Education Endowment Foundation

Early Number with Numicon

Pilot report

December 2025

Sashka Dimova, Maya Agur, Jenny Hull, and Daisy Woods

IFF Research

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

We do this by:

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

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

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

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

info@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents

About the evaluator

The Early Number with Numicon programme was independently evaluated by IFF Research. The research team comprised of Dr Sashka Dimova who led the evaluation, Maya Agur and Jenny Hull who managed the evaluation day-to-day, and Daisy Woods who supported the team with the day-to-day running of the project. The team was also supported by Sophie Harvey-Rich and Simran Sangha in collecting data.

Contact details:

Dr Sashka Dimova
IFF Research
5th Floor, The Harlequin Building,
65 Southwark Street,
London,
SE1 0HR

Email: Sashka.Dimova@iffresearch.com

Executive summary

The project

Early Number with Numicon is a professional development programme designed and delivered by Oxford University Press (OUP). The programme aims to improve early years practitioners' knowledge and confidence in developing early numeracy skills among children aged three and four years, focusing on pattern, counting, comparison, and composition. The programme facilitators trained practitioners to deliver short structured weekly small group sessions, while also enhancing continuous mathematics provision throughout the setting. The 14-week programme included four training sessions (two in person and two online) and provided all necessary Numicon materials. Practitioners were encouraged to adapt activities to suit individual children's needs and to embed numeracy into everyday practice. Delivery of the training and activities took place between September 2024 and February 2025.

The Education Endowment Foundation (EEF) supported this pilot based on evidence from its Improving Mathematics in the Early Years and Key Stage 1: Guidance Report, which highlights the effectiveness of using manipulatives and representations in teaching mathematics (Clark *et al.*, 2020). Early numeracy approaches are also among the most impactful strands in the EEF Early Years Toolkit, with an average learning gain of seven months (EEF, 2025). However, the security of this evidence is currently rated as limited (2 out of 5 padlocks), making further research a priority.

The pilot, completed by IFF Research, assesses the programme's feasibility of implementation, evidence of promise, and readiness for trial. A total of 31 settings were recruited to take part, comprising of both private, voluntary, and independent (PVI) nurseries (17) and childminder settings (14). The evaluation focused on practitioners, and no data related to the children in the programme was collected.

The evaluation used a mixed-method design involving online surveys with practitioners at three key points (before, during, and after delivery). Setting leads in PVIs were also surveyed twice, before and after programme delivery. Case studies were conducted with eight purposively sampled settings, including in-depth interviews and observations before, during, and after delivery. All four training sessions were observed, and two OUP delivery team members were interviewed after programme delivery.

Table 1: Summary of pilot findings

Research question	Findings
Evidence of promise: Is there evidence of promise that the programme may lead to the changes expected in the Theory of Change?	There is strong qualitative and quantitative evidence suggesting that Early Number with Numicon can improve practitioners' knowledge and confidence in developing early numeracy as well as children's early numeracy development. Practitioners reported increases in their self-rated knowledge of specific numeracy areas and greater confidence in teaching early numeracy skills from baseline to endline. While some practitioners began with relatively high levels of confidence, improvements were observed across all measured areas. All surveyed practitioners reported that Numicon positively impacted children's early numeracy development in most aspects of children's numeracy. By the end of the programme, all surveyed practitioners reported using and adapting the Numicon resources in their setting, indicating that the Numicon approach had become embedded in their practice.
Feasibility of implementation: Is Early Number with Numicon feasible and acceptable to PVIs and childminders?	Early Number with Numicon was considered feasible to implement and acceptable to practitioners. Most practitioners attended all training sessions and reported feeling well-prepared to deliver the programme in their setting. The programme comprises 11 structured learning sequences, designed to be delivered progressively. Reaching at least learning sequence 7 (around two-thirds of the programme) was identified as a key benchmark for meaningful participation. By the end of the programme, 15 settings (out of 21) had reached this point or beyond, indicating strong engagement. Fidelity was rated medium to high across all settings, with no major barriers preventing delivery. However, some settings did encounter some practical challenges in delivery, such as managing

Research question	Findings
	the participation of children who attended part-time or inconsistently, and balancing delivery with the care needed of children aged two years old or under. Despite these challenges, practitioners adapted creatively to ensure the programme was accessible to most children. Early Number with Numicon was considered relevant and easy to integrate into daily practice with three–four-year-olds, with only minor adjustments needed. Overall, the programme was considered both practical to implement and appropriate for a wide range of early years settings.
Readiness for trial: How feasible is it to scale delivery (and undertake evaluation using a randomised controlled trial design) with this population (i.e. PVIs and childminders)?	The pilot evidence suggests that Early Number with Numicon is showing strong readiness for scale-up. Engagement with the programme activities was high, with 29 of the 31 recruited settings remaining engaged with the programme through to the end in some capacity. While data collection was adequate for the pilot (e.g. attendance monitoring, short post-training surveys) enhancements should be considered for an efficacy trial. These may include standardised fidelity tools, more formal monitoring of implementation, and processes for tracking children's progress across settings. The evaluation concluded that the programme can be scaled for an efficacy trial with minimal changes for implementation into PVI and childminder settings. Recommended refinements include adapting training for larger and more diverse practitioner groups, providing specific guidance for working with younger children and those with Special Educational Needs and Disabilities (SEND) or English as an Additional Language (EAL), improving recruitment of childminders through better outreach and messaging, and reviewing how programme fidelity is measured to ensure it reflects meaningful progress.

Additional findings

The training element of the programme received high praise, with practitioners valuing the blend of in-person and online sessions. Most preferred in-person sessions, but the mixed delivery format made it easier for practitioners to attend and engage with the training.

Practitioners valued the programme's flexibility, and ability to adapt the activities to suit their setting and children's needs, including younger children. Many practitioners reported a shift towards more structured and confident numeracy teaching as a result of using Early Number with Numicon.

The programme was inclusive, with many practitioners successfully adapting delivery for children with SEND and EAL. However, some requested more guidance for how to adapt with these groups where there were more complex or individual needs. Continuous provision using Early Number with Numicon resources was widely adopted, and this was achieved in a variety of ways by different settings (e.g. using large Numicon shapes during play time, or putting Numicon manipulatives in the sand pit outside of structured activities). The impact of this depended on how well it was integrated into daily practice.

The programme was considered low-cost and manageable in terms of time commitment. Most settings did not incur additional expenses, and those that did typically spent small amounts on supplementary materials. Practitioners were compensated for their time attending the four training sessions, which consisted of two face-to-face sessions (three hours each) and two online sessions (1 hour 45 minutes each).

Introduction

Existing evidence

Studies have shown that high-quality mathematics education in the early years can have a significant impact on children's later mathematical performance (Raghubar and Barnes, 2017). Early numeracy practices are of particular interest for the Education Endowment Foundation (EEF), as they are one of the most impactful approaches in improving children's numeracy development. A recent review undertaken by the EEF identified several elements, which can be effective in improving early numeracy development (Clark *et al.,* 2020). These include professional development for practitioners, creating opportunities for practitioners to discuss mathematical ideas with children, and the use of physical manipulatives to increase children's ability to understand and engage with mathematical ideas (Hodgen *et al.,* 2020). However, there remain substantive evidence gaps in relation to which specific approaches to early mathematics learning are the most effective.

Evidence also highlights several structural challenges in the early years sector. These include a high variability in qualifications among staff, high proportions of children eligible for Early Years Pupil Premium (EYPP) (Department for Education, 2024), and inconsistent training across the early years' workforce in different setting types (due to limited available funding). For example, less than half of private, voluntary, and independent (PVIs) settings employ a qualified early years teacher or equivalent, while all maintained nursery schools must employ a member of staff with qualified teacher status (Cattoretti *et al.,* 2019; Bonetti and Landen, 2020). These challenges were further compounded by the Covid-19 pandemic, which put increased financial pressure on the sector and exacerbated issues around recruiting and retaining highly qualified staff (Parliamentary Office for Science and Technology, 2021).

The Early Number with Numicon pilot is an extension to the pre-pilot work undertaken with Ryders Hayes Thrive Together Early Years Stronger Practice Hub over the Spring Term and Summer Term in 2023, with 12 settings, (four childminders, and eight private and school-based nurseries). The pre-pilot work was undertaken by the Oxford University Press (OUP) delivery team. Support was given from the early-stage pipeline development (ESPD) team at the EEF.

Findings from the pre-pilot highlighted that children acquire numeracy skills at varying rates, and some settings required more time to consolidate children's understanding before progressing to the subsequent activity. In response, the OUP decided to extend the length of the programme for the pilot from 8 to 14 weeks. The additional time was intended to allow more flexibility in delivery and to provide additional support to children who need to work at a slower pace or lack understanding of underlying mathematical concepts.

In 2023, IFF Research was commissioned by the EEF to carry out a pilot evaluation of the Early Number with Numicon programme. Although early findings from the pre-pilot suggest that the programme has promise, this pilot evaluation is the first independent evaluation with a specific focus on PVIs and childminders, and as such offers a unique opportunity to strengthen the evidence base for effective delivery in early years mathematics.

Intervention background

The pilot workplan followed the Template for Intervention Description and Replication (TIDieR) framework. This helped ensure that Early Number with Numicon is well-defined, replicable, and can be properly evaluated. The full TIDieR checklist is provided in Appendix A. The programme design reflects key elements identified in existing evidence as effective for early numeracy, including professional development for practitioners, opportunities for mathematical talk with children, and the use of physical manipulatives (Clark *et al.,* 2020). In this way, Early Number with Numicon directly addresses evidence gaps by combining these features into a structured and replicable professional development offer.

What

Early Number with Numicon is a professional development programme that aims to improve practitioners' knowledge and confidence in supporting early numeracy skills in three–four-year-olds. It provides training for practitioners to deliver short (c. 15 minutes) weekly small group (ideally not more than four children) sessions and enhance their continuous

mathematics provision based on the Numicon resources.[1] The programme aims to provide practical support specifically to practitioners in PVIs and childminders (Paull and Popov, 2019).

The Early Number with Numicon programme was developed by the OUP, who own the wider Numicon brand and associated resources. The Early Number with Numicon approach helps children see connections between numbers through a multi-sensory approach, which means children learn through seeing and feeling, for example, by exploring physical resources or mimicking language used in the session.

How

The training is designed to develop practitioners' knowledge and confidence in delivering mathematics sessions to children. The training took practitioners through the process of planning and delivering programme activities and covered five topics: i) getting started (understanding basic concepts of size, colour, and difference); ii) patterns (recognising, copying, and creating patterns); iii) counting (understanding counting principles); iv) comparison (same and different, more and less); and v) composition (working with numbers up to and beyond 5). The training supported and encouraged practitioners to enhance the mathematics opportunities available to children, through their continuous provision. The training was delivered by the head of professional development at the OUP and a lead Numicon trainer. Both trainers have extensive experience working with Numicon and in delivering professional development to practitioners.

The training followed a blended model, combining two face-to-face (three hours each) and two online sessions (1 hour and 45 hours each). The trainers offered individual catch-up sessions to practitioners who missed a training session, and online sessions were recorded and shared with those who missed them or needed a re-cap. The sessions were designed in this blended way to be flexible and accessible, and mindful of the time pressures and other demands of practitioners' lives. Between training sessions, practitioners delivered the activities they had been trained on to the children at their settings. Each activity was delivered to a small group of children (ideally no more than four children). Practitioners were asked to include all pre-Reception-aged children (aged three and four) in activities. Practitioners had the option to repeat content if children had not grasped the concept being taught and to reinforce learning through linked activities in continuous provision.

During the delivery of the pilot there was continued support to practitioners via individual check-ins by the OUP delivery team (one visit at each setting) and a professional learning community was encouraged between practitioners to provide a platform for sharing ideas and good practice. Additionally, each practitioner was encouraged to have a buddy from another setting in the programme, who could support them if they missed one or more of the training sessions.

Practitioners were encouraged to deliver the sessions as guided by the book and their training. However, there was flexibility in the design of the programme, to enable practitioners to make adaptations and tailor the sessions to align better with the needs of the children they are working with and the context of their setting. Adaptations were mainly made to link the activities to the topics of focus at the time in the setting or to the resources available at the settings.

Who

The programme specifically targeted practitioners working in PVI nurseries and childminder settings. These settings were targeted because they typically have less access to regular Continuing Professional Development (CPD) opportunities compared to maintained settings. As a result, there is limited evidence on what works effectively in PVI and childminder contexts, particularly in areas such as practitioner qualifications and CPD offer, which was of particular interest for the EEF. A total of 29 practitioners from 29 settings completed the programme: 13 from PVI settings and 16 from childminder settings. Taking part involved attending four training sessions (two sessions in person at a venue in Bradford and two sessions online) and delivering activities using a workbook in their setting. In the PVI settings, a senior practitioner, nursery

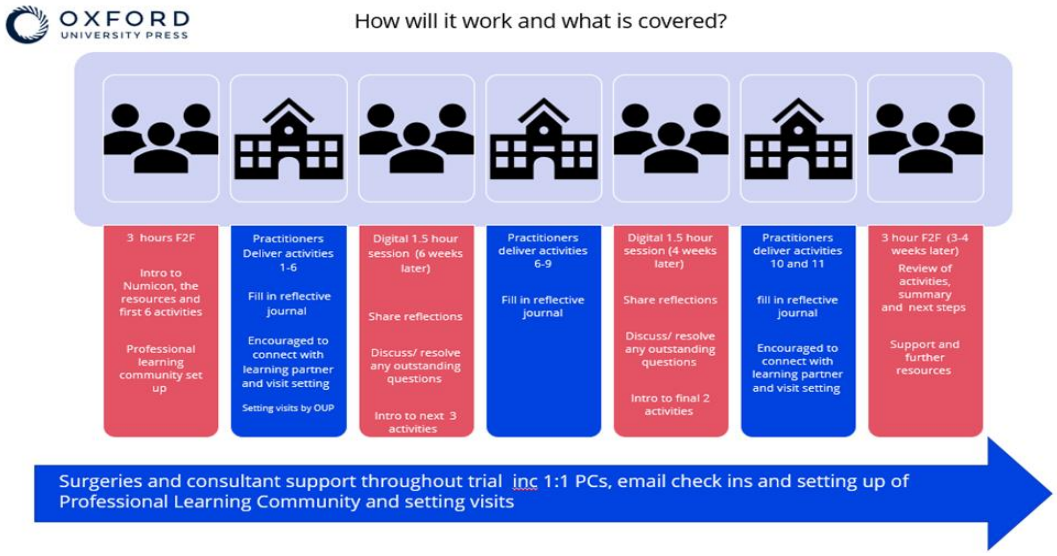
[1] Each participant received the activity book with guidance for practitioners and a set of Numicon resources including two boxes of Numicon shapes, Pegs and baseboard, Number overlays, Picture overlays, Numicon large foam shapes, Zig-Zag book, Cuisenaire (number rods), counters, and Pan balance.

manager, or member of the senior leadership team (SLT) was also invited to attend the training. This was to promote senior buy-in and enable a cascade model of passing on the knowledge to other practitioners at the settings and ensure the sustainability of the approach beyond the pilot. Each individual practitioner was compensated £7.50 per hour to attend the training. No prior knowledge of mathematics teaching was required. All participating settings received Numicon manipulatives and other mathematics apparatus to implement the programme.

Practitioners were asked to deliver Numicon activities to the three–four-year-old children in their settings between training sessions, based on their learning at the training. They were also asked to take notes and send feedback on their reflections of the programme, how the activities were received by the children, and what they might do differently next time. These reflections were collected by the developers and were used to further develop the programme.

The programme also provided continuous support to practitioners via email exchange and via individual check-ins by the OUP delivery team, and a professional learning community that was created with the support of the OUP and was maintained, to some extent, by practitioners. Practitioners were encouraged to pair-up with another practitioner in the pilot (by way of a buddy system) to observe, support each other's delivery, and share ideas and feedback.

Figure 1: Delivery overview



Source: The OUP.

Where

Delivery of the training and activities was spread across 14 weeks between September 2024 and February 2025. The first training session lasted three hours and was face-to-face (the session took place on a Saturday, a non-working day to enable participation in person). The middle sessions (sessions 2 and 3) were online twilight sessions that were one and a half hours long. The final session was face-to-face and three hours long (this session also took place on a Saturday to enable participation in person). Activities for children took place at least once a week with each activity lasting 15 minutes.

The pilot recruited settings from the West Yorkshire area. The first and final training sessions took place in Bradford, a location central to all of those taking part in the training, to ensure it was easily accessible to all. The activities for children were delivered by the practitioners in their own settings.

Research questions

Following the Implementation, Delivery, and Evaluation Analysis (IDEA) workshop and set-up meetings with the EEF and the OUP, a set of research questions and success indicators were developed in close collaboration and in line with the three pilot criteria for this evaluation.

Table 2 provides an overview of the research questions.

Table 2: Research questions

Pilot criteria	Research question
Evidence of promise	1.1 Is there a change in practitioners' knowledge and confidence in delivering numeracy sessions after participating in the programme? • Are the four personal development sessions perceived to be relevant in influencing the practitioners' knowledge and confidence? • Is the professional learning community perceived to be useful? • Is the buddy system perceived to be useful? • What new knowledge and practices do practitioners report learning? • Does this vary by setting type or practitioner's prior knowledge and experience?
	1.2 What, if any impact, do practitioners feel the programme has had on increasing children's engagement and knowledge in early numbers? • What is the contribution of the small group activities? • What is the contribution of the additional three introductory activities? • What is the contribution of the continuous provision activities? • What impact do practitioners feel the programme is having on children from disadvantaged backgrounds especially for children with EYPP, English as an Additional Language (EAL), and Special Educational Needs and Disabilities (SEND)?
	1.3 How different is the approach to other mathematics resources and programmes used? • What is the usual setting practice of early numeracy approaches?
	1.4 Did practitioners' early numeracy approach change (and if so, how) as a result of the programme? • Did practitioners use the resources? • Did this vary by setting type?
Feasibility of implementation	2.1 What is the engagement with the training sessions? • What are the completion rates for the training? • What are practitioners' views on the number of personal development sessions, duration, and hybrid training model?
	2.2 What is the engagement with the programme activities (i.e. small group activities and continuous provision for children)? • Did practitioners deliver the required activities? • What are practitioners' views on the programme's dosage and content? • What adaptations, if any, were made?
	2.3 What contextual factors act as barriers and facilitators for practitioners in terms of: • Attending the personal development sessions. • Delivering the small group activities and continuous provision. • Do these vary for PVIs and childminders?
	2.4 What are the barriers and enablers for children from disadvantaged backgrounds accessing the programme especially for children with EYPP, EAL, and SEND?
	2.5 How acceptable is Early Number with Numicon to PVIs and childminders? • Is the training accessible to practitioners in all settings? • How is the programme received in different settings?
Readiness for trial	3.1 To what extent does the recruitment strategy recruit and retain the settings and practitioners? • What are the different challenges depending on the type of setting?
	3.2 What does the pilot tell us about the feasibility of data collection?
	3.3 How feasible is it to conduct an evaluation using a randomised controlled trial (RCT) design with PVIs and childminders? • What are the benefits and disadvantages of estimating one treatment effect for PVIs and childminders combined, compared to estimating separate effects? • What are the implications for sample size?

Pilot criteria	Research question
	3.4 What costs/resources are required by settings implementing Early Number with Numicon and is this feasible and acceptable, at a larger scale?

Ethical review

The evaluation of Early Number with Numicon was considered low risk as the evaluation did not collect personal data on children, only on adults, on a small scale, and did not include personal or sensitive data. Thus, a full ethical review was deemed unnecessary. Instead, the IFF Research evaluation team developed an ethics and safeguarding statement that set out how the researchers will ensure the evaluation activities will be accessible to participants and will not add any undue stress or burden.

In setting out the ethical statement, IFF Research was informed by the requirement of the Government Social Research's (GSR's) 'Ethical Assurance for Social and Behavioural Research' (GSR, 2021) as well as the ethical guidance of the Market Research Society 'Code of Conduct' (Market Research Society, 2023) and the Social Research Association's 'Research Ethics Guidance' (Social Research Association, 2021). IFF Research are actively involved in both organisations. Key ethical considerations that were considered include:

- ***Informed consent.*** *Researchers ensured that respondents gave explicit and informed consent.*
- ***Ensuring accessibility of participation.*** *IFF Research adopted the social model of accessibility. IFF Research Ethos means that researchers firmly believe there are no hard-to-reach participants.*
- ***Avoiding personal and social harm.*** *The statement took the widest definition of harm.*
- ***Protecting confidentiality.*** *Data protection protocols for protecting confidentiality are described below.*
- ***Upholding public sector equality duty.*** *IFF Research believe that equality, diversity, and inclusion is a vital part of a successful organisation.*

The pilot was registered with the Open Science Framework (OSF) registry in August 2024 (OSF, 2024).

Data protection

Data protection procedures were developed for the project in collaboration with the EEF. For this pilot, IFF Research was the data controller who also processed the data. As such, IFF Research produced privacy notices for those elements of the evaluation involving personal data and ensured the wording (and derivatives such as consent questions for recontact) and the implication of that wording was discussed with the EEF. In line with the UK General Data Protection Regulation (GDPR), personal information was securely stored and was only used for the purpose for which it was collected. All personal data will be destroyed when no longer required and no later than 12 months after the publication of the final report.

IFF Research data protection officer reviewed the data specifications for the pilot evaluation and concluded that it was a low-risk project with respect to its data security requirements. Therefore, the pilot evaluation did not require a full Data Protection Impact Assessment (DPIA) to be conducted. In delivering the pilot evaluation, IFF Research followed its usual data processing protocols, which are ISO 27001 and Cyber Essentials Plus certified.

The legal basis for processing personal data (GDPR Article 6) for this pilot evaluation is legitimate interest (GDPR, 2016). The processing of data collected in the pilot had clear social benefits for understanding how to undertake this type of research, with a limited privacy impact on the individual.

Project team

Delivery team: The OUP

- ***Will Power, Head of Maths:*** *Responsible for contracts, budget, recruitment, advice, and liaison for the delivery team.*
- ***Jayne Jarvis, Head of Maths:*** *From January 2025 took over Will Power's role following their departure from the OUP.*
- ***Louise Pennington, Head of Professional Development:*** *Responsible for recruitment, delivery lead, resourcing and personal development, and visit and support to settings.*
- ***Dr Penelope Woolf, Director of Impact and Learning Design:*** *Responsible for peer review, fed into research design from an OUP perspective.*
- ***Janine Wilson, Project Manager (Freelance):*** *Supported the team with the day-to-day running of the project including admin and documentation, timelines, budgets, and tracking for recruitment and delivery.*
- ***Helen Laflin, Lead Numicon Trainer for Early Years (Freelance):*** *Responsible for the creation of personal development training materials and delivery of training.*

Evaluation team: IFF Research

- ***Dr Sashka Dimova, Research Director:*** *Led the evaluation, with overall responsibility, covering design, analysis, and reporting deliverables.*
- ***Maya Agur, Research Director:*** *Lead project manager ensuring a seamless communication with the OUP and the EEF. Specific responsibility for project timelines, processes, and budget, as well as being hands-on in fieldwork, analysis, and reporting.*
- ***Jenny Hull, Senior Research Manager:*** *Supported Maya Agur in managing the evaluation in all elements of the pilot, with a particular focus on the qualitative fieldwork, analysis, and reporting.*
- ***Daisy Woods, Senior Research Executive:*** *Supported the team with the day-to-day running of the project.*

Methods

Recruitment

Recruitment for the pilot evaluation was conducted by the OUP delivery team with support from their existing contacts in the Bradford area. These included but were not limited to: The Bradford Research School, a Multi-Academy Trust; and the regional Early Years Stronger Practice Hub (Yorkshire and the Humber). As a first step for the recruitment, the OUP created a list of eligible settings to contact using various sources (GOV.UK website, county Education websites, and PVI websites). This yielded a list of 560 settings. Throughout this process, the recruitment lead sent out emails to each identified setting, providing information about the pilot and providing a link to the expression of interest (EOI) form. This was followed by phone calls to settings, where there was no response, when emails bounced-back or when further information was required to support recruitment. Once interested settings completed the EOI, the OUP followed up with additional information (information sheets for settings and parents, please see Appendix B for full documents) and a copy of the Memorandum of Understanding (MoU) for the settings to consider, complete, and sign. The recruitment lead from the OUP then prompted settings who did not return the signed MoU, to secure their participation in the pilot.

The initial target areas were geographically close to the centre of the Bradford borough such as Allerton, Girlington, Manningham (i.e. in and around the BD8 postcode area). However, once recruitment for the pilot study began, interest was received from settings outside of Bradford. As a result, the OUP delivery team widened recruitment to cover West Yorkshire and certain areas of Greater Manchester (e.g. Oldham), but keeping within a short driving distance from Bradford to enable attendance at the training sessions and trainers' visits to all settings. To ensure that the final list of settings in the pilot included an even number of PVIs and childminders, the recruitment lead focused on one type of setting at a time and moved their focus depending on the numbers signing up.

The recruitment target was 30 settings, with a roughly equal split between PVIs and childminders. This target was deemed feasible due to a sufficient number of eligible PVIs and childminders (i.e. there were 109 PVIs and 73 registered childminders in Bradford alone). An equal split between PVIs and childminders was preferred to enable exploration of feasibility of implementation and readiness for trial across different setting types.

PVIs and childminders were considered eligible for the pilot evaluation if they met the following criteria:

- *Did not participate in another funded programme by the EEF or a Stronger Practice Hub-funded programme delivered in the 2024/2025 academic year.*
- *Had at least one child that was three–four-years old in the 2024/2025 academic year.*
- *Committed to releasing at least one early years practitioner to attend two face-to-face training sessions and two online sessions. Other staff members could also attend, but at least one practitioner had to attend all the sessions.*
- *Committed to one senior member of staff or setting manager attending the first training session.*
- *Provided time for the practitioner to deliver the weekly sessions with all suitable groups of children (three–four-year-olds).*
- *Provided time for the practitioner to plan their sessions and complete weekly activity logs and reflection sheets.*
- *Supported the pilot evaluation by taking part in the research activities detailed in the evaluation's MoU (Appendix C).*

A total of 34 settings initially signed the MoU. However, three settings were later excluded as they were identified as maintained nurseries and therefore, not eligible for inclusion in the pilot. Two additional settings dropped out: one (childminder) prior to the first training session; and another (PVI) after attending the first training session and completing the first tasks. This was due to resource and capacity constraints at the settings. One recruited setting participated in another numeracy programme. However, the OUP team confirmed that separate staff were responsible for different age groups, and

the setting had sufficient capacity to deliver both programmes. Given the pilot nature of the study, this setting was allowed to participate.

Ultimately, 31 settings started the Early Number with Numicon pilot and 29 of recruited settings remained engaged with the programme through to the end in some capacity (13 PVIs and 16 childminders; see Table 7 for a breakdown of the demographic profile of settings that took part in the pilot). Participants were all practitioners working with children three–four years old in the recruited settings. Children were not directly recruited to the study.

All settings who participated in the pilot had committed to taking part in the evaluation. This commitment was set out in the MoU they signed, which was developed by the OUP and IFF Research, collaboratively. To ensure informed consent was given by settings, IFF Research developed a pilot specific Privacy Notice detailing the data processing protocols for the pilot and setting out the participants' rights in line with the GDPR (GDPR, 2016). In addition, IFF Research developed information flyers for families who had children in settings that were selected as case studies. The flyers provided information about the pilot that was taking place, and informed families that researchers would be visiting the setting to observe activities, giving them the option to opt out of their children participating in any activity that was being observed.

Measures

IFF Research in collaboration with the OUP identified 13 success indicators for assessing the effectiveness of the programme. These were grouped under the three pilot criteria (EEF, 2022). Table 3 below lists the 13 success indicators and aligns each with the evaluation research questions and the respective assessment approach.

Table 3: Success indicators assessment framework

Pilot criteria	Research question	Success indicator	How this was assessed
Evidence of promise	1.1	Improvement in practitioners' knowledge and confidence in supporting children's early numeracy development across the majority of the PVIs and childminders	Surveys with practitioners; surveys with leads; trainer interviews; observations
	1.2	Practitioners (who take part in the training) in the majority of the PVIs and childminders perceive Early Number with Numicon to have a positive impact on children's early numeracy development	Surveys with practitioners; practitioner interviews; observations
	1.3; 1.4	Practitioners in the majority of PVIs and childminders make changes to their practice as a result of taking part in Early Number with Numicon	Surveys with practitioners/leads; case studies
Feasibility of implementation	2.1	Practitioners in the majority of PVIs and childminders attend all training sessions	Surveys with practitioners; observation of training; case studies
	2.2	Practitioners (who take part in the training) in the majority of PVIs and childminders complete the programme activities (i.e. small group activities and continuous provision for children)	Surveys with practitioners; case studies
	2.1; 2.2; 2.3	Early Number with Numicon is delivered with medium to high fidelity in all setting types as assessed with the fidelity measure (see Table 14)	Fidelity measure; surveys with practitioners; case studies
	2.3; 2.4	Absence of any major barriers[a] to delivery for majority of PVIs and childminders and for children from disadvantaged backgrounds	Discussions with the OUP delivery team; surveys with practitioners, case studies
	2.5	Practitioners in the majority of PVIs and childminders consider the intervention practical to implement (with minor amendments)	Surveys with practitioners; surveys with leads; case studies
Readiness for trial	3.1	High-level engagement with the pilot in terms of recruitment and low attrition	Surveys with leads ; surveys with practitioners; case studies; observation of training
	3.2	There are viable strategies in place to collect sufficient data to monitor delivery	The OUP delivery team and evaluator own assessment

Pilot criteria	Research question	Success indicator	How this was assessed
	3.3	Early Number with Numicon can be scaled for an efficacy trial (with no or minor amendments)	Case studies; the OUP delivery team and IFF Research evaluation team own assessment
	3.4	The resource commitment[b] required to deliver Early Number with Numicon is acceptable to majority of PVIs and childminders	Surveys with leads/practitioners; case studies; the OUP delivery team interviews

[a]*Some major barriers include high dropout among practitioners or children from disadvantaged backgrounds.*

[b]*It includes staff time required to take part in the training and time required to deliver the programme.*

Data collection

The evaluation brought together data from a range of research methods to answer the evaluation research questions, and to assess the success indicators. The methods were selected to collect views from all stakeholders and to evidence any difference in delivery across PVIs and childminders. The evaluation included multiple rounds of data collection to ensure valuable information captured change over time. The range of research methods used is illustrated in Table 4 below, where each research method is discussed in more detail.

Table 4: Summary of data collection and analysis

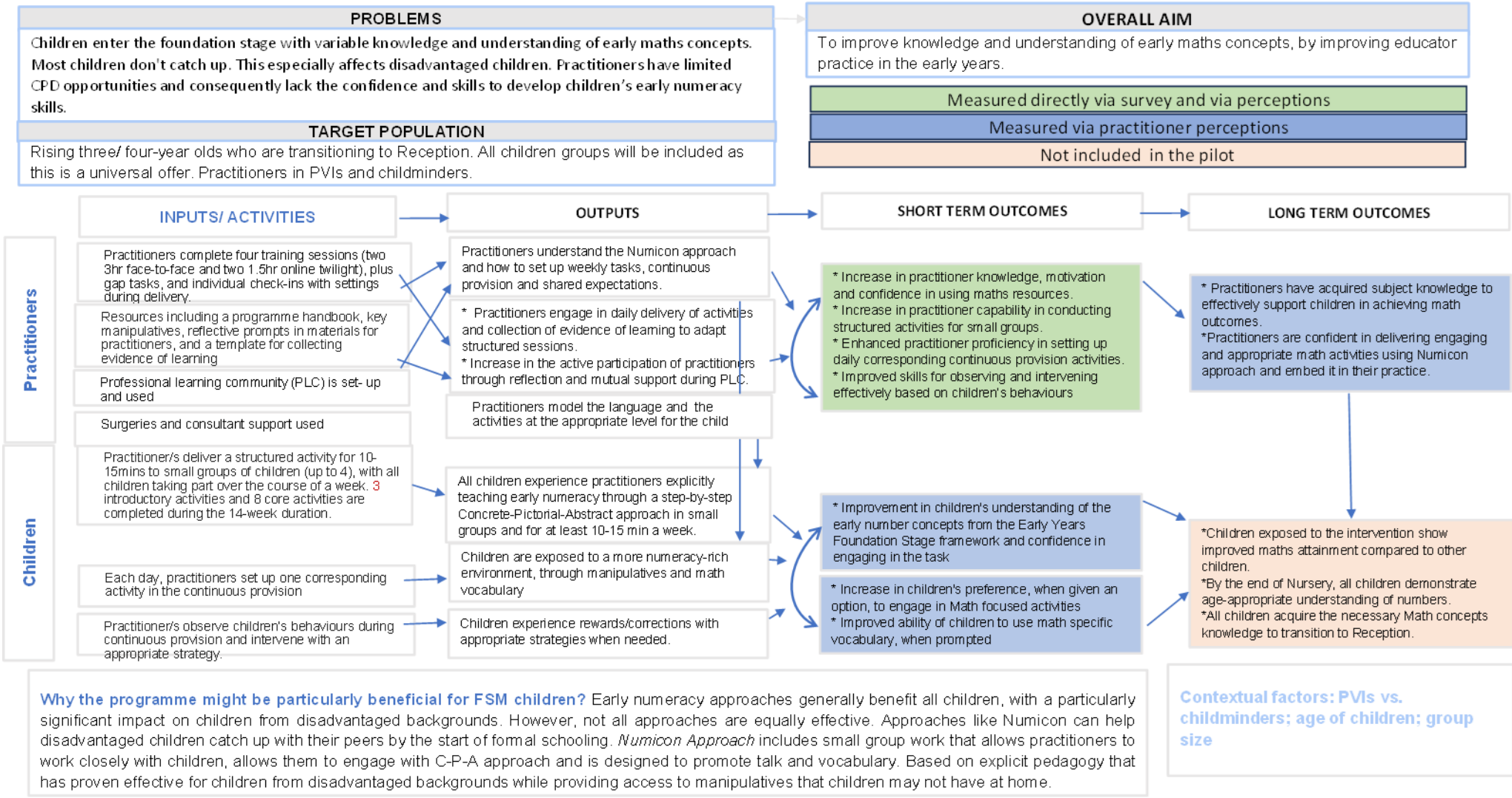
Research methods	Data collection methods	Sample size and sampling	Data analysis
Theory of Change updates	IDEA workshop and set-up meetings	Not applicable	Not applicable
Survey with practitioners (three rounds)[a]	Online questionnaire	All practitioners in recruited PVIs/childminders (29)	Descriptive statistics; inductive analysis of open-text responses; framework / thematic analysis
Survey with leads (two rounds)	Online questionnaire	All leads in PVIs (13)	Descriptive statistics; inductive analysis of open-text responses; framework / thematic analysis
Case studies with PVIs and childminders (three rounds)	Structured interviews	Practitioners in four PVIs and four childminders	Inductive coding; thematic analysis
	Semi-structured observations	Four PVIs and four childminders	Observation framework
Discussion with the OUP delivery team (at the end of the programme)	Semi-structured interviews and progress updates	Two members of the OUP delivery team (programme developers and trainers)	Framework/thematic analysis

[a]*The online survey for practitioners included modules that were routed to childminders only. These included some of the questions from the 'leads survey', such as motivation to sign up to the pilot, concerns about implementation, barriers, and costs, to ensure the study captures this information from childminders as well as the larger PVIs.*

Theory of Change updates

IFF Research reviewed the draft Theory of Change of Early Number with Numicon, following the IDEA workshop and the first set-up meeting, and made some amendments. These were discussed at the second set-up meeting, and the revised Theory of Change is provided below in Figure 2. It sets out the main inputs/activities and outputs that need to be in place to lead to the short-term and long-term outcomes.

Figure 2: Early Number with Numicon logic model and Theory of Change



Online survey with setting leads (two rounds), and practitioners in all participating settings (three rounds)

Survey data from **practitioners** were collected at three different time points to capture how the settings progressed during the pilot: from recruitment to post-delivery. The surveys were administered:

- *before the training commenced in September 2024;*
- *at midpoint after the delivery of sessions at settings had started, around October 2024–November 2024; and*
- *post-delivery in February 2025–March 2025.*

The main objective of the first survey was to gain insights into prior knowledge and confidence in delivering similar approaches and 'business as usual' practices. The second survey probed practitioners about their experience of training and their use of the programme materials, the professional learning community, and the buddy system, in order to gauge take up and how useful these resources had been. The last survey provided further insight on delivery, the use of the ongoing support offered, and the perceived effects of the programme on practitioners' and children's outcomes. A small incentive was provided for the completion of each survey (£10 voucher).

The survey with **setting leads**[2] at the start of the academic year (September 2024–October 2024) examined the setting's motivation for joining the programme, and their understanding of the requirements so that potential barriers to recruitment and engagement could be better understood. The second survey with setting leads was administered post-delivery in March 2025. It covered the following topics: acceptability; affordability of the approach; and other aspects of the practicalities of delivery including how fidelity to the programme was facilitated and/or hindered by various factors. Surveys consisted primarily of Likert-type questions (Joshi *et al.,* 2015), with minimal open questions to reduce burden. A small incentive was provided for the completion of each survey (£10 voucher).

Table 5 below summarises the response rates achieved in each stage of the evaluation.

Table 5: Survey response rate

Respondent group	Time point	Sample size	Completed	Response rate (%)
Practitioners	Baseline	31	26	83.9%
	Midline	29	24	82.8%
	Endline	29	21	72.4%
Lead survey (PVI only)	Baseline	14	9	64.3%
	Endline	13	8	61.5%

Observation of training

To understand the engagement with the training as well as any differences across the face-to-face and remote training, two members of the IFF Research evaluation team attended all training sessions. An observation guide was used by the evaluation team to ensure targeted focus on the delivery of key programme and training elements.

Case studies

Eight settings were purposively sampled for more in-depth case studies to capture the experience of delivering the intervention across PVIs and childminders, gathering a more holistic view of the programme. Four PVIs and four childminders were selected according to specific characteristics in September 2024 before training commenced, so that the evaluation team could observe the entire journey of the setting including the initial training sessions. The sampling

[2] If the lead and the practitioner were the same person (i.e. very likely in childminder settings) they were not asked to fill in the setting lead survey, instead they were routed through the practitioner survey to cover some of the topics from the lead survey.

strategy was defined after recruitment was completed and was designed to ensure variation in key setting characteristics, for example, setting type (PVI and childminders), level of disadvantage using the Income Deprivation Affecting Children Index (IDACI, derived from the settings postcode), and geographical location. The case studies included in-depth interviews with the practitioner and wider setting staff for PVIs, setting managers, and/or other relevant practitioners and observation of practice. In-depth interviews and observations of practice were conducted at three time points: in September 2024; November 2024; and February 2025.

The **interviews** covered a range of topics, including engagement with training, support, and practicalities of delivery, including how fidelity to the programme was facilitated and/or hindered by various factors (such as time constraints, lack of resources, child absence, or children disengagement). The IFF Research evaluation team developed a tailored topic guide for each setting. In addition, during interviews, researchers referred back to the respondent's responses in previous interviews to gather evidence on progress from the individual practitioner's own personal experience. **Observations** were guided by a semi-structured protocol examining teaching practices and children's engagement. The observations also allowed us to understand implementation fidelity, noting divergencies from the approach. Practitioner feedback from the case studies is summarised and paraphrased from detailed notes taken during case study visits, as interviews were not audio-recorded, so direct quotes are not available. To compensate PVIs and childminders for the additional burden of these elements, each case study setting received a £300 payment as a thank you for their time.

Progress updates and semi-structured interviews with the OUP delivery team

The views of the OUP delivery team were captured regularly through monthly progress updates in check-in calls. The aim was to capture their views on programme delivery across the project, any challenges they faced, any changes they made to the programme during delivery, and how actual delivery compared to their expectations. In addition, a semi-structured group interview was conducted with the OUP delivery team members at the end of the programme to obtain a more detailed understanding of their overall experiences of providing training and support, as well as any barriers or enablers of programme delivery.

Data analysis

Responses from the surveys were aggregated, tabulated, and broken down according to setting type (PVIs and childminders). Any open-text responses in the survey were analysed using a general inductive approach (Thomas, 2006). The purpose for using an inductive approach was to condense open responses with the aim of deriving emerging themes. A general inductive approach was also used when assessing responses to the interviews.

The analysis approach implemented for this evaluation was rooted in the multi-phase/dynamic method (Humphrey *et al.,* 2016), which is well suited to integrating mixed-method data. The analysis blended data of all strands to generate insights covering both the breadth of all participating settings and the depth of experiences and impacts for different types of settings. The analysis was ongoing and iterative and conducted in parallel to all stages and strands of the methodology, informing the design of the topic guides in all subsequent rounds of data collection. IFF Research designed an analysis framework to incorporate the information from all strands of data collection. The framework was structured around the research questions and success indicators for each pilot research area. It was set-up to enable triangulating evidence from across the data sources and identify differences across settings and stakeholder groups.

Significance testing was not carried out on the survey data, due to the small size of the sample and due to the large variance in the number of responses from each subgroup at each time point.

Analysis limitations: It is important to note that while the findings of the evaluation are indicative of the outcomes of the pilot, they should be interpreted with caution and understood as directional rather than definitive. This is due to:

- *the small sample size of the pilot;*
- *the pilot taking place in a specific geographic location;*
- *the opt-in nature of taking part;*
- *the specific recruitment criteria; and*

- *the limitations of capturing data by observation*.

The limitations of the evaluation are discussed in further detail in the 'Limitations' subsection under the 'Conclusion' section below.

Timeline

Table 6: Timeline

Date	Activity
January 2024	IDEA workshop
January 2024 – February 2024	Evidence review: Review of Theory of Change
February 2024	Set-up meeting
May 2024 – June 2024	Ethics, safeguarding, and data protection protocols drafting
July 2024	Evaluation protocol sign-off
June 2024 – September 2024	Participants recruitment
August 2024	Development of data collection tools
September 2024	Baseline survey for setting leads and practitioners
September 2024	Baseline case study visits
September 2024	Observation of first training session
November 2024	Observation of second training session
November 2024 – December 2024	Midline survey for practitioners
November 2024 – December 2024	Midline case study visits
December 2024	Observation of third training session
February 2025	Observation of fourth training session
February 2024 – March 2025	Endline case study visits
February 2024 – March 2025	Endline survey for settings leads and practitioners
March 2025	Interview with the OUP delivery team
April 2025 – May 2025	Analysis
May 2025	Learning workshop
June 2025 – December 2025	Reporting

Findings

Participants

Overall, 29 of the 31 recruited settings remaining engaged with the programme through to the end in some capacity. As shown in Table 7, there were slightly more childminder settings (16) than PVI settings (13) who completed the programme. The settings had a spread of locations as shown in Table 7, although more than half of settings were located in Bradford. There was a fairly even distribution across different levels of deprivation, according to the IDACI levels.

Table 7: Demographic profile of settings that took part in the pilot

Settings type	Number of settings	
	Participated	Completed
Childminder	17	16
PVI	14	13
Location		
Bradford	18	16
Halifax	1	2
Huddersfield	2	2
Leeds	5	5
Manchester	2	1
Oldham	3	3
IDACI decile		
Most deprived (1–3)	10	9
Mid-range deprived (4–6)	12	11
Least deprived (7–10)	9	9

Eight of the 29 settings were selected for a case study. This included an even split between childminders (four) and PVI settings (four). In terms of location, three settings were based in Bradford, two in Huddersfield, and one each in Oldham, Manchester, and Leeds. The case study sample also reflected a range of deprivation levels, with three settings in the most deprived deciles (1–3), two in the mid-range deprived deciles (4–6), and three in the least deprived deciles (7–10).

In terms of staff qualifications, the most common qualification among practitioners was Level 6 or higher, with almost half of practitioners (12 out of 26) in the baseline survey reporting holding this level of qualification. There was a fairly even split among practitioners in PVIs (five out of ten) and childminders (7 out of 16) holding a Level 6 or qualification or higher. This indicates that the majority of practitioners participating in the pilot were highly qualified, with degree-level or equivalent qualifications. A smaller proportion (3 out of 26) of practitioners reported holding a Level 5 qualification—with two (out of ten) practitioners in PVIs and 1 (out of 16) childminders doing so. About a third (9 out of 26) of all practitioners reported holding a Level 3 qualification, again with a similar split across practitioners in PVIs (three out of ten) and childminders (6 out of 16). Only two childminders reported holding a Level 1 or 2 qualification. Overall, these findings suggest that those taking part in the pilot were well qualified, with the majority holding at least a Level 3 qualification.

Evidence of promise - is there evidence of promise that the programme may lead to the changes expected in the ToC?

Baseline practices and approaches to early numeracy

Research question: 1.1 Is there a change in practitioners' knowledge and confidence in delivering numeracy sessions after participating in the programme? Including:

- *Are the four PD sessions perceived to be relevant in influencing the practitioners' knowledge and confidence?*
- *What new knowledge and practices do practitioners report learning?*
- *Does this vary by setting type or practitioner's prior knowledge and experience?*

At the beginning of the programme, survey and observation data found practitioners and setting leads had a varied range of prior experiences in delivering early numeracy. On a scale of 0 to 10, practitioners rated their prior experience with mathematics delivery at an average of 7.0. The majority (16 out of 26) of practitioners reported high levels of experience (scoring 7–10), suggesting a generally confident workforce. Despite this, familiarity with Numicon specifically was more limited. Although 20 (out of 26) practitioners had heard of Numicon, only 6 (out of 26) had used it before. Among practitioners already using it at the start of the programme (3 out of 26), all were childminders, indicating limited use of Numicon in PVI settings.

Among PVI leads, six (out of nine) were aware of Numicon, and two (out of nine) of settings were already implementing it.[3] This suggests that while Numicon was not entirely new to the sector, its practical application was not widespread prior to the programme.

No practitioners reported currently being involved in any other formal numeracy programmes at the time of joining Early Number with Numicon. However, some settings used other programmes in the past such as Maths Champions[4].

Many PVI practitioners in case study settings used the Early Years Foundation Stage (EYFS) framework for numeracy, which helped them use Numicon's language and concepts. For example, familiarity with EYFS terminology, such as subitising, made the transition to Numicon easier. However, one setting noted perceived gaps with the EYFS framework. For instance, one setting felt that the concept of 'shape' was less emphasised in Numicon, prompting them to adapt the programme to include their own shape-based activities, but without following a specific programme. Some settings also reported getting ideas for resources or activities online, including ideas from the social media platforms TikTok or Pinterest, or in Twinkl.

In terms of regular practice before joining Numicon, most practitioners delivered structured activities to teach mathematics. Nearly all (25 out of 26) used manipulatives, such as counters, blocks, shapes, clock dials, and number wheels in settings; resources that align closely with the tactile and visual nature of Numicon. Digital tools (11 out of 26) and paper-based resources (10 out of 26) were also common, indicating a blended approach to numeracy learning. Observation data captured during the baseline visits highlighted a wide variety of numeracy delivery methods, including both structured and unstructured activities, and the use of diverse materials across settings.

Numeracy-related professional development in the settings was also relatively limited. Most (23 out of 26) of practitioners had not participated in any mathematics-related professional development in the three years preceding Numicon with 15 (out of 16) childminders and eight (out of ten) nursery practitioners not having done so. Among the few who had (one childminder and two practitioners), the most common programmes were Maths Champions, the Early Years Professional Development Programme (EYPDP), and training linked to Early Years Initial Teacher Training (EYITT). Similarly, only three

[3] The data presented reflects survey responses to two distinct questions: one directed at practitioners and childminders; and one question specifically to PVI leads, asking about their familiarity with Numicon.
[4] Maths Champions | EEF

(out of nine) PVI leads had undertaken mathematics-specific professional development in the past two years, with Maths Champions again being the most commonly cited. However, there was evidence of growing interest with five (out of nine) leads in PVIs at the baseline and five (out of eight) at the endline planning to engage in mathematics-related professional development activities during 2024/2025.

While many practitioners were experienced in delivering early numeracy and familiar with hands-on resources, Numicon represented a relatively new and structured approach for most, particularly in settings where numeracy programmes and professional development opportunities had been limited. This was more commonly reported in childminder settings.

Practitioner knowledge and confidence

Research question: 1.1 Is there a change in practitioners' knowledge and confidence in delivering numeracy sessions after participating in the programme?

As shown in Table 8 and Table 9 below, practitioner self-report rating of their knowledge in specific items related to numeracy delivery, increased between baseline and endline, as did self-report of their confidence.[5] Practitioners were asked to rate their knowledge and confidence in supporting children's early numeracy development on a Likert scale from 'strongly disagree' to 'strongly agree' (Table 8) and to rate their knowledge on specific pieces of numeracy on a scale of 0–10, with scores of 1–3 rated as low, scores of 4–6 rated as mid, and scores of 7–10 rated as high (Table 9). For context, prior to the programme, practitioners reported a mean score of 7.0 on a scale of 0–10 regarding their previous confidence of delivering numeracy, with 16 (out of 26) indicating high levels of experience (scoring 7–10) and 10 (out of 26) mid-levels of experience (scoring 4–6). The mean score for childminders before the programme was 7.1, it was 6.7 for PVI practitioners. More detail about the prior experience of practitioners can be found in the 'Baseline practices and approaches to early numeracy' section above.

Many practitioners had high confidence (scoring 7–10) in their ability to measure, build on, assess, and set numeracy goals for children at baseline, although their confidence increased at endline for all measures:

- *set reasonable numeracy goals (21 out of 21 at endline, from 24 out of 26 at baseline);*
- *build on children's spontaneous numeracy comments/discoveries (21 out of 21 at endline, from 25 out of 26 at baseline); and*
- *decide how to assess numeracy knowledge and understanding (21 out of 21 at endline, from 23 out of 26 at baseline).*

The survey data suggested greater increases in practitioner confidence for measures related to their own knowledge and self-confidence in numeracy, compared with their ability to measure, build on, assess, and set numeracy goals for children. For example, by endline, all practitioners reported they were confident in their knowledge of what to do if a child did not understand a mathematical concept (21 out of 21, from 18 out of 26 at baseline), and how to incorporate numeracy into other activities (21 out of 21, from 22 out of 26 at baseline).

Childminders showed especially significant gains. Many reported that this was the first numeracy-specific training they had received, and the first structured programme they had implemented (numeracy or otherwise). Survey results and interviews highlighted a transformation in confidence, delivery, and professional identity. For example, while 13 out of 16 childminders reported at baseline that they could plan activities to help nursery-aged children learn numeracy, this rose to 14 out of 14 by endline. One childminder summed it up well:

[5] A minor inconsistency occurred in the midline practitioner survey. Although the question instructed respondents to rate on a scale from 0 to 10, the response options available ranged only from 1 to 10. As a result, '0' was not selectable at midline, despite being implied by the question wording. This error did not happen at baseline or endline where 0 to 10 scales were correctly implemented. In practice, the score of '1' was rarely selected (only once for one item), and no respondents selected '0' at endline, suggesting the issue had minimal impact on findings. However, this inconsistency should be borne in mind when interpreting comparisons across time points.

I don't know what I would do if I didn't have Numicon. (Childminder, Endline)

Table 8: Practitioners' self-reported knowledge and confidence in supporting children's early numeracy development, on a scale from 'strongly disagree' to 'strongly agree'

Measure	Percentage of respondents that agreed at baseline			Percentage of respondents that agreed at endline		
	Total	Practitioners	Childminders	Total	Practitioners	Childminders
What to do if a child doesn't understand a mathematical concept	18	9	9	21	7	14
How to incorporate numeracy into other activities	22	9	13	21	7	14
Best practices for learning mathematics	20	9	11	20	7	13
Local or national numeracy standards	19	9	10	20	7	13
What children know about mathematics when they enter nursery	22	8	14	19	6	13
Set reasonable numeracy goals	24	10	14	21	7	14
Build on children's spontaneous numeracy comments/discoveries	25	10	15	21	7	14
Decide how to assess numeracy knowledge and understanding	23	10	13	21	7	14
Plan numeracy activities	23	10	13	20	6	14
Gauge what children know about nursery	23	10	13	19	6	13

Source: Practitioner baseline survey. Total practitioners (n=26): PVI practitioners (n=10), childminders (n=16). Practitioner endline survey. Base: Total practitioners (n=21): PVI practitioners (n=7), childminders (n=14). The percentage of agreement on this table includes ratings of 'agree' and 'strongly agree'.

Table 9: Practitioners' self-reported ratings in specific pieces of knowledge, on a scale of 0 to 10

Measure	Percentage of respondents that rated high at baseline			Percentage of respondents that rated high at endline		
	Total	Practitioners	Childminders	Total	Practitioners	Childminders
Ability in mathematics	15	6	9	20	7	13
Ability to teach children to count	22	9	13	19	7	12
Knowledge of how to deliver mathematics	16	6	10	18	7	11
Confidence in delivering mathematics	16	5	11	18	7	11
Confidence with mathematics	14	6	8	17	6	11
Ability to teach simple calculations	18	7	11	17	5	12

Source: Practitioner baseline survey. Base: Total practitioners (n=26): PVI practitioners (n=10), childminders (n=16). Practitioner endline survey. Base: Total practitioners (n=21): PVI practitioners (n=7), childminders (n=14).

The self-reported increases in confidence in supporting children's early numeracy development among practitioners in PVIs were echoed in quantitative data from PVI leads.[6] Following attending all four training sessions, eight out of eight PVI leads stated that their practitioner was[7]:

- *confident in delivering the Early Number with Numicon activities;*
- *confident in adapting the Early Number with Numicon activities, if required; and*
- *enjoyed delivering the Early Number with Numicon activities*.

This triangulates well with the practitioner survey and suggests the training was successful in increasing both competence and enthusiasm for numeracy delivery.

Practitioners also self-reported increased confidence in their general comfort with mathematics, from 14 (out of 26) to 17 (out of 21), and their own mathematical ability, from 15 (out of 26) to 20 (out of 21). However, survey data showed smaller gains over time in terms of how practitioners rated their ability to teach simple calculations, from 18 (out of 26) to 17 (out of 21), and their ability to teach children to count, 22 (out of 26) to 19 (out of 21). These increases, while positive, were lower relative to the improvements seen in practitioners' general comfort with mathematics and their own mathematical ability. This could be explained by the focus in the training on foundational concepts such as subitising, patterning, and comparative language. One practitioner described how the training prompted them to reflect on the over-reliance on counting as a sole measure of numeracy development and helped them to assess children's learning more holistically. Therefore, while most practitioners were already confident in their ability to teach children to count, it did not grow so much over time compared to other areas of knowledge.

Finally, both survey and observational findings suggest that the programme had a broader influence on professional practice. Many practitioners expressed a motivation to share what they had learned with colleagues, indicating a potential ripple effect beyond individual delivery, and contributing to a wider shift in early years numeracy practice within settings. Observation of practice showed that the programme encouraged practitioners to embed numeracy in various activities and to use mathematical language thoughtfully and precisely. Baseline observations suggested that some practitioners were already using numerical language with children before the training began, but this was inconsistent and sometimes inaccurate. By midline and endline, observations of practice showed that the use of numerical language had become more frequent and precise, demonstrating a clear improvement in practice over the course of the programme. Overall, practitioners felt better equipped to deliver numeracy sessions and were motivated to share their knowledge with colleagues, which could indicate a positive shift in their professional practice.

Relevance of training to practitioner confidence and knowledge

Research question: 1.1 Is there a change in practitioners' knowledge and confidence in delivering numeracy sessions after participating in the programme? Including:

- *Are the four PD sessions perceived to be relevant in influencing the practitioners' knowledge and confidence?*

Observations of the training sessions, along with quantitative data collected from practitioners, indicates that the training sessions were highly relevant in influencing the practitioners' increased knowledge and confidence throughout the course of the programme. This was observed throughout the duration of the programme and confirmed through both midline and endline data. Practitioners consistently rated the training as relevant to their understanding of Numicon and numeracy delivery.

For example, all 24 practitioners reported that the first training session enhanced their understanding of how the Early Number with Numicon programme supports early mathematics learning, which makes sense given their self-reported limited knowledge beforehand. Nearly all (22 out of 23) practitioners reported this to be true of the second training session,

[6] Childminders did not complete the PVI lead specific survey, as they were often sole providers.
[7] PVI leads were invited to complete a survey at baseline, and at endline.

and all (18) reported this about the final training session. Similarly, after attending the first training session, 23 (out of 24) practitioners reported they became familiar with some of the key teaching strategies used in the Early Number with Numicon programme (e.g. creating Numicon patterns), as a result of participating in the first training session.

All (24) practitioners were clear on the requirements in relation to setting up the delivery of Early Number with Numicon in their setting after the first two training sessions.

Table 10 highlights the high level of confidence (scoring 7–10) that the training provided practitioners regarding delivery. For example, all practitioners (24) reported after the first training session and all practitioners (18) after the final training session that the information provided to them was clear, and that trainers were able to provide useful answers to their questions. Table 10 covers practitioners who attended the training sessions and filled in the survey at each time point. For information on how many practitioners attended each training session and survey response rates, please refer to Table 13 and Table 5, respectively.

Table 10: Extent to which practitioners agreed about the usefulness of the training session, on a scale from 'strongly disagree' to 'strongly agree'

Measure	First training session (rating of agree)			Final training session (rating of agree)			Percentage point change
	Total	Practitioners	Childminders	Total	Practitioners	Childminders	
The training session has enhanced my understanding of how the Early Number with Numicon programme supports early mathematics learning	24	8	16	18	5	13	0%
I am familiar with some of the key teaching strategies used in the Early Number with Numicon programme (e.g. creating Numicon patterns, etc.) as a result of participating in the training session	23	7	16	18	5	13	4%
The training session was useful in preparing me to deliver the Early Number with Numicon programme effectively	23	7	16	18	5	13	4%
Following the training sessions, how clear are the requirements in relation to Early Number with Numicon delivery?	24	8	16	18	5	13	0%

Measure	First training session (rating of agree)			Final training session (rating of agree)			Percentage point change
	Total	Practitioners	Childminders	Total	Practitioners	Childminders	
The information provided during the training session was useful for delivering the Early Number with Numicon sessions	24	8	16	18	5	13	0%
I could ask questions specific to my setting	23	7	16	17	4	13	-2%
Trainers provided useful answers to my questions	24	8	16	18	5	13	0%

Source: Practitioner midline survey. Base: Total practitioners who attended the first training session (n=24): PVI practitioners (n=8), childminders (n=16). Practitioner endline survey. Base: Total practitioners who attended the final training session (n=18): PVI practitioners (n=5), childminders (n=13). The percentage of agreement on this table includes ratings of 'agree' and 'strongly agree'.

Alongside relevance, there was high praise for the training provided from the practitioners, in terms of the style of delivery, and the content covered. Many practitioners acquired new concepts and materials and felt equipped to return to their setting to deliver Numicon activities. This was true from the first training session, where 23 (out of 24) practitioners agreed that the training session was useful in preparing them to deliver the programme effectively. This suggests high-quality training from the start, and a high level of buy-in, as to the value of the training for practitioners. The same was true when practitioners considered the final training session, all (18) agreed with the same metric.

> *The science around the Cuisenaire was great, the way they taught it was so clever. It was an interactive session, there was a lot of playing around with it [the Cuisenaire].* (PVI, Endline)

As reported, many practitioners were already quite confident and knowledgeable about numeracy delivery. However, the training was still seen as useful in cementing the pedagogical principles and concepts. Observations of the training sessions suggest a growth in confidence as the programme moved on, with more conversation and sharing between practitioners as each session went on.

The buddy system and the professional learning community

Research question: 1.1 Is there a change in practitioners' knowledge and confidence in delivering numeracy sessions after participating in the programme? Including:

- *Is the professional learning community perceived to be useful?*
- *Is the buddy system perceived to be useful?*

The programme included an informal buddy system, in which the trainers invited practitioners in the first training session to share contact information with other local practitioners. The intention was to encourage mutual support and the exchange of best practice and ideas throughout delivery. Following the first two training sessions, the majority of practitioners 21 (out of 24) were aware of the support offered by buddies. Around half of the practitioners who were aware of the support (10 out of 21) reported engaging with it. All of those who accessed buddy support were childminders (ten), while no participants from PVI settings utilised it, despite most being aware of it. It is worth noting that those who did not engage with the buddy system did not express negative views about it; rather, their non-participation appeared to be due to time constraints or a lack of perceived need for additional peer support. Observations of the training session highlight that this support was mentioned quite informally (often at the end of the sessions), which may have contributed to the low uptake.

Among those who used the buddy system, eight (out of ten) found it very useful, suggesting where it was accessed it had clear value in supporting implementation. The buddy support mostly took place via WhatsApp groups, where some

childminders shared ideas and reassured one another about progress made (where there were concerns that progress was slower than anticipated), particularly ahead of the training sessions.

While the buddy system was not utilised by all practitioners, there was a lot of conversation in the training sessions between practitioners sharing experiences and ideas. The trainers actively encouraged this exchange by asking practitioners to email photos and descriptions of their activities, which were then shared with the wider group as examples of best practice. This peer interaction helped practitioners reflect on their approaches, adapt activities more confidently, and gain reassurance from others with similar experiences. For childminders in particular, who often work alone, the opportunity to speak with others in similar roles was especially valued. One childminder noted that they liked speaking with other childminders and sharing ideas and best practice, especially when they too had experiences with younger children who might be less engaged.

Impact on children's engagement with numeracy

Research question: 1.2 What, if any impact, do practitioners feel the programme has had on increasing children's engagement and knowledge in early numbers?

Across both PVIs and childminders, practitioners consistently reported high levels of engagement and enjoyment with Numicon activities among children participating in Early Number with Numicon. Most children aged three to four in participating settings took part in at least some Numicon activities. At midline, the mean engagement score was 8.2 out of 10,[8] with 20 (out of 24) practitioners noting high engagement (scoring 7–10) and 4 (out of 24) reporting mid-engagement (scoring 4–6). By endline, the mean engagement score slightly increased to 8.3, with 20 (out of 21) reporting high engagement and 1 (out of 21) mid-engagement. Enjoyment of activities was also high, with 22 (out of 24) practitioners reporting high enjoyment at midline, and 17 (out of 21) at endline, showing a slight decrease, though children were widely described as engaged and enthusiastic to take part in the Numicon activities.

> *It's more interesting for the children. The book has given me so many great ideas about what activities to do. It's become part of our day-to-day routine. It's really helped me to teach maths.* (PVI, Endline)

Practitioners reported that most children involved in Early Number with Numicon activities learned something new and improved their understanding of numbers, with many using patterns. Practitioners in case study settings noted specific impacts, such as increasing confidence in using Numicon pieces and manipulatives throughout the programme, improvements in counting, and recognition of patterns and colours. Additionally, as the programme progressed, they observed the integration of Numicon language into everyday language within the setting. Observation data indicated that children had a good level of familiarity with Numicon resources and concepts by the end of the programme.

Practitioners often reported adapting the content in the booklet to engage children, to both keep them interested in the activity, and to ensure they are solidifying concepts. Multiple practitioners adopted the approach of 'the messier the better', in terms of creating sensory activity that is fun and exciting for children. For example, involving paint in activities. Children were observed clearly enjoying taking part in Numicon activities.

> *Anything that's messy the children really engage with and enjoy.* (PVI, Midline)

A few practitioners noted that it could be more difficult to engage boys in the activities, which reflected the practitioners broader experience of engaging boys in more structured activities in settings.

[8] Practitioners were asked: 'On a scale of 0–10, where 0 is very low and 10 is very high, how would you rate…' different statements, and a mean score was produced. A high score is a rating of 7–10, a medium score is a rating of 4–6, and a low score is a rating of 0–3.

Contribution of small group activities

Research question: 1.2 What, if any impact, do practitioners feel the programme has had on increasing children's engagement and knowledge in early numbers? Including:

- *What is the contribution of the small group activities?*
- *What is the contribution of the additional three introductory activities?*

The small group activities appear to have had a positive contribution towards children's engagement in the programme and understanding of numeracy. Practitioners were asked whether all, most, some, or a few of their children could do something during or after a Numicon activity. Overall, the proportion of practitioners who felt that all children were able to take something away from activities grew over time, as shown in Table 11 below. At midline, 13 (out of 24) practitioners reported that all children in activity groups learned something new, rising to 18 (out of 21) at endline. Similarly, there was an increase in the proportion of practitioners reporting that all children improved their understanding of numbers, from 10 (out of 24) at midline to 18 (out of 21) at endline. This suggests that not only did the children benefit from the programme, but that practitioners likely improved their delivery of the activities over time.

Table 11: Percentage of practitioners that reported on outcomes for all participating children during or after a Numicon activity, at midline and endline

Measure	Midline			Endline			Percentage point change
	Total	Practitioners	Childminders	Total	Practitioners	Childminders	
Children who have been part of an activity group have learned something new	13	3	10	18	6	12	32%
Children who have been part of an activity group have improved their understanding of numbers	10	3	7	18	5	13	44%
Children recognise objects as big or small	18	7	11	17	7	10	6%
Children notice colour in the environment	16	5	11	16	7	9	9%
Children can recognise the physical attributes of objects as the same or not the same	11	2	9	13	5	8	16%
Children notice patterns in the environment	9	2	7	12	4	8	19%
Children recognise Numicon patterns	4	1	3	12	4	8	40%
Children create Numicon patterns using objects	4	1	3	10	4	6	31%

Children use patterns to see how many objects without counting	3	–	3	10	4	6	35%
Children know that objects within a collection can be individually different and still make the same pattern	3	–	3	10	4	6	35%

Source: Practitioner midline survey. Base: Total practitioners (n=24): PVI practitioners (n=8), childminders (n=16). Practitioner endline survey. Base: Total practitioners (n=21): PVI practitioners (n=7), childminders (n=14). This table shows the proportion of practitioners who felt that 'all' children had achieved each outcome.

Recognition of Numicon patterns also grew, with 4 (out of 24) practitioners reporting that all children could recognise patterns at midline, increasing to 12 (out of 21) at endline. By endline, all (21) practitioners noted that at least a few children recognised Numicon patterns. Finally, at midline, 3 (out of 24) practitioners felt all children in their setting could use patterns to determine quantities without counting, which grew to 10 (out of 21) at endline, with all (21) confident that at least a few children could do this.

As further explored in 'Delivery of required activities' section below, not all settings reached the end of the programme within the recommended timeline of the programme, so were yet to deliver all of the small group activities. The barriers to progressing at pace included timetabling, the range of need within a group of children, and having to involve younger children (mostly in childminder settings).

Observation data showed a change over time in practitioner confidence with Numicon, but also children's recognition of Numicon pieces and confidence within the programme. It was clear through observation in settings that children were well aware of the Numicon pieces and could recognise their values, shape, and colour. Some settings had less previous experience with delivering structured small group activities (especially childminders) and did not have the most appropriate physical environment to deliver them (e.g. very small spaces). For a few settings, getting the children's engagement in this new style of activity was difficult at first, however by endline, this felt natural for most practitioners we observed. Practitioners were keeping the session short when they sensed the children were getting distracted. They recommended this as good practice with delivering the programme to this young age going forward.

In the pre-pilot stage, the findings suggested that an additional three introductory activities at the start of the programme would be beneficial. These were introduced in the pilot, under Topic 1 'Getting started'. This included learning sequence 1 (big and small), learning sequence 2 (colour), and learning sequence 3 (the same and not the same). The contribution of these activities were beneficial as it allowed for practitioners with children of a lower numeracy ability to introduce concepts at a slower pace. Many practitioners we observed were spending a lot of time on solidifying these Topic 1 learning sequences before moving on to Topic 2, suggesting the relevance of these additional activities.

Contribution of continuous provision activities

Research question: 1.2 What, if any impact, do practitioners feel the programme has had on increasing children's engagement and knowledge in early numbers? Including:

- *What is the contribution of the continuous provision activities?*

While the available data does not provide a full picture of the specific impact that continuous provision had on children's learning, it suggests that the integration of Numicon into continuous provision was widespread. However, variation in how and how often it was used indicates that its impact likely varied across settings, and further research would be needed to explore this in more depth.

Practitioners demonstrated strong uptake of continuous provision activities using Numicon. At midline, 22 (out of 24) practitioners reported implementing Numicon into continuous provision, with seven (out of eight) practitioners in PVIs and 15 (out of 16) childminders doing so. This rose to all practitioners (21) at endline, indicating universal adoption. Despite widespread use, the frequency of delivering continuous provision with Numicon remained relatively stable. At midline, 6 (out of 22) practitioners incorporated Numicon via continuous provision daily, 9 (out of 22) did so a few times a week, and 7 (out of 22) once a week. By the endline, these figures shifted slightly to 5 (out of 21) daily, 11 (out of 21) a few times a week, and 4 (out of 21) once a week. These figures show a slight increase in mid-range use, with fewer using it only weekly and the majority using it more frequently (daily or a few times a week). Satisfaction levels among practitioners regarding the frequency of Numicon use in continuous provision were mixed; four (out of seven) practitioners in PVIs and 9 (out of 14) childminders were content with their current frequency, while three (out of seven) practitioners in PVIs and 5 (out of 14) childminders expressed a desire to increase their use of Numicon in continuous provision. This feedback highlights both the successful integration of Numicon and the potential for further enhancement in its application.

We observed a variety of ways in which practitioners embedded Numicon into continuous provision. In some settings Numicon pieces were always 'out' and available for children to access. This included taking the big foam shapes outside in play time or having Numicon pieces out to play with in the sand or put in water. Practitioners attempted to use mathematical language during play supporting children's informal learning. However, some practitioners noted that when Numicon was used without adult support, children often engaged with it as a toy just 'for play' rather than as a numeracy resource.

In other settings, particularly where there were safety concerns or fears about losing small pieces, Numicon manipulatives were only available during small group activities rather than being included in free-play areas. These variations suggest that setting-specific factors such as physical environment, staffing, and profile of children (e.g. the age of the children), shaped how continuous provision was implemented.

Practitioners ensured there was one or two additional opportunities in the wider environment linked to the taught activities each week was a fidelity measure for the pilot. Based on the evidence, this measure was rated as 'high'. More information regarding the fidelity measures can be found in Table 13.

Programme impact on children from disadvantaged backgrounds

Research question: 1.2 What, if any impact, do practitioners feel the programme has had on increasing children's engagement and knowledge in early numbers? Including:

- *What impact do practitioners feel the programme is having on children from disadvantaged backgrounds especially for children with EYPP, EAL and SEND?*

Practitioners felt Early Number with Numicon was suitable and appropriate for delivery to most children in their setting (including those with EAL and SEND). Across the participating settings, leads of PVIs (of which there were nine) estimated on average 22% of children in their settings were identified or suspected to have SEND. In childminders, of which there were 16, the average was 1%. Similarly, on average PVI leads estimated 26% of children were entitled to the EYPP. Childminders estimated 1% of children on average were entitled to EYPP, highlighting the diverse needs within these educational settings.

Engagement among children with EAL remained steady throughout the programme. The mean engagement score (as reported by practitioners) increased from 7.3 to 7.6 on a scale of 0–10, between midline and endline.[9] Initially, 4 (out of 24) practitioners reported high engagement (scoring 7–10) among EAL children, 3 (out of 24) observed mid-level engagement (scoring 4–6), and 17 (out of 24) found this measure not applicable to their setting (likely due to not having EAL children in the setting). By the endline, high engagement increased to 5 (out of 21), mid-engagement decreased to 2 (out of 21) practitioners, and non-applicable responses reduced to 13 (out of 21), indicating a slight increase in engagement where EAL were present.

[9] Practitioners were asked: 'On a scale of 0–10, where 0 is very low and 10 is very high, how would you rate…' different statements, and a mean score was produced. A high score is a rating of 7–10, a medium score is a rating of 4–6, and a low score is a rating of 0–3.

Engagement among children with SEND also improved slightly over time. At midline, the mean engagement score was 6.6. Practitioners reported high engagement in 7 (out of 24) settings, mid-engagement in 4 (out of 24) of settings (scoring 4–6), low engagement in 2 (out of 24) settings (scoring 0–3), and not applicable in 11 (out of 24). By the programme's end, the mean score rose to 7.3, with high engagement reported by 8 (out of 21) practitioners (scoring 7–10), mid-engagement by 2 (out of 21), low engagement by 1 (out of 21), and non-applicability by 9 (out of 21). The high proportion of not applicable responses particularly among childminders, reflect that some settings did not have children with these characteristics, or did not know whether they had.

The survey asked practitioners to report on the prevalence of EYPP children in their setting but did not include further questions about their engagement or outcomes. Practitioners did not tend to speak explicitly about EYPP children as a distinct group. EYPP status was not commonly mentioned by practitioners, and they did not make explicit links between EYPP and programme impact.

The 'Barriers and enablers for children from disadvantaged backgrounds' section below covers the experiences of SEND and EAL children in more detail.

Changes to practitioner's approach and use of resources

Research question: 1.3 How different is the approach to other mathematics resources and programmes used? Including:

- *What is the usual setting practice of early numeracy approaches?*

Participation in Early Number with Numicon led to a shift in the way practitioners approached early numeracy teaching. Practitioners in case study settings reported taking a more structured approach in teaching numeracy than they had used previously. In some settings, practitioners who had not planned specific numeracy activities prior to Numicon began preparing and documenting activity plans each week, as encouraged by the training.

Another observed change in case study settings was in the use of language. Practitioners were making more deliberate use of mathematical vocabulary in daily interactions with children. Examples included consistently using terms such as 'bigger' and 'smaller' and encouraging number recognition using terms like 'find me 4' rather than assuming children know the numeral. Practitioners attributed this shift to learning in the Numicon training, which emphasised the importance of language use in supporting numeracy development.

Use of resources

Research question: 1.4 Did practitioners' early numeracy approach change (and if so, how) as a result of the programme? Including:

- *Did practitioners use the resources?*
- *Did this vary by setting type?*

In terms of resources, by midline, practitioners had begun integrating a range of Numicon materials into their practice. The most widely adopted resource were the Numicon manipulatives, used by all settings, and there was a strong uptake of the pegs (20 out of 24), baseboards (17 out of 24), and large foam shapes (15 out of 24), as shown in Table 12 below. However, more specialised resources such as the pan balance and Cuisenaire rods saw limited uptake at this stage, with only (4 out of 24) and one setting, respectively reporting their use.

Table 12: Resources used by practitioners

Resources used	Midline			Endline		
	Total	Practitioners	Childminders	Total	Practitioners	Childminders
Box of Numicon shapes	24	8	16	21	7	14
Baseboard	17	6	11	21	7	14
Number overlays	9	2	7	16	5	11
Picture overlays	6	1	5	16	4	12
Pegs	20	7	13	20	7	13
Zig-Zag book	10	3	7	18	7	11
Cuisenaire (small box)	1	1	–	9	2	7
Big foam shapes	15	5	10	20	7	13
Counters	8	2	6	15	5	10
Pan balance	4	–	4	19	6	13

Source: Practitioner midline survey. Base: Total practitioners (n=24): PVI practitioners (n=8), childminders (n=16). Practitioner endline survey. Base: Total practitioners (n=21): PVI practitioners (n=7), childminders (n=14).

By endline, after all training sessions had been completed and settings had the opportunity to work through the full sequence of learning materials, the use of Numicon resources had become more widespread and diverse. The manipulatives remained universally used, and all settings had now incorporated the baseboards. Notably, the use of the pan balance saw a significant increase—from just 4 (out of 24) at midline to 19 (out of 21) at endline—indicating growing confidence and familiarity with a broader range of tools. The Cuisenaire rods, however, remained the least used resource, adopted by only 9 (out of 21) settings. This included 7 (out of 14) childminders but just two (out of seven) practitioners in PVI settings. Childminders also made greater use of the picture overlays by the endline. Although all settings were provided with all of the listed resources (in Table 12 above) at the beginning of the programme, they were trained to use them in a staggered fashion across the four training sessions. For example, practitioners were shown how to use the pan balance during the third training session, and the Cuisenaire during the final training session. This may explain the lack of use of certain resources at midline.

In terms of perceived usefulness, feedback was consistently positive. At midline, all practitioners who had used each resource reported finding them useful. This included unanimous positive feedback for the Cuisenaire rods (despite low usage), the big foam shapes, and the Zig-Zag book. The pan balance, while less used at midline, was also rated positively by those who had tried it. By endline, this trend continued, with all resources being rated as useful by those who used them. The Numicon manipulatives stood out as the most valued, with 19 (out of 21) settings describing them as 'extremely useful'. PVI practitioners were more likely to rate the pan balance, Zig-Zag book, and picture overlays as 'extremely useful', suggesting variation in how different settings integrated and valued specific resources.

In the case studies, Cuisenaire rods were regarded less positively by settings. A few practitioners and setting leads felt that they were too complex for the children in the setting, or that they would easily be lost, contributing to lower uptake. Additionally, some regarded the Cuisenaire pieces as potentially hazardous for younger children, so would not use them with a younger age group. Another resource that was reported as less useful was the seaside book. Some practitioners felt the setting of the seaside was not relatable to the children's experiences either because of the geographic location of the setting or that some children may not have experience of visiting the seaside due to family working patterns or lack of funds. This meant some children were less familiar with certain terms related to this resource (such as 'lighthouse'), which made it difficult for them to engage with the resource. While these comments are sometimes related to disadvantaged children,

it should be noted that specific information on children's backgrounds was not collected, so firm conclusions about the prevalence of disadvantage among children cannot be drawn.

Other negative feedback on resources included the distracting nature of certain design elements. For example, the colours of the Numicon manipulatives sometimes confused children (e.g. calling the same piece blue or purple). Additionally, the pegs, which needed to be attached to the baseboard in a specific way were considered too fiddly and occasionally disrupted focus during activities.

Feasibility of implementation - is Early Number with Numicon feasible and acceptable to PVIs and childminders?

Engagement with training sessions

Research question: 2.1 What is the engagement with the training sessions? Including:

- *What are the completion rates for the training?*
- *What are practitioners' views on the number of PD sessions, duration and hybrid training model?*

Delivery of the training and activities was spread across 14 weeks between September 2024 and February 2025. The first training session lasted three hours and was face-to-face. The middle sessions (sessions 2 and 3) were online twilight sessions that were one and a half hours long. The final session was face-to-face and three hours long. Activities for children took place at least once a week with each activity lasting 15 minutes.

Before the programme began, childminders and PVI leads were confident in their ability to participate in all professional development sessions and to deliver the necessary programme activities.[10] Most practitioners attended all training sessions, as shown in Table 13 below. Absences were rare and this was due to unforeseen circumstances (e.g. sickness) and only accounted for a handful of practitioners for each session.

Around half (6 out of 13) of PVI leads attended the first training session, with slightly fewer attending subsequent sessions. While attendance was encouraged for PVI leads, it was not mandatory.

The training was delivered across four sessions, in a blended format:

1. In person (weekend), held in Bradford.
2. Online (weekday evening).
3. Online (weekday evening).
4. In person (weekend), held in Bradford.

[10] For PVI leads, this was confidence in their staff's ability to attend.

Table 13: Training attendance

Training session	Required attendees	Number of attendees	Number of leads in attendance	Number of practitioners in attendance (including childminders)	Number of childminders in attendance	Number of staff from PVIs in attendance (leads or practitioners)
1	Practitioners	38[a]	6	32	18	20
2	Practitioners[b]	30	2	28	18	12
3	Practitioners	32	2	30	17	15
4	Practitioners	33	4	29	17	16

Source: Attendance data from the OUP.

[a]This included the setting lead from a setting that dropped out after the first training session. A practitioner did not attend with them.

[b]N.B. some PVI leads did also attend the practitioner only sessions.

Practitioners attending all four training sessions was a fidelity measure for the pilot (with a threshold for high fidelity being if 100% of practitioners attend all training sessions either during scheduled delivery or afterwards). On the basis of the attendance data, the fidelity measure was rated as 'medium to high'. More information regarding the fidelity measures can be found in Table 13 above.

As a result of high attendance at the training, practitioners were well-prepared to deliver the Numicon programme. Practitioners were positive about the training sessions, and reported they gained valuable information, resources, and ideas for implementing the Early Number with Numicon programme. For example, at endline, of those that attended the final training sessions, all (18) practitioners agreed that the training session was useful for delivering the programme and enhanced their understanding of how the Early Number with Numicon programme supports early mathematics learning.

While most practitioners appreciated the flexibility of online sessions, some practitioners found online sessions less engaging than in-person ones. Others noted that in-person sessions were more time-consuming due to travel and weekend commitments. There was value in the in-person sessions, which were lively and involved many practitioners asking questions and sharing practice with other practitioners. For online sessions, observations suggested that a few practitioners joined the session through phones or tablets, which might have been more challenging than using laptops. However, there was no indication that they were not engaged.

> *They weren't too long—sometimes when you've worked all day it can be tiring. But it didn't feel like the online ones were too long.* (PVI, Endline)

The mixed approach of online and in-person training was generally well-received, taking into account the benefits and shortcomings of both online and in-person approaches. There was however, a slight preference for in-person sessions. At the midline survey, practitioners who had attended both an in person and online training were asked, which they preferred. Just over half (12 out of 23) preferred in-person attendance, 2 (out of 23) online sessions, and a sizeable 9 (out of 23) had no preference. At the endline survey, most practitioners who attended both formats preferred in-person sessions (15 out of 21), while 5 (out of 21) had no preference, and 1 (out of 21) favoured online sessions. It should be noted that practitioners were not asked how they would feel if all training sessions were in person.

Engagement with the programme activities

Research question: 2.2 What is the engagement with the programme activities (i.e. small group activities and continuous provision for children)? Including:

- *Did practitioners deliver the required activities?*
- *What are practitioners' views on the programme's dosage and content?*
- *What adaptations if any, were made?*

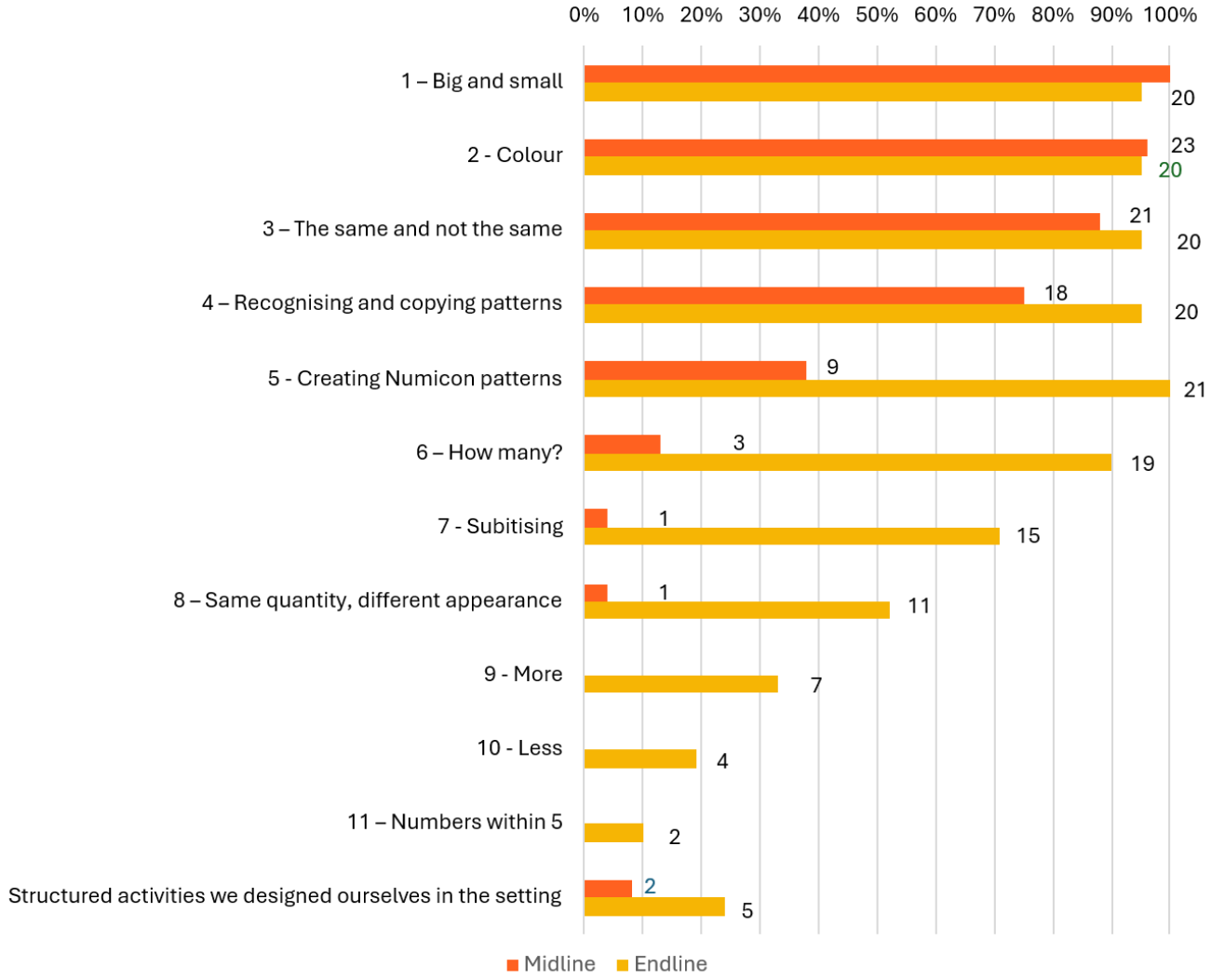
Motivation to take part in the programme

Reasons for engagement in the programme were explored through understanding of motivations for programme sign up. The primary motivation for signing up for the programme among PVI leads was to improve numeracy skills among children and to seek more mathematics/numeracy-related professional development for their staff, with all (nine) indicating these as key reasons. Interest in Numicon specifically was a less common motivator, with four (out of nine) citing its reputation and a desire for more professional development in using Numicon. Among childminders, the main reason was also to enhance children's numeracy skills all (16), followed by improving their confidence in teaching numeracy (13 out of 16). The least common reason among childminders was addressing specific needs in children's mathematics/numeracy development (2 out of 16), and interest in reasons specifically related to the Numicon product/brand were also low.

Delivery of required activities

Progress through the programme activities was tracked via the midline and endline practitioner surveys. At the time of responding to the midline survey, nearly all (21 out of 24) practitioners had completed up to learning sequence 3 with their children, with 18 (out of 24) reaching sequence 4, 9 (out of 24) sequence 5, and smaller percentages completing sequences 6 through sequence 8 (as shown in Figure 3). By endline, most practitioners had progressed to learning sequence 6, with 15 (out of 21) completing sequence 7, and 11 (out of 21) completing sequence 8. Progress beyond sequence 8 was lower, with fewer practitioners reaching sequences 9 through 11. Additionally, 2 (out of 24) practitioners had completed structured activities they had designed at midline, which increased to 5 (out of 21) by endline. Figure 3 shows this progress through the learning sequences, based on the self-report of practitioners at midline, and then again at endline. Note, the sample is small, and not all the same practitioners responded at midline and endline. This means that there can be small differences in conclusions based on responses at midline/endline (e.g. 100% [24] of practitioners reported at midline that they completed learning sequence 1, and only 95% [21] of practitioners reported this at endline).

Figure 3: Activities delivered at midline and endline



Source: Practitioner midline survey. Base: Practitioners who delivered Numicon activities in settings (n=24). Practitioner endline survey. Base: All practitioners (n=21).

These findings show that not all practitioners in all settings delivered all 11 activities by the end of the programme. The findings indicate that the delivery approach varied. Some practitioners worked through each individual activity within a learning sequence before moving ahead to the next learning sequence, others picked specific activities they felt would most engage the children, based on observed interest and readiness. This flexible approach in delivery led to variation in programme completion. For example, by the end of all training sessions, some settings had 'completed the programme', and others were in the early stages of delivery and felt delivery in full could take one to two years. The booklet was often used as a flexible guide rather than a fixed script for the activities. This was due to a lack of specific resources, practitioner preference, or children's interest and needs.

In training sessions, case study visits and the OUP visits, some practitioners shared concerns about whether they were progressing through the activities quickly enough, given the number of activities and weeks in the programme. However, many practitioners demonstrated sound professional judgement, making informed decisions to adjust, revisit, or advance in the programme activities based on the children's readiness, demonstrating responsiveness to children's needs.

> *We want to build a base before we move on to the next step.* (PVI, Endline)

One of the fidelity measures for the pilot was related to whether the teaching sequence was progressive, and children reached learning sequence 7 or beyond. The threshold for this was that practitioners follow the teaching sequence (with the exception of the familiarisation activities when the children are already familiar or they are not necessary). At a minimum, it

was expected that a child will reach and complete learning sequence 7. This did not need to hold true for children with SEND or EAL. However, during the programme, there were discussions about the importance of moving beyond learning sequence 7. The OUP delivery team emphasised during delivery that the quality of teaching and depth of children's understanding were more important than strictly reaching learning sequence 7. This approach was emphasised in training sessions and visits to reassure practitioners that adapting to children's needs was appropriate and encouraged. As a result, reaching learning sequence 7 was not viewed as essential, and relatedly, this was rated as having medium fidelity.

At the time of the midline survey, all practitioners had begun implementing Numicon activities in their settings. Most 17 (out of 24) delivered these activities a few times a week, with 3 (out of 24) delivering daily, and 4 (out of 24) once a week. By the endline, the frequency shifted to 1 (out of 21) daily, 13 (out of 21) a few times a week, and 6 (out of 21) once a week. Practitioners were split on whether they wanted to increase the frequency of delivery, with 11 (out of 21) who expressed interest in delivering activities more often and 10 (out of 21) content with the current frequency.

Numicon activities were primarily delivered to small groups, with 11 out of 24 practitioners at midline working with groups of two to three children, and this increased to 10 (out of 21) by endline. The ideal group size, according to 11 (out of 21) practitioners, was two to three, aligning with actual practice. While the programme was designed for three–four-year-olds, many practitioners included children from other age groups, particularly two-year-olds or older preschoolers, based on the setting composition. At endline, 9 (out of 21) practitioners focused solely on three–four-year-olds, while others included additional age groups. By endline, after attending the training sessions and delivering the programme, 20 (out of 21) of practitioners thought that Numicon was suitable for three–four-year-olds, as well as children of other age groups in their settings. This was true for all (14) childminders.

Practitioners delivering one taught session (10–15 minutes) to all three–four-year-olds in the setting per week was one of the fidelity measures of the pilot (with the threshold for high fidelity being that practitioners deliver one taught session to all three–four-year-olds each week). While a majority of practitioners reported delivering taught sessions frequently—with 17 (out of 24) at midline and 13 (out of 21) at endline reporting delivering a few times a week, and 3 (out of 24) at midline and 1 (out of 21) at endline delivering daily—the fidelity threshold required one taught session per week to all three–four-year-olds in the setting. In fact, delivering precisely once a week was not very commonly reported, this was the case of 4 (out of 24) settings at midline, and 6 (out of 21) settings at endline. All six of these settings were childminders at endline, whereas PVI settings at endline were all delivering at least a few times a week. As a result, while overall delivery was strong and often exceeded the frequency requirement, the programme did not consistently reach *all* three–four-year-olds every week. For this reason, fidelity to this specific measure was rated as 'medium to high' rather than 'high'.

While feedback on the content of the programme was largely positive, some aspects were viewed less favourably by practitioners. For example, a small number of practitioners noted that the beach-themed activities did not feel relevant for their children either due to geography or the season in which they were delivered.
The programme is designed in a way that means practitioners can deliver the sessions that they have been trained on, but they also have the freedom to tailor the sessions based on the needs of the children they are working with and the context of their setting. Adaptations were mainly made to link the activities to the topics of focus at the time in the setting (e.g. mini beasts, weather). The way some of these adaptations were delivered depended on the resources and environment the setting has, for example, using flour or placing the pieces in water during the activity, if sand was not available. Sometimes this was done to make activities more engaging. Childminders, in particular, were required to make adaptations to their ways of delivery, due to their unique setting. Childminders were often a single practitioner looking after a small group of children of different ages, and usually just one or two children of the target age for Numicon. As a sole practitioner, they were required to adapt the timing and frequency of the sessions in an agile way to fit in with their responsibilities to the other children in their setting. When there was one child of the target age, the group activities needed to be adapted to align with a one-on-one set-up.

Barriers and facilitators for practitioners

Research question: 2.3 What contextual factors act as barriers and facilitators for practitioners? Including:

- *Attending the PD sessions*

- *Delivering the small group activities and continuous provision*
- *Do these vary for PVIs and childminders?*

Most practitioners in both PVIs and childminder settings were able to deliver Early Number with Numicon and attend the training sessions without major barriers. However, some challenges were noted. Travel time to face-to-face training sessions, competing responsibilities, staffing shortages, and the presence of younger children, sometimes made training attendance or programme delivery more difficult. These challenges were generally manageable, and practitioners demonstrated flexibility and adaptability to maintain delivery.

Concerns about travel distance and time were raised in a few cases, particularly in relation to the weekend face-to-face sessions. There was no evidence that training cost or access to resources were a barrier. More information about training attendance can be found in the section on 'Engagement with training sessions' above, and further discussion of resources and costs is provided in the section on 'Cost and resources: feasibility and acceptability' below.

While most practitioners were able to deliver the programme, they encountered a range of barriers that sometimes affected delivery. One of the most frequently cited challenges was limited time to deliver the programme alongside other responsibilities. Half of PVIs leads (four out of eight) and 6 (out of 14) childminders identified time demands as a key challenge. Another common challenge was coordinating the participation of all children, particularly part-time attendees, which disrupted the delivery flow of the programme. These factors contributed to some settings not completing all activities, suggesting a need for more time to thoroughly cover the content. The varied schedules and limited flexibility in these settings necessitate adjustments in programme delivery. Extending the programme duration or offering flexible scheduling options could help ensure that all children benefit from the Early Number with Numicon programme. This approach would reduce the time-related pressures on practitioners, allowing them to provide high-quality delivery without compromising content coverage.

Practitioners also reported challenges related to the presence of younger children in their settings, which was particularly common among childminders. Balancing the needs of children outside the three–four-year age range sometimes disrupted delivery, with 8 (out of 14) childminders and four (out of eight) practitioners reporting this as a challenge. Tasks such as feeding and nappy changes sometimes diverted focus from delivery of Numicon activities, making it difficult to balance educational needs with care requirements, across a range of children. Additionally, some of the programme materials were considered unsafe for younger children. For example, smaller materials like Cuisenaire pieces and manipulatives were seen as hazardous for younger children. Practitioners found ways to manage some of these challenges by employing strategies like scheduling activities during nap times or adapting activities to suit younger children's development levels or needs where appropriate and maintaining programme flow while accommodating all children's needs. However, including two-year-olds in sessions was viewed positively. Even if younger children did not grasp all concepts, their early exposure to numeracy was seen as beneficial. Overall, while younger children's presence could complicate programme delivery, practitioners often found solutions by adjusting activities to fit schedules and developmental stages, ensuring the programme's benefits reach as many children as possible.

Some settings faced challenges related to limited or crowded/busy spaces. Observations suggested that practitioners in busy settings often faced challenges due to distractions from other children. In many nurseries, multiple activities were happening at once, often in relatively small and crowded spaces. This could make it hard for children to focus on Numicon activities, as they were distracted by other activities. While practitioners were skilled at keeping children engaged, it was not always possible, especially with child-led activities, where children sometimes chose other activities. Despite these challenges, practitioners showed resilience and adaptability, and the observation data shows them using a range of strategies to sustain children's participation even when full engagement wasn't always possible.

Staffing issues in some settings have led to some delays in delivering the Early Number with Numicon programme. In PVIs, if the only trained practitioner was off sick or had other priorities, delivery of Numicon could be paused for a few weeks. These issues were made worse by sector-wide staffing shortages. Another related factor is the need to support children with SEND, requiring practitioners to prioritise these immediate needs over the Numicon programme. This challenge was

especially pronounced for childminders who worked alone and managed all childcare aspects. Balancing these responsibilities with programme delivery could be difficult in understaffed settings.

> *We've had a few staff leave, and some staff off sick, which impacts anything else. When that happens, we need to make sure we've got the basics done, safeguarding in place, before anything extra. We shouldn't really see Numicon and maths as extra, we should see it as an everyday thing, but when you're learning and trying to embed it, it is an extra thing.* (PVI, Endline)

Barriers and enablers for children from disadvantaged backgrounds

Research question: 2.4 What are the barriers and enablers for children from disadvantaged backgrounds accessing the programme especially for children with EYPP, EAL, and SEND?

Practitioners across both PVI and childminder settings demonstrated an ability to adapt the programme effectively to support diverse groups of children. Children with SEND and EAL generally engaged well with the Early Number with Numicon activities (for more information see 'Programme impact on children from disadvantaged backgrounds' section above).

Language barriers between children and practitioners, and parents and practitioners were noted as a barrier in some settings. This occasionally resulted in a slower pace of delivery for EAL children. Practitioners highlighted that the repetitive, structured, and multi-sensory nature of the programme supported EAL learners by reinforcing key vocabulary and mathematical concepts. One setting enhanced home learning by providing parents with a vocabulary list in their native language, which was well-received.

Some children with SEND did not participate due to other developmental priorities, particularly where children needed to focus on skills like self-regulation or sitting quietly for group activities. Where possible to engage a SEND child with the activities, practitioners were able to make effective adaptations to support inclusion of children with SEND by increasing repetition, shortening activity durations, or by replacing group activities with smaller group sizes or one-on-one delivery. Despite these adaptations, some practitioners requested more structured guidance in the training materials, specifically related to SEND.

SEND and EAL emerged more organically in the interviews with practitioners more readily identifying and reflecting on barriers and enablers for these groups. In contrast, EYPP status was not commonly mentioned by practitioners. Where issues related to disadvantage were raised, these were often framed more generally without direct reference to EYPP eligibility. For example, one practitioner felt the programme was particularly effective in their setting because children came from 'good homes' and already had a strong foundation in mathematics at home. Conversely, another noted that some children in their setting were less familiar with everyday concepts (e.g. not knowing what a lighthouse was), which they linked to broader disadvantage.

Acceptability of the programme

Research question: 2.5 How acceptable is Early Number with Numicon to PVIs and childminders? Including:

- *Is the training accessible to practitioners in all settings?*
- *How is the programme received in different settings?*

The Early Number with Numicon programme received positive feedback, with all (14) of childminders stating they would recommend it to other professionals, highlighting strong practitioner endorsement. Nearly all PVI leads (seven out of eight) stated they would recommend it to other professionals. Additionally, all participating settings reported that they would continue using the programme in the next academic year reflecting its relevance, practicality, and ease of integration.

No major amendments were suggested, as settings adapted the programme using available resources to better suit their own setting. For example, some tailored the content to suit younger children, or children with different developmental levels. Others used accessible materials like flour or water in place of sand, and some omitted content (e.g. beach-themed content) that was not relevant for their setting. Half of PVIs (four out of eight) and 6 (out of 14) childminders reported making

or planning to make such adaptations. These adjustments were generally seen as practical and necessary rather than problematic, with many viewing the flexibility to adapt as a strength of the programme. While the content was largely seen as suitable, some practitioners suggested that future iterations of the programme could offer more guidance for adapting activities to younger or older age groups or children with EAL and SEND, as well as alternatives to the beach theme, which was considered less relevant for non-coastal settings.

Delivery approaches varied across settings. Practitioners welcomed the flexibility of the programme structure and resource booklet, often using it as a guide rather than a strict curriculum. This adaptive use allowed them to align delivery with children's readiness and interest levels, rather than progressing linearly booklet's structure. Instead, practitioners used their professional judgement to prioritise depth of understanding over speed of delivery, a shift reinforced by the programme's delivery team during training and support visits. More information about this can be found in the section on 'Engagement with the programme activities' above.

The training component, which was delivered in a blended format of online and in-person sessions was also well-received. The hybrid format helped balance flexibility with peer learning and was acceptable to most. As discussed in the section on 'Engagement with training sessions' above, while preferences leaned towards in-person delivery, some noted that in-person sessions were more time-consuming due to travel and weekend commitments. Practitioners were not asked how they would feel about all sessions being in person, so it is unclear if full in-person delivery would be more acceptable.

In conclusion, both the programme and its associated training were acceptable to practitioners across all settings. Minor refinements, such as offering more adaptation guidance or expanding thematic content, may enhance the acceptability further in future iterations.

Fidelity measures

The pilot had four fidelity measures, which assessed how feasible the implementation of the pilot was. Table 14 summarises the findings across each of the four fidelity measures and gives a rating of fidelity from low to high. Further discussion of how particular measures could be adapted for scale-up, is included in the 'Interpretation' and 'Future evaluation work' subsections under the 'Conclusion' section below.

Table 14: Fidelity measures high-level summary

Fidelity measure	Threshold	Summary of evidence	Rating
1. Practitioner(s) in the setting attend all four training sessions	100% of practitioners attend all training sessions either during scheduled delivery or afterwards	Nearly all practitioners (and leads where appropriate) attended all four training sessions. Where training was missed, there was often good reason for this (e.g. illness). Only one setting dropped out once the training had begun. There is no clear evidence on whether 'catch-up' happened for those that missed the sessions (no-one who completed the surveys missed the training)	Medium to high fidelity
2. Practitioner(s) deliver one taught session (10–15 minutes) to all three–four-year-olds in the setting per week	Practitioners deliver one taught session to all three–four-year-olds each week	Practitioners were most commonly delivering taught sessions a few times a week. It was not always possible for all three–four-year-olds to have received every activity (due to schedules, their ability/interest, SEND/EAL). In many settings, children aged below the age of three were	Medium to high fidelity

Fidelity measure	Threshold	Summary of evidence	Rating
		receiving a taught session (often childminders)	
3. Practitioner(s) ensure there are one or two additional opportunities in the wider environment linked to the taught activities each week	Practitioners ensure there is at least one additional opportunity in nine or more out of the 11 weeks	At the midline—92% of practitioners had used the Numicon resources more widely in the setting, growing to 100% at endline. All (100%) practitioners at the midline reported that they delivered some play or continuous provision with Numicon at least once a week, this was 95% at endline	High fidelity
4. The teaching sequence is progressive and children reach activity sequence 7 or beyond	Practitioners follow the teaching sequence with the exception of the familiarisation activities when the children are already familiar, or they are not necessary. At a minimum it is expected that a child will reach and complete activity sequence 7. This does not need to be applied for children with SEND or EAL	By endline, 71% of settings delivered activity sequence 7. All practitioners were looking to continue Numicon after the pilot and anecdotally shared they were going to continue working through the activities	Medium fidelity

Readiness for trial -how feasible is to scale delivery (and undertake evaluation using a randomised controlled trial design) with this population (i.e. PVIs and childminders)?

Recruitment and retention of practitioners

Research question: 3.1 To what extent does the recruitment strategy recruit and retain the settings and practitioners? Including:

- *What are the different challenges depending on the type of setting?*

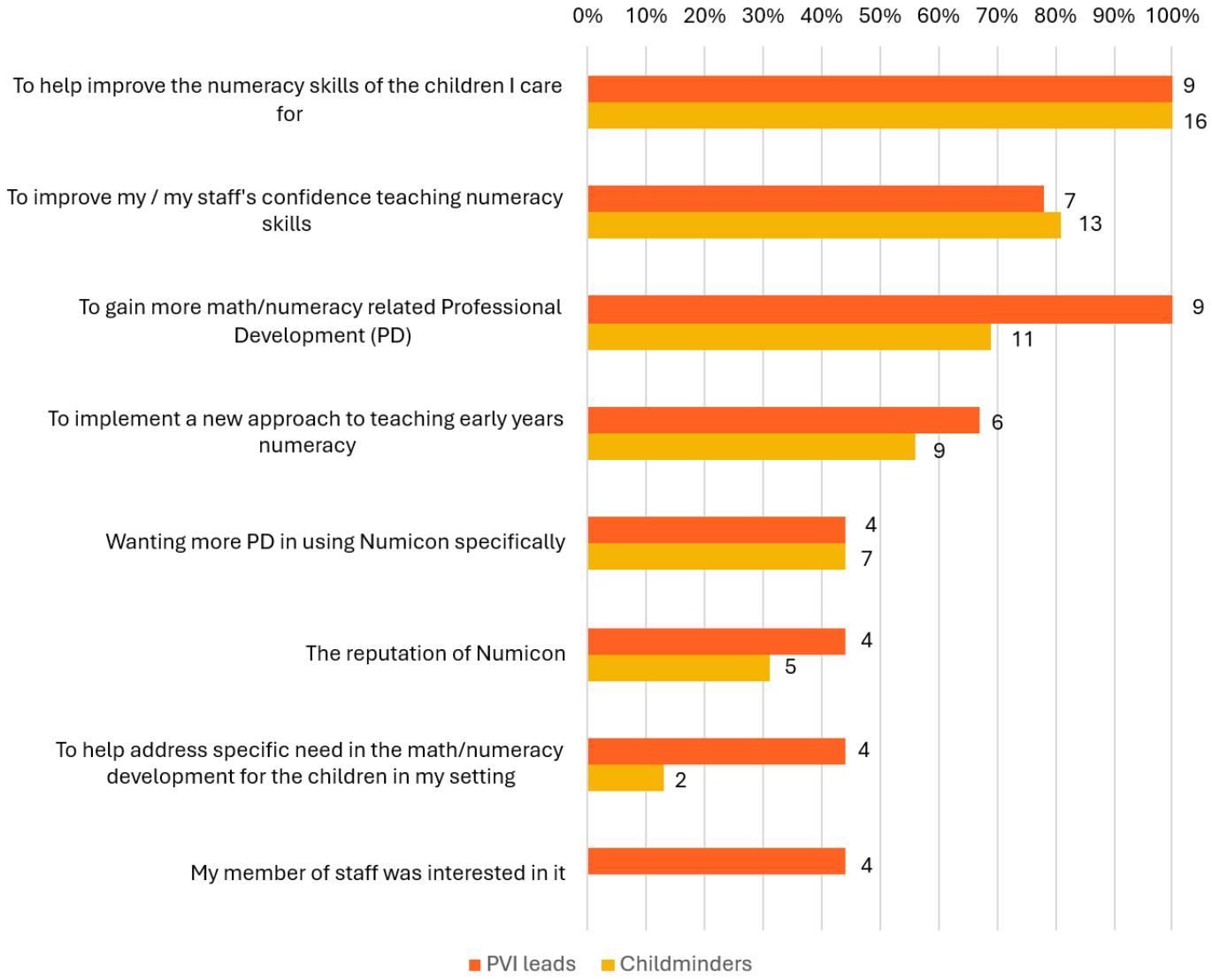
The pilot aimed to recruit 30 settings, with an approximately equal split between PVIs and childminders. This sample size, determined jointly by the evaluation team, the OUP, and the EEF, was deemed sufficient for the pilot, allowing for a minimal level of attrition. Recruitment responsibilities were led by the OUP delivery team. Engagement with Early Number with Numicon was generally strong across both PVIs and childminders, with low overall attrition. A total of 34 settings initially signed the MoU.

However, three settings were later deemed ineligible based on the recruitment criteria, as they were identified as maintained nurseries and therefore, not eligible for inclusion in the pilot. Two additional settings dropped out: one prior to the first training session; and another between the first and second training sessions. Ultimately, 29 settings remained engaged with the programme through to the end in some capacity.

Recruitment feasibility was confirmed through survey data. All PVI leads (nine) indicated that they had no concerns about joining the programme, and the majority of childminders (14 out of 16) expressed similar confidence. A small number of childminders (2 out of 16) reported uncertainty, rather than specific concerns. These findings align with the delivery team's experience, who noted that recruiting PVIs was more straightforward than recruiting childminders. While PVIs often had an established online presence and connections to other settings, which facilitated more efficient outreach, childminder recruitment was more resource-intensive, requiring individual contact sourcing due to the lack of a centralised database. In some cases, childminders also expressed initial uncertainty about the relevance of the programme or low confidence in engaging with professional development opportunities, particularly compared to PVI settings.

In terms of motivation, survey data further highlighted the strong alignment between participant goals and programme aims as shown in Figure 4 below. Among PVIs, all leads (nine) cited a desire to improve numeracy skills among children and access professional development for staff as key drivers for taking part. Additionally, two-thirds (six out of nine) identified interest in trying a new approach to teaching mathematics and numeracy as a further motivating factor. Among childminders, similar motivations were observed: all respondents (16) reported that they joined the programme to help improve the numeracy skills of the children in their care. A large majority (13 out of 16) also sought to improve their confidence in teaching numeracy, and 11 (out of 16) childminders reported an interest in gaining access to more mathematics- and numeracy-related professional development.

Figure 4: Motivations for signing up to the Early Number with Numicon programme



Source: Leads baseline survey. Base: Childminders (n= 16). Base: PVI leads (n= 9).

Recruitment began slowly and was initially very challenging. However, the recruitment effort gained significant momentum when a team member was allocated to recruitment and concentrated on this until it was achieved. Their focused time and persistence proved to be a turning point in the process. This suggests that having a dedicated recruitment lead that is able to prioritise recruitment, maintain consistent follow-up, and build relationships with potential participants can be key in securing the required number of settings.

Despite the additional effort required to recruit childminders, the OUP delivery team successfully recruited the target number of settings. For a future scale-up, more robust and scalable recruitment strategies will be needed, particularly for

childminders, to reduce reliance on time-intensive outreach and local area knowledge. Nonetheless, the pilot demonstrates that recruitment and retention are feasible, and that settings are motivated by the opportunity to support children's numeracy development and enhance practitioner skills.

Feasibility of data collection

Research question: 3.2 What does the pilot tell us about the feasibility of data collection?

The OUP delivery team monitored how each setting was progressing primarily through recording training attendance and via their own post-training surveys. Attendance at training sessions was consistently tracked, and the four post-training surveys received a good response rate, providing useful insights into practitioner engagement and immediate feedback on the training content.

During the pilot, the OUP trainers conducted one visit to all but one of the participating settings, to offer support and guidance. These visits were not considered an essential part of the programme design by the OUP team but were seen as a helpful addition where resources allowed. Feedback from case study settings suggested that these visits were appreciated primarily for providing reassurance and a sense of connection, rather than contributing substantial new learning or insights into the programme content. Practitioners rarely referred to the visits unprompted, indicating that they were not a central component of their experience.

Looking ahead to a scaled-up trial, the OUP would need to consider the value and role of these visits. If such visits are to be included, there should be a clear purpose, whether to provide implementation support, collect data, or conduct quality assurance. Alternatively, monitoring activities of this kind could be delegated to an external research or evaluation partner to reduce the delivery burden.

There were no formal systems in place to robustly track how settings were progressing through the programme or to assess the quality of delivery. Informal mechanisms, such as email exchanges between the OUP and practitioners prior to training sessions and occasional discussions during training, were used to gauge progress. However, these ad hoc approaches offer limited visibility and consistency and did not allow for systematic evaluation of fidelity or implementation quality.

More broadly, a scaled-up trial would benefit from the development of formalised data collection processes that would allow for tracking where each setting is in the programme, identifying any delays or challenges, and gaining a clearer picture of implementation quality. This could include structured tools to support regular reporting, capture fidelity of delivery, and ensure consistent data collection across settings.

In conclusion, the pilot demonstrates that basic data collection processes such as tracking attendance and gathering immediate feedback through post-training surveys are in place and they are valuable. These mechanisms provide essential information about engagement and can inform ongoing improvements to training delivery. However, for a scaled-up trial, it will be important to go beyond these elements. Specifically, the OUP will need to consider how to systematically track the progress of individual settings through the programme, as the pilot provides evidence that progression can vary significantly across contexts (see ‘Delivery of required activities’ subsection above).

In addition, while setting visits offered useful reassurance and support to practitioners during the pilot, replicating this level of contact in a larger trial may not be practical. As such, it will be important to evaluate the role and purpose of site visits and explore alternative ways to provide similar forms of support (e.g. peer networks, virtual check-ins, etc).

Feasibility of a randomised controlled trial (RCT)

Research question: 3.3 How feasible is it to conduct an evaluation using an RCT design with PVIs and childminders? Including:

- *What are the benefits and disadvantages of estimating one treatment effect for PVIs and childminders combined, compared to estimating separate effects?*
- *What are the implications for sample size?*

If the evaluation is scaled to a full RCT, several elements will require careful planning to ensure feasibility and effectiveness. Key areas that will need attention include training delivery, recruitment strategy, and data collection and monitoring. These are discussed in turn below.

Training delivery

- ***Format:*** *The format and logistics of training delivery will be one of the key elements that may need to change at a scaled version of Early Number with Numicon. During the pilot, both in-person and online training formats were used, and at endline, most practitioners (15 out of 21) expressed a preference for in-person training, while only a small proportion indicated no strong preference (5 out of 21). For a larger trial, the feasibility of delivering in-person sessions to a significantly larger group would need to be explored. Similarly, online or hybrid delivery may present logistical and engagement challenges, particularly in terms of replicating the interactive and responsive nature of smaller-scale training. Addressing absence from training and offering consistent catch-up opportunities would need to be more formalised to maintain consistency in delivery. To support training, the delivery team could explore the development of an online portal or website to host training resources and recordings as it could provide practitioners with more flexible access.*
- ***Content:*** *Moreover, tailoring training content to the varying levels of practitioner knowledge often achieved during the pilot on an ad hoc basis, may become more difficult at scale, as would ensuring content is appropriately pitched to a more diverse group of attendees.*

Recruitment strategy

While both PVIs and childminders engaged with Early Number with Numicon, the challenge of identifying and recruiting childminders especially in areas where the delivery team does not have existing networks may be exacerbated in a scaled-up trial. Geographic spread may also affect attendance at in-person training, and delivery support visits would become increasingly costly and logistically difficult if they were to be maintained across multiple regions.

To support recruitment in a scaled trial, several strategies may be helpful. These include working with local authorities, early years networks to access contact lists, and extend reach and assigning dedicated recruitment roles to maintain consistent engagement. Tailored communications that highlight relevance and flexibility, alongside options such as flexible training times or small incentives, may also help address common barriers, particularly for childminders.

Data collection and monitoring

- ***Fidelity:*** *The pilot demonstrated that monitoring of training via attendance tracking and gaining timely feedback via post-training surveys are effective. However, tracking which activities have been delivered and when, will be beneficial when delivered to larger number of settings, in order to monitor implementation fidelity in a consistent way.*
- ***Baseline data:*** *Furthermore, to meaningfully assess the programme's impact, it would be important to assess children's numeracy skills before delivery. This would help isolate the effect of Early Number with Numicon. As highlighted by case study participants, it was difficult to determine whether observed numeracy improvements were specifically attributable to the programme or part of expected developmental trajectories.*
- ***Outcome measures:*** *In discussion with the OUP delivery team, key child outcomes for a future evaluation were identified. These include children reaching the essential 'must have level' by developing early concept knowledge, engaging in pattern work, and confidently subitising quantities of 1 to 3. The delivery team emphasised the importance of children applying these concepts independently during play as a clear indicator of success. In the short-term, outcomes are expected to be more visible in practitioner learning and shifts in practice. For meaningful midpoint evaluation, baseline and follow-up assessments ideally at six and twelve months would be needed to track children's progress over time.*

Overall, the pilot provides a strong foundation for conducting an RCT with both PVIs and childminders. However, scaling up to an efficacy trial will require thoughtful design and planning across training, recruitment, and data collection. With appropriate adjustments and dedicated infrastructure, an RCT would be feasible.

Cost and resources: Feasibility and acceptability

Research question: 3.4 What costs/resources are required by settings implementing Early Number with Numicon and is this feasible and acceptable, at a larger scale?

The cost and time requirements associated with implementing Early Number with Numicon were minimal and largely deemed acceptable by participating practitioners and setting leads. Both financial and time investments were manageable within the setting, suggesting strong feasibility for scale-up, provided similar support remains in place.

Costs

There were minimal direct costs incurred by settings related to delivering Early Number with Numicon activities. Most childminders (11 out of 14) and leads (six out of eight) reported that they did not need to purchase any other resources or supplies in addition to the resources shared with them at the first session to deliver the Early Number with Numicon activities. Similarly, the majority of childminders (8 out of 14) and leads (seven out of eight) did not spend any money for delivery.

In cases where settings did make purchases, these were typically low-cost items such as downloadable worksheets or additional Numicon materials (e.g. manipulatives, and baseboards). The most spent by a setting was £40, which covered a mix of supplementary materials like puppets, sand, coloured objects, and food colouring to support practical, engaging, numeracy activities. Many practitioners reported drawing inspiration from online platforms (e.g. Pinterest, TikTok, and Twinkl) and then adapting these ideas using materials already available in their settings. This creative, resourceful approach to resource preparation further helped limit additional expenditure.

Importantly, all practitioners were paid by the OUP to attend training, which significantly reduced the perceived financial burden and contributed to overall acceptability. All practitioners and leads were compensated £7.50 per hour to attend the training, and all participating settings received a full set of Numicon resources to implement the programme, free of charge.

Table 15: Table showing whether childminders/PVI leads purchased additional Numicon resources since the first session

	Childminder/PVI lead	Yes	No
Whether purchased additional Numicon resources	PVI lead	2	6
	Childminder	3	11

Source: Leads endline survey. Base: Childminders (n=14).Base: PVI leads (n= 8).

Table 16: Table showing whether childminders/PVI leads purchased any other resources or supplies to deliver Numicon activities

	Childminder/PVI lead	Yes	No
Whether purchased any other resources or supplies to deliver Numicon activities	PVI lead	1	7
	Childminder	6	8

Source: Leads endline survey. Base: Childminders (n= 14). Base: PVI leads (n= 8).

Time

The programme required attendance at four training sessions: two in-person sessions held in Bradford; and two online sessions delivered after working hours. Each session lasted between two and three hours. Despite some participants being based in Lancashire, attendance and engagement did not appear to be negatively impacted by travel or timing requirements.

Practitioners reported that they had enough time to prepare for both delivery and training attendance. However, many noted a desire to spend *more* time delivering the programme to avoid a sense of 'rushing through' activities. This reflected a strong engagement with the content rather than a lack of time to prepare for delivery.

Weekend or evening training was considered acceptable by most participants, and in-person sessions were particularly valued due to the perceived quality of learning and opportunities for discussion. These preferences align with the realities of early years work patterns and suggest that similar scheduling and format considerations would remain effective at scale.

Conclusion

Early Number with Numicon was implemented in 29 settings with the aim to improve early years practitioners' knowledge and confidence in developing early numeracy skills among children in their care, and to improve children's early numeracy skills/development. The pilot evaluation of Early Number with Numicon used a mixed-method approach to examine the feasibility of delivering the programme across PVIs and childminders. It focused on three overarching areas of interest: i) evidence of promise; ii) feasibility of implementation; and iii) readiness for trial. As detailed above (see 'Research questions' subsection in the 'Introduction' section), each of these areas had specific research questions. The research questions and the corresponding findings are summarised in Table 17.

Table 17: Summary of pilot findings

Research question	Finding
Evidence of promise: Is there evidence of promise that the programme may lead to the changes expected in the Theory of Change?	The pilot evaluation evidence is suggesting that Early Number with Numicon can improve practitioners' knowledge and confidence in early numeracy as well as children's early numeracy development. Practitioners reported increases in specific pieces of knowledge or confidence in mathematics and numeracy teaching from baseline to endline, even when starting from high confidence. All surveyed practitioners reported that Early Number with Numicon positively impacted children's early numeracy. By the end, all practitioners reported using and adapting the Numicon resources, indicating that the approach has become embedded in practice.
Feasibility of implementation: Is Early Number with Numicon feasible and acceptable to PVIs and childminders?	The programme was considered feasible to implement and was well-received by practitioners in both PVIs and childminders. Most practitioners attended all training sessions and felt well-prepared to deliver Early Number with Numicon. The 11 structured learning sequences were delivered progressively. Reaching learning sequence 7 (around two-thirds of the programme) was identified as a key benchmark for meaningful participation. By the end, 15 settings (out of 21) had reached this point or beyond, indicating strong engagement. Fidelity was rated medium to high, with no major barriers preventing PVIs and childminders from delivering Early Number with Numicon. However, some settings faced practical challenges (e.g. part-time/inconsistent attendance, balancing delivery with care of children two years old). Practitioners adapted creatively to ensure accessibility for most children. The programme was relevant and easy to integrate into daily practice, requiring only minor adjustments. Overall, the programme was considered both practical to implement and appropriate for a wide range of early years settings.
Readiness for trial: How feasible is it to scale delivery (and undertake evaluation using an RCT design) with this population (i.e. PVIs and childminders)?	Early Number with Numicon shows strong readiness for scale-up. Engagement was high, with 29 of the 31 recruited settings engaging with the programme through to the end in some capacity. Pilot data collection (attendance tracking, post-training feedback surveys) were adequate, but would require strengthening for an efficacy trial, including standardised fidelity monitoring tools to accurately capture how closely the programme is delivered as intended, tracking children's progress in numeracy skills, and improving attendance records to better track implementation. Early Number with Numicon can be scaled with minimal changes. Recommended refinements include adapting training for larger and diverse groups, providing tailored guidance for working with younger children and those with SEND or EAL, improving childminder recruitment and outreach, reviewing fidelity measurement, and enhancing data collection to better track implementation and outcomes.

Furthermore, for each of the three areas of interest of the pilot evaluation and the corresponding research questions, indicators of success were mapped out. The high-level assessment of the achievement of each of these indicators is presented in Table 18 below.

Table 18: Summary of the findings against the success indicators

Research question	Success indicator	High-level summary finding
Evidence of promise: Is there evidence of promise that the programme may lead to the changes expected in the Theory of Change?	Improvement in practitioners' knowledge and confidence in supporting children's early numeracy development	Practitioner self-report rating of their confidence in specific pieces of knowledge increased between baseline and endline for all measures, although some measures were already high at baseline
	Practitioners perceive Early Number with Numicon to have a positive impact on children's early numeracy development	All practitioners felt that Early Number with Numicon had a positive impact on children's early numeracy development, to at least some extent
	Practitioners make changes to their practice as a result of taking part in Early Number with Numicon	By the end of the programme, all practitioners used Early Number with Numicon (compared to very few at the start). Many settings did not follow the activity booklet rigidly and adapted the activities to the interests and needs of their children
Feasibility of implementation: Is Early Number with Numicon feasible and acceptable to PVIs and childminders?	Practitioners attend all training sessions	Nearly all practitioners attended all training sessions and felt that it prepared them well to deliver the Early Number with Numicon programme
	Practitioners complete the programme activities (i.e. small group activities and continuous provision for children)	Almost all practitioners completed up to and including learning sequence 6 by the end of the programme. While most did not progress beyond this point within the pilot time frame, this should not be interpreted as a lack of engagement. Many practitioners reported taking the programme at their own pace
	Early Number with Numicon is delivered with medium to high fidelity as assessed with the fidelity measure	There is medium to high fidelity in all setting types across all fidelity measures (see Table 14)
	Absence of any major barriers to delivery and for children from disadvantaged backgrounds	There are no major barriers that prevent PVIs and childminders from being able to deliver Early Number with Numicon. Some challenges made delivery more difficult, but settings were creative in finding ways to ensure delivery ran smoothly. This enabled most children, including those from disadvantaged backgrounds, to take part. Children with SEND and EAL were generally able to engage in Early Number with Numicon activities without significant obstacles, with practitioners adapting the content as needed to make it suitable and at the correct level
	Practitioners consider the intervention practical to implement (with minor amendments)	Practitioners found the programme relevant to their setting and practical to implement with the majority of their three–four-year-old children
Readiness for trial: How feasible is it to scale delivery (and undertake evaluation using an RCT design) with this population (i.e. PVIs and childminders)?	High-level engagement with the pilot in terms of recruitment and low attrition	Overall, there was a high level of engagement, and low levels of attrition in the pilot—with 29 settings of the targeted 30 reaching the end of the programme in some capacity.
	There are viable strategies in place to collect sufficient data to monitor delivery	Some monitoring work is already in place and sufficient for the scale of the pilot, this should be further developed for a scaled-up trial
	Early Number with Numicon can be scaled for an efficacy trial (with no or minor amendments)	The programme can be scaled for an efficacy trial with minimal amendments for PVIs and childminders. The following factors should be considered in the scale-up, which could differ from this pilot: • Refining the training format

Research question	Success indicator	High-level summary finding
		• Addressing recruitment challenges for childminders • Making minor refinements to the programme content • Strengthening data collection for delivery monitoring
	The resource commitment required to deliver Early Number with Numicon is acceptable	The cost and time commitment of the programme were deemed acceptable by the majority of practitioners and setting leads

Interpretation

Evidence to support the Theory of Change

There is good evidence to support the Theory of Change. The short-term outcome related to increases in practitioner knowledge, confidence, capacity, and skills were all realised to some extent after participating in the programme. Initially, many practitioners were unfamiliar with Numicon, but by the end, all reported confidence in delivering the programme, and in their own ability. Childminders, in particular, experienced substantive growth, often citing this programme as their first structured numeracy training. Similarly, the outcomes identified in the Theory of Change related to children were realised. Practitioners felt the programme had a positive impact on increasing children's engagement and knowledge in numeracy. Most children had high levels of interest and enjoyment. Practitioners noted enhancements in counting, pattern recognition, and the use of Numicon-related language. The programme encouraged practitioners to make adaptations to keep activities engaging for children. Evidence suggests that these benefits extended beyond the children directly involved in the programme, with ripple effects observed in other children within the settings, indicating that practitioners' improved practice had a wider influence on numeracy learning. The approach to mathematics teaching in this programme was different from other mathematics resources and programmes used by practitioners prior to the programme. Previously, practitioners relied on general resources like the EYFS and other types of manipulatives. Numicon introduced a more systematic method, which was new to many settings (particularly childminders). This structured approach was particularly beneficial in settings with limited prior numeracy programmes. Based on this, it is clear that practitioners' numeracy approach changed as a result of involvement in the programme. Most embraced the training and delivery and adopted more structured planning and mindful use of mathematical language as a result of this. Numicon resources became integral to most practitioners' practice, with increased use of resources throughout the programme. Variations in resource use were observed, with childminders and PVI settings integrating different tools based on their setting, and the needs of the children.

Feasibility of approach

The programme was delivered mostly as planned with few adjustments in the settings. The programme was acceptable to PVIs and childminders, with positive feedback and plans for continued use. Overall, the programme was well-received, with practitioners stating they would recommend it to others. An interesting divergence from planned delivery was the involvement of younger children, than originally planned, in the structured activities, due to the way settings were organised (e.g. childminders having children with a range of different ages involved in one activity at once). Practitioners acknowledged that the programme (training and activities) was effectively designed for three–four-year-olds but fed back that it would be helpful to have guidance for implementing with younger children as well, alongside guidance for delivery with SEND children.

Engagement with the programme initially was driven by a desire to enhance children's numeracy skills. Practitioner engagement with the training sessions was medium to high, with most practitioners attending, with occasional absences occurring due to unforeseen circumstances. Practitioners valued the training, rating both trainers and materials positively. The hybrid model was generally well-received, with some tweaks to be considered for scaling up. Practitioners progressed through learning sequences at varying rates, adapting activities to suit children's interests. They prioritised delivering the

programme to the children in a flexible manner, that was responsive to the level and ability of the children, over strict adherence to the programme schedule of learning sequences.

Contextual factors such as time demands and the existence of younger children in settings posed challenges for practitioners in programme delivery, as did scheduling and staffing shortages. However, practitioners were creative and were able to adapt delivery around these barriers (e.g. by integrating activities during nap times or adjusting sessions for developmental stages). EAL and SEND children generally engaged well, benefiting from the programme's repetitive nature. Practitioners adapted content to suit individual needs.

Readiness for trial

The pilot successfully recruited and retained the vast majority of settings and practitioners for the pilot. A total of 34 settings signed the MoU, with minimal attrition due to eligibility issues or other reasons, meaning 29 settings took part in the pilot and remained engaged with the programme through to the end in some capacity. This suggests the current recruitment and engagement techniques (given the scale of the pilot) were suitable. However, PVIs were easier to recruit than childminders, who required more personalised outreach. When scaling up for a larger trial, the delivery team would need to allocate greater resource for the recruitment efforts, to allow time for sourcing contact information and for the engagement with the settings.

The pilot indicates that the feasibility of data collection processes is promising. Basic mechanisms like tracking attendance and collecting feedback through surveys provided valuable insights. However, a more formalised system is needed for a larger trial to ensure consistent monitoring of implementation quality. This would help track progress and identify any challenges systematically, to help the delivery team be reactive to the needs of practitioners.

Financial and time investments required by settings when implementing Early Number with Numicon were manageable and acceptable for settings, indicating strong feasibility for scaling up. Weekend or evening training sessions were well-received, aligning with the working patterns of those working in the early years. Additionally, most settings did not incur significant expenses taking part, with only a few purchasing low-cost supplementary materials. Practitioners were paid for training attendance, which reduced any financial burdens (the training was free to attend, and they were given resources for free). This contributed to the programme's overall acceptability. The findings suggest that scaling up is feasible and acceptable, provided similar support is maintained.

Looking to the future, conducting an evaluation of this programme using an RCT design including both PVIs and childminders is feasible and would provide valuable evidence on the programme's effectiveness across diverse early years settings. Given that the pilot suggested positive outcomes in both groups—and recognising the limited CPD opportunities available to childminders—it would be beneficial and equitable to include them. An equal split between PVIs and childminders may be efficient, but the final design will need to consider statistical power and the smaller number of children typically cared for by childminders, which could make achieving balance more challenging. Careful planning around recruitment, training delivery, baseline data collection, and fidelity monitoring will be essential to ensure a robust evaluation. It is also important to note that the programme may have a cascade effect, with practitioners sharing learning with colleagues. For a full-scale trial, this could influence outcomes and would need to be considered in the study design to minimise potential contamination.

The pilot did not include maintained settings, as these were deemed ineligible. Given that the recruitment processes, programme delivery, and contextual factors identified in the pilot were specific to PVIs and childminders the pilot would not recommend including maintained settings in a future trial without further preparatory work.

Formative findings

Based on evidence emerging in the pilot evaluation we would recommend some areas for further consideration if the programme were to be scaled up, particularly in relation to recruitment and engagement, training, session delivery, and programme duration. This section consolidates these findings and outlines recommendations for refinement to inform future delivery and evaluation.

Recruitment and engagement

While practitioners engaged well with the Early Number with Numicon programme, the pilot found that identifying and recruiting childminders proved difficult. This may be more challenging in areas where the delivery team lacks existing networks (which was the case in this pilot). To support recruitment at scale, the delivery team could consider partnering with local stakeholders in new areas, to expand access to contact lists as well as to establish trust in new areas. It is also recommended to have a dedicated team member as in the pilot that can focus on recruitment and engagement, ensuring consistent communication. These measures could include more comprehensive communication of the programme's benefits (e.g. enhanced training and improved knowledge), and additional drop-in session for catching up.

Training

The structured training sessions led by experienced trainers were viewed positively. Both the training sessions and the practitioner booklet were instrumental in equipping practitioners with the knowledge and skills necessary for effective programme implementation. However, in a scaled-up trial, several adjustments to both format and content may be necessary to ensure effectiveness and accessibility across a larger, more diverse practitioner cohort. This diversity includes variation in the location and type of settings, previous experience with numeracy delivery and/or specific resources such as Numicon, differing levels of interest in the programme, and differences in how and why practitioners join (e.g. their manager could have signed them up, or they may sign up for different reasons than those in this pilot). Addressing these factors will be important to ensure training is relevant and engaging for all participants.

Format of the training

During the pilot, training was delivered through a blend of in-person and online sessions. While this blended model worked well in a smaller-scale context, its scalability warrants careful consideration. Expanding the programme geographically could affect attendance at in-person training sessions and increase the cost and logistical complexity of on-site delivery support.

It will also be important to formalise processes for practitioners who are unable to attend training at the scheduled time, ensuring that catch-up opportunities are consistently available. This could include access to recorded sessions, supplementary online resources, or scheduled follow-up support, helping to maintain fidelity while providing greater flexibility at scale.

Content of the training

Tailoring the training content to practitioners' varying levels of experience and prior knowledge was achieved on an ad hoc basis in the pilot. However, this approach may become unsustainable at scale, particularly given the likely increase in variation across roles, qualifications, and familiarity with early mathematics pedagogy. When considering an expanded cohort of practitioners on the programme, it may be beneficial to adjust materials to better accommodate varying levels of practitioner expertise and experience, that are more likely to present at a larger scale.

Session delivery and engagement

The delivery team might consider providing further support and advice to practitioners on adapting the content of activities to specific profiles of children. For example, the pilot found that many childminders were engaging younger children in the programme, and they wanted more support on how to deliver the programme successfully with a younger age group. Despite the pilot being aimed at three–four-year-olds, childminders were likely to engage younger children due to time and staff constraints. Practitioners indicated that younger or older children could participate when activities were adapted as needed but emphasised that any engagement with these age groups relied on their own judgement and adaptations. Guidance on how to deliver sessions for children outside the three–four-year-old range was highlighted by practitioners as a useful addition to support effective delivery in diverse settings.

Practitioners also expressed a desire for more support in tailoring activities to children with SEND or EAL. It is recommended that trainers continue emphasising the importance of tailoring programme content to meet the specific needs of the children within practitioners' settings, both during training sessions, within the practitioner booklet, and through additional communications as required.

Programme duration

The delivery team may wish to consider whether the programme duration should be extended. The current 14-week structure left some practitioners unable to complete all the activities. Delivering one or two sessions a week was not always feasible given time constraints and the individual pace of children's learning (which may require a lot of repetition and delivery of multiple activities within a sequence). An extended timeline, such as incorporating additional time between training sessions, or adjusting expectations regarding the pacing of delivery for practitioners, could be explored. This adjustment would ensure practitioners maintain confident and do not feel pressured to complete approximately one learning session per week, if it does not work for them and their setting.

Limitations

The limitations of the pilot evaluation of Early Number with Numicon are akin to that of any small-scale pilot that operates in a singular geographic location. On this basis, it is important to keep in mind that the findings of the pilot are indicative of the outcomes of the pilot. They should be interpreted with caution and understood as directional rather than definitive. Alongside a small sample size and location, the pilot recruitment exercise was conducted as an opt-in study, which could have created a self-selection bias in the sample. The recruitment criteria, as detailed in the 'Methods' section, meant that only certain settings within the area were able to take part. For example, those that are not able to commit staff to training, or those that undertook other funded work by the EEF in 2024/2025, could not participate.

A further limitation to the small sample size is that the differences between childminders and PVI practitioners in the survey responses cannot be compared for statistically significant difference. Additionally, the evaluation relied on self-report survey data to assess changes in practitioner knowledge and confidence. However, the number of respondents at endline was lower than at baseline, which limits the comparability of results over time and may reflect differences in respondent characteristics rather than true change. As such, these findings should be interpreted cautiously.

There is an additional limit to capturing data by overt observation in early years settings and training sessions, namely creating bias in terms of the behaviour of children and practitioners.

Future evaluation work

A key next step for the Early Number with Numicon programme would be to conduct an efficacy trial, ideally using an RCT design. We recommend randomisation at the setting level, with a 'business as usual' control condition, to avoid contamination and to allow for a rigorous assessment of the programme's added value.

Notably, the pilot highlighted that some programme components (e.g. the use of manipulatives or foundational numeracy concepts) were already familiar to practitioners. As such, a critical task in future evaluation will be to explore and clearly define 'business as usual' practice, in order to isolate the effect of the Early Number with Numicon programme.

To support robust outcome measurement, baseline assessment of children's numeracy skills will be essential. This is particularly important as many practitioners expressed uncertainty about children's starting points. For validity and objectivity, numeracy assessments should be conducted by independent evaluators, rather than relying on practitioner-administered measures. Moreover, careful consideration must be given to the choice of outcome measures as robust, reliable instruments to assess early numeracy were not piloted in this study.

Ideally the trial design should allow for subgroup analyses, particularly to investigate differences in impact between PVIs and childminders. This will require a sufficiently large sample size—something that may be challenging, especially for the childminder group, given the smaller numbers of eligible children per setting. Despite this challenge, we recommend

retaining childminders in any future trial, as pilot findings suggest they experienced substantial benefit from the programme and showed high levels of engagement.

In addition, fidelity monitoring should be refined. During the pilot phase, four fidelity measures were used. All four achieved medium or high fidelity. Attendance at training was treated as a critical component of programme delivery, with the measure defined as 100% of practitioners attending all four sessions, either during scheduled delivery or through catch-up arrangements afterwards. While this benchmark provided a clear standard, it may not be fully realistic at scale. Similar, it is advisable for future evaluation teams to reassess a fidelity measure that requires settings to reach a specific sequence in the programme at a specific time point, as it may not accurately capture the overall quality or success of the programme implementation. It would be sensible to develop an alternative fidelity indicator(s) that would be able to reflect the quality of engagement with the programme (by practitioners and children) rather than strict adherence to attendance or reaching specific sequence. Such indicators could include consistency in delivery, or the effective adaptation of the programme to meet the individual needs of children.

References

Bonetti, S. and Landen, J. (2020) *'Early Years Workforce Qualifications and Children's Outcomes: An Analysis Using Administrative Data'*. London: Education Policy Institute. Available at: https://epi.org.uk/publications-and-research/early-years-qualifications-and-outcomes/ (accessed 13/3/2024)

Cattoretti, G., Paull, G. and Marshall, L. (2019) *'Providers' Finances: Evidence from the Survey of Childcare and Early Years Providers 2018'*. London: Department for Education. Available at: https://assets.publishing.service.gov.uk/media/62a9cf538fa8f50396173333/Frontier_SCEYP_2018_Finance_Report_Revised_April_2022.pdf (accessed 14/3/2024).

Clark, A., Henderson, P. and Gifford, S. (2020) '*Improving Mathematics in the Early Years and Key Stage 1: Guidance Report'*. London: Education Endowment Foundation. Available at: https://d2tic4wvo1iusb.cloudfront.net/production/eef-guidance-reports/early-maths/EEF_Maths_EY_KS1_Guidance_Report.pdf (accessed13/3/2024).

Department for Education. (2024) *'Methodology: Childcare and Early Years Provider Survey'*. GOV.UK. Available at: https://explore-education-statistics.service.gov.uk/methodology/childcare-and-early-years-provider-survey (accessed 04/07/2025).

Education Endowment Foundation (EEF). (2022) '*Implementation and Process Evaluation Guidance for EEF Evaluations'*. London: Education Endowment Foundation. Available at: https://d2tic4wvo1iusb.cloudfront.net/production/documents/evaluation/evaluation-design/EEF-IPE-Guidance-August-2022.pdf?v=1750279461 (accessed 14/3/2024).

Education Endowment Foundation (EEF). (2025) *'Early Years Toolkit'*. London: Education Endowment Foundation. Available at: https://educationendowmentfoundation.org.uk/early-years/toolkit (accessed 04/07/2025).

General Data Protection Regulation (GDPR). (2016) *'Council Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data (United Kingdom General Data Protection Regulation) (Text with EEA relevance)'*. Available at: www.legislation.gov.uk/eur/2016/679 (accessed 04/07/2025).

Government Social Research (GSR). (2021) '*GSR Professional Guidance: Ethical Assurance for Social and Behavioural Research in Government'*. Available at: https://assets.publishing.service.gov.uk/media/60e808d0d3bf7f56801f3c3f/2021-GSR_Ethics_Guidance_v3.pdf (accessed 14/3/2024).

Hodgen, J., Barclay, N., Foster, C., Gilmore, C., Marks, R. and Simms, V. (2020) *'Early Years and Key Stage 1 Mathematics Teaching: Evidence Review'*. London: Education Endowment Foundation. Available at: https://educationendowmentfoundation.org.uk/education-evidence/evidence-reviews/early-years-and-key-stage-1-mathematics-teaching (accessed 14/3/2024).

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R. and Kerr, K. (2016) *'Implementation and Process Evaluation (IPE) for Interventions in Education Settings: An Introductory Handbook'*. London: Education Endowment Foundation. Available at: https://d2tic4wvo1iusb.cloudfront.net/documents/evaluation/evaluation-design/IPE_Handbook.pdf (accessed 14/3/2024).

Joshi, A., Kale, S., Chandel, S. and Pal, D.K. (2015) 'Likert Scale: Explored and Explained'. *British Journal of Applied Science & Technology,* 7: 4, 396–403. https://doi.org/10.9734/BJAST/2015/14975

Market Research Society. (2023) *'Code of Conduct'*. Available at: www.mrs.org.uk/pdf/MRS-code-of-conduct-2023.pdf (accessed 14/3/2024).

OSF Registries. (2024) *'Early Number with Numicon: Pilot Evaluation'*. OSF. August 27. https://doi.org/10.17605/OSF.IO/ZMG3Q

Parliamentary Office for Science and Technology. (2021) 'Impact of Covid-19 on Early Childhood Education & Care'. POSTnote 682, 27 October. UK Parliament: Parliamentary Office for Science and Technology. https://doi.org/10.58248/RR74

Paull, G. and Popov, D. (2019) *'The Role and Contribution of Maintained Nursery Schools in the Early Years Sector in England'*. London: Department for Education. Available at: https://assets.publishing.service.gov.uk/media/5f48d4398fa8f57fb653f840/Frontier_Economics_MNS_report_REVISED_v2.pdf (accessed 04/07/2025).

Raghubar, K.P. and Barnes, M.A. (2017) 'Early Numeracy Skills in Preschool-Aged Children: A Review of Neurocognitive Findings and Implications for Assessment and Intervention'. *The Clinical Neuropsychologist,* 31: 2, 329–351. https://doi.org/10.1080/13854046.2016.1259387

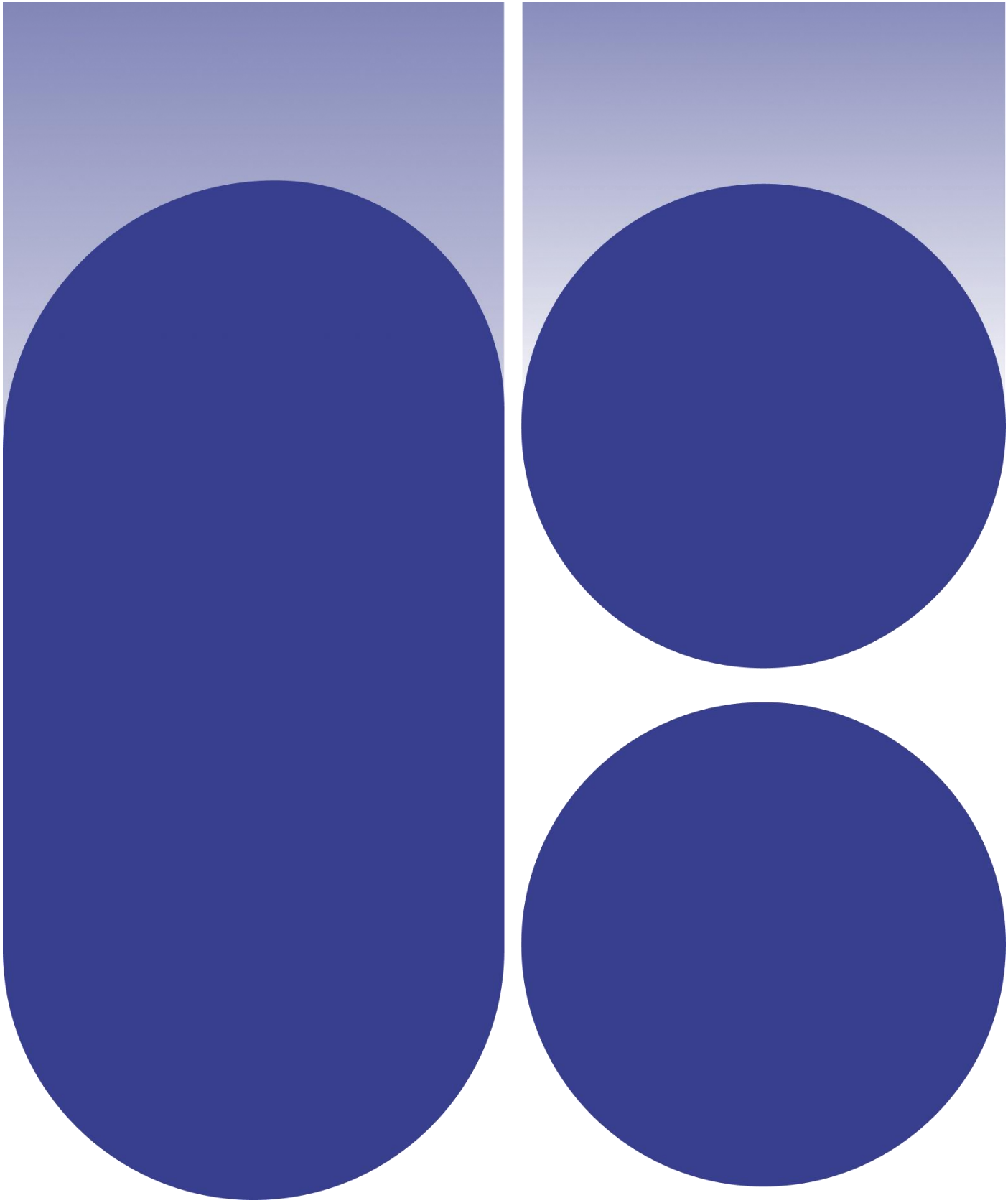
Social Research Association. (2021) *'Research Ethics Guidance'*. Available at: https://the-sra.org.uk/common/Uploaded%20files/Resources/SRA%20Research%20Ethics%20guidance%202021.pdf (accessed 14/3/2024).

Thomas, D.R. (2006) 'A General Inductive Approach for Analyzing Qualitative Evaluation Data'. *American Journal of Evaluation,* 27: 2, 237–246. https://doi.org/10.1177/1098214005283748

Appendices

Appendices are published as a separate document available on project page.

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



Education Endowment Foundation

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

https://educationendowmentfoundation.org.uk

@EducEndowFoundn

Facebook.com/EducEndowFoundn