has also been made available at: https://www.niesr.ac.uk/projects/pilot-evaluation-early-years-toolbox

Data collection

Drawing on the EEF Implementation and process evaluation guidance, (Humphrey at et al., 2016) the study uses a multiphase design, based around a triangulation of mixed methods, to help fully develop and explore a theory of change model and then test the three main EEF pilot research questions as outlined above.

To do this we will use the following data collection methods:

---

[1] This will also include children who were 3 at the start of the school year but who turn 4 during the life of the project.

- A series of 3 IDEA workshops in a development package to help construct and test a theory of change model for the intervention
- Case study visits to 12 nurseries to interview practitioners and setting managers, 6 visits will take place midway through the programme and the remaining 6 towards the end of the programme
- Telephone interviews with key staff at 12 settings (again, 6 of these to take place midway through the programme and the remaining 6 towards the end)[2]
- Observations of the group training sessions
- Observations of some of the individual follow-up training (3 observations)
- Observations of some of the quarterly meetings, to take place in the final term (3 observations)
- Review of programme materials
- Sharing and analysis of support provided by AfC via the helpline and emails (recorded by AfC via a log)
- Sharing and analysis of support provided by AfC via the quarterly monitoring meetings/ regular phone calls (recorded by AfC via a log)
- Sharing and analysis of data collected by the delivery team of feedback surveys from the training
- Online survey (baseline and post-treatment) of setting managers and staff to capture change in practices
- Analysis of data collected by delivery team such as progress reports used with the quarterly monitoring meetings, monthly phone calls, intervention suggestions or feedback surveys
- Analysis of outcome data as collected through the Early Years Toolbox

An initial logic model was provided by the delivery team prior to the start of the pilot. This will be reviewed throughout the course of the pilot, including in response to feedback from expert stakeholders and through a series of IDEA workshops. An initial discussion of the theory of change for the intervention took place at the first set up meeting for the pilot study. The theory of change was assessed again in the first IDEA workshop, which took place after the practitioner training, and also included a review of the programme materials as well as looking at the feedback from practitioners on the training sessions, as collected by Action for Children. The final two workshops will use Connell and Klem's (2000, cited in Humphrey et al., 2016) four criteria for assessment of quality of theory of change: Plausible (is it logical and based on good assumptions?), Doable (can the settings resource it?), Testable (what methods can be used to test it?) and meaningful (is it worth the time/money?) using external feedback. The second workshop will focus on expert stakeholder feedback to the theory of change and will update it as necessary based around these criteria. The final workshop will focus on the practitioner feedback from the interviews and case studies and allow final adaptations to the theory of change model as necessary. This final workshop will also provide an opportunity to collect cost data from the delivery team. This feedback loop will allow the intervention to be fully developed over the course of the 11 months.

Given that this is a pilot programme, it is necessary and appropriate at this stage to use qualitative research methods (in this instance, case studies, semi-structured interviews and observations) as the primary modes of data collection. Broadly, the data will provide rich insight into nurseries' experience of training and implementation, the feasibility of the intervention and its readiness for trial. This will be combined with quantitative and qualitative survey data and child outcome data to triangulate learning.

The case studies will allow the evaluation team to explore the perspectives of nursery managers and practitioners, especially their experiences of the training and support mechanisms, using the EYT resources and integrating these into their routine practice. The semi-structured interviews will explore the feasibility of the intervention including any challenges or barriers to delivery. In addition, we will

---

[2] Therefore in total, 24 of the 27 settings will receive either a visit or a telephone interview (ultimately the number of settings recruited was slightly higher than initially anticipated). The 3 settings that will not be interviewed will largely be dependent on which settings respond; where possible, we will endeavour to ensure that the 3 settings not interviewed are not systematically different from those that are, but we acknowledge that settings where it is more difficult to arrange interviews may potentially be those which are less engaged with the intervention or where there are other issues.

explore perceptions of potential impact on practices and outcomes, and ask whether they recommend the programme in its current format, and identify any recommendations or formative findings for subsequent rollouts. We would also seek to understand in particular whether settings believed the intensity of the programme was appropriate, especially in terms of the number and frequency of support sessions. We will also collect cost and time data from these settings. We will visit 12 nurseries. We propose to visit 6 of these settings midway through the programme (around February/March), after they have initially started using the EYT, to explore their initial experiences. By doing so, this would give the potential to make adaptations to delivery within the life of the pilot (if, for example, this showed that settings needed more support in terms of signposting to materials to enhance children's development). The remaining 6 visits will take place towards the end of the programme, to capture experiences over the life of the pilot.

The semi-structured telephone interviews of nursery managers and practitioners in the other 12 settings will allow an exploration of the same topics as the case studies but provide a more cost effective way of including all settings. The interviews will be short (approx. 30 minutes) and informal to avoid overburdening staff, and will be conducted at a time convenient for the participants. Again, we propose that 6 telephone interviews are conducted midway through the programme with the remaining 6 taking place towards the end of the programme. We have also allowed for telephone interviews with up to 5 settings that drop out of the intervention over the course of the year. This will explore barriers to continuing with the intervention and any potential solutions that could be offered.

The observations of the training and quarterly monitoring visits as well as a review of project materials, will supplement the semi-structured interviews by increasing the evaluation team's understanding of what the pilot intervention looks like in practice, and how it evolves during the course of the programme. The observations will also provide naturalistic impressions of how contextual factors affect implementation, what aspects of delivery are most salient and which are somewhat removed from the potentially biased perceptions of interviewees. We will observe the initial training day in all three regions, the follow-up session in three settings and then three quarterly monitoring visits in settings. The observations of the quarterly monitoring visits will take place in the final term.

The **baseline and post-treatment surveys of practitioners** will be focused on establishing an indicator of existing practices, and assess how practice and behaviours have changed during the programme. As we gather in-depth data on implementation and experiences in each of the participating nurseries during the interview stage, the survey will be very brief and focused almost exclusively on aspects such as understanding of child skills, confidence and leadership in their settings. We welcome any discussions with the delivery team as to how we can liaise to avoid duplication of data collection. Both the baseline and post-treatment survey will be conducted online.

Finally, we will review and analyse **data shared by the delivery team** during the pilot such as the data (and attendance) from the monthly calls, helpline and email enquiries (including, where possible, a log of the types of issues discussed during these), quarterly progress meetings and any feedback forms completed by the practitioners. These data will be analysed in light of the purpose of its collection, and data would be triangulated with the main evaluation findings. The delivery team will also inform the evaluation team if any nursery settings decide to withdraw from the programme, in which case they would be contacted by the evaluation team and asked about their reasons for doing so.

The following table summarises how the methods described above will be used to address each of the research questions, along with suggestions for indicators as to how these will be assessed.

Research question	Indicator	Data collection method
Evidence of promise		
Is there evidence of an improvement in practitioners' practice and behaviour to better support children's early development?	Self-reported change in practice by practitioners Setting managers' perception of change in practice	Surveys Case study visits and telephone interviews
Is there evidence of an improvement in practitioners' understanding of children's competencies and skills?	Change in understanding, based on practitioner self-report	Surveys

	Setting managers' perception of change in understanding of practitioners	Case study visits and telephone interviews
Do participants perceive a change in work practices from engaging in the intervention?	Percentage of practitioners perceiving change in work practices	Surveys Case study visits and telephone interviews
Do participants believe that the intervention has had (or will have) benefits to children compared to previous practices?	Percentage of practitioners who agree intervention has benefits for children	Surveys Case study visits and telephone interviews
Are there any unintended or negative effects of the intervention?	Type and prevalence of any negative or unintended effects	Case study visits and telephone interviews
Are practitioners more confident that they are correctly targeting children in need of support?	Change in percentage of practitioners who agree they are confident in correctly identifying children in need of support	Surveys Case study visits and telephone interviews
Feasibility		
Is the intervention feasible for practitioners to engage in and implement?	Experiences of practitioners and setting managers in relation to the intervention Issues covered in providing support Frequency of use of apps	Case study visits and telephone interviews Analysis of log of issues covered in monthly calls, emails and helpline enquiries and observation of quarterly monitoring visits Analysis of Toolbox data
Can the intervention be delivered within routine practice?	Participants' views on ease of using app within routine practice Frequency and timing of use of apps	Surveys Case study visits and telephone interviews Analysis of Toolbox data
What are the facilitators/barriers to the programme?	Participants' reports of facilitators and barriers	Case study visits and telephone interviews Telephone interviews with any settings that drop out of the intervention
What level of support is needed to enable practitioners to effectively interpret assessment data?	Type and extent of support required by practitioners Participant satisfaction with support received	Case study visits and telephone interviews Analysis of log of issues covered in monthly calls and quarterly monitoring visits Surveys
What level of support is needed to enable practitioners to effectively change their practice based on this data?	Type and extent of support required by practitioners	Case study visits and telephone interviews Analysis of log of issues covered in monthly calls and quarterly monitoring visits Surveys

What is the optimal / most feasible frequency and timing of the assessments within the school term?	Practitioners' reports on experiences regarding timing/frequency of use Frequency and timing of use of apps	Case study visits and telephone interviews Analysis of Toolbox data
Scalability		
Does the training and resources constitute a standard intervention that can be replicated and scaled up?	Extent to which training could be standardised and scaled up Participants' experiences of training and using the app	Observation of training and quarterly monitoring visits with training Review of materials Feedback forms from training days Case study visits and telephone interviews
What, if anything, should be updated or changed for a trial evaluation (e.g. theory, materials, procedures)?	Participants' experiences of the programme	Case study visits and telephone interviews Surveys Telephone interviews with any settings that drop out of the intervention
What will the cost of the intervention be to settings (including direct costs of training, indirect costs e.g. of interventions bought as a result of the assessments, and costs in terms of staff time)?	Direct costs of provision of training and resources as reported by delivery team Additional costs reported by settings (direct, indirect and time required)	Cost information collected from delivery team Case study visits and telephone interviews

Data analysis

- *Describe your approach to analysing the data collected as part of the pilot, including qualitative and any quantitative data.*

Qualitative data analysis: Interviews will be digitally recorded with the agreement of participants and transcribed verbatim. We will analyse the data using a 'framework' approach, drawing themes and messages from an analysis of interview transcripts, observations of training and quarterly monitoring sessions, and other materials collected by evaluation and project teams.

Analysis of survey data: We propose to keep this light touch looking at change over time for practitioners in the key outcome areas.

Analysis of child outcome data, collected through the EYT app, will also provide a means of monitoring children's progress over the course of the intervention. Given the stage of the trial, the main purpose of this analysis will be to explore frequency and use of the app, rather than for the assessment of evidence of promise. We will consider use and frequency of the app both for the 3 games in which practitioners are trained and the 3 games in which practitioners are not trained, but will have access to. We will also look for any variation across settings, both in terms of use and in

terms of outcomes, where possible linked to other information collected from settings as part of the evaluation. We understand that we will be able to access this data directly from the delivery team, and that we will be provided with two sets of data, one midway through the study (around February 2019) and the other at the end of the programme. Outcomes will also be compared against the published norms for the EYT. In addition to the analysis conducted for the purposes of this study, it will also be valuable for the pilot to establish whether the EYT may be a feasible means of collecting children's outcomes for other early years' trials.

Ethics and registration

The project has received ethical approval from the NIESR Ethics Committee. This requires the submission of an application form to the ethics committee, outlining the key features of the project, and setting out the ethical issues involved and mitigations.

Settings were provided with a Memorandum of Understanding, explaining what the project entails and the responsibilities of the evaluators, the delivery team and participating settings. Parents of eligible children were sent a letter explaining the study and giving the option for parents to withdraw their children from the study, which they are able to do at any time.

Personnel

Delivery team:

- Sarah Read (Action for Children)
- Fliss James (Action for Children)
- Professor Edward Melhuish (Oxford)

Evaluation team:

- Anneka Dawson (Co-PI, IES)
- Lucy Stokes (Co-PI, NIESR)
- Nathan Hudson-Sharp (NIESR)
- Clare Huxley (IES)
- Chiara Manzoni (NIESR)
- Johnny Runge (NIESR)
- Helena Takala (IES)
- Ceri Williams (IES)
- Heather Rolfe (NIESR)

Risks

The table below sets out the key risks associated with the project, and how they might be addressed. Risks will be actively monitored and re-evaluated, emerging or growing risks identified and discussed with EEF and appropriate control measures put in place.

Issue/risk	Action to address issue/reduce risk	Likelihood	Impact
Data quality issues relating to information collected as part of the pilot (for example, sufficient and complete	We will work with the delivery team to develop a format for logging the issues covered during support calls/visits. These will be reviewed	Low	Medium

Issue/risk	Action to address issue/reduce risk	Likelihood	Impact
information on log of topics covered during calls)	early on in the project to check they are working as intended. We will also receive an initial set of Toolbox data partway through the pilot, which will enable early exploration of any issues that may need to be addressed.		
Lapse in data security	NIESR and IES both have rigorous data security policies which meet Government (e.g. DWP) guidelines and the requirements of the Data Protection Act 2018. We will ensure there are robust measures in place to collect and transfer data securely and that participants in the research are kept anonymous (using password protected codes for interview files for example). We will detail these measures in full in a data sharing agreement with EEF and the delivery team.	Low	High
Reluctance of settings to participate in case study research/ survey research/ observations	Settings will have their responsibilities clearly laid out in a Memorandum of Understanding which they will sign when recruited. We will work to reduce the burden on participants as much as possible with short surveys and interview times and having a wide staff across two teams means we can be flexible about dates for case studies	Low	High
Settings drop out of pilot	Evaluators have limited ability to affect participant numbers other than explore reasons for low participation, identify good practice in maximising participant engagement and share across settings. We can work to maximise research participation among those who take part in the pilots. We will speak to any settings dropping out to establish why and seek to resolve any issues with the support of the delivery team.	Low	Medium
Illness of key personnel or staff leaving	The team has an experienced pair of co-principal investigators who can cover for each other. Sufficiently large team to ensure other team members can take the place of any leaver. In the event of long-term illness or staff departures, both organisations will bring in additional staff from their pools of 30+ other researchers – a number of whom also have experience of conducting similar research.	Low	Low

We recognise that data protection is of the utmost importance and are fully committed to complying with the Data Protection Act 2018 and GDPR legislation.

Our proposed approach includes the collection of personal data. Our grounds for processing personal data is legitimate interests, and the evaluation team has conducted a legitimate interest assessment.

Practitioners and other participants interviewed for the research will be asked to sign a consent form or give verbal consent in the case of telephone interviews to indicate that they have understood the aims of the research and agree to the interview being recorded and transcribed. They will also be given written assurance of anonymity and confidentiality for themselves and their setting. Contact details of settings and staff taking part in the research will be kept on password protected files accessed only by the research team. No setting or individual will be identified in the report or any other outputs of the evaluation.

A privacy notice has been made publicly available explaining how information collected from participants would be used and stored, and communicates to participants and parents their right to withdraw from data processing at any time.

A detailed data sharing agreement is being developed between both evaluation teams, the delivery team and EEF which states which data will be shared by who, how and why to ensure full data security throughout the project.

Timeline

Dates	Activity	Responsibility
Aug 2018	Set- up meeting	Delivery team and evaluation team
Aug-Sept 2018	Recruitment of settings including preparation and signature of MoU and other recruitment materials	Delivery team and evaluation team
Sept 2018	Usual practice questions and baseline setting survey	Evaluation team
Oct 2018	Deliver training for practitioners	Delivery team Evaluation team to observe training session
October	Observations of initial training sessions	Evaluation team
Nov 2018	IDEA workshop 1- feedback on training	Delivery team and evaluation team
Oct- Dec 2018	Observations of follow- up training	Evaluation team
Oct 18- June 2019	Programme delivery	Delivery team
Nov 2018 -Jan 2019	Stakeholder feedback on ToC model	Evaluation team
Feb- March 2019	First set of case study visits and telephone interviews	Evaluation team
Feb-Mar 2019	IDEA workshop 2- stakeholder feedback integrated	Delivery team and evaluation team
Jan - Mar 2019	Observation of quarterly monitoring visit and monthly phone calls	Evaluation team
May-July 2019	Second set of case study visits and telephone interviews with other settings	Evaluation team
June-July 2019	Post-tests of practitioner survey Attendance at end of project event	Evaluation team
July 2019	IDEA workshop 3- final session with review of year and cost data collected	Delivery team and evaluation team
October 2018-July 2019	Collection of project data from delivery team (including monthly report data and support data) and collection of EYT data	Evaluation team
July-Aug 2019	Analysis of project and evaluation data	Evaluation team
Aug 2019	Presentation of findings	Evaluation team
Sept - Nov 2019	Report writing	Evaluation team
Nov 29th 2019	First draft of evaluation report	Evaluation team