Counting Collections
Evaluation Protocol

Sheffield Hallam University
Dr Martin Culliney, Dr Karen Daniels, Joanne Robson

Education Endowment Foundation

PROJECT TITLE[1]	Evaluation of Counting Collections, a two-armed cluster randomised trial
DEVELOPER (INSTITUTION)	University of Nottingham
EVALUATOR (INSTITUTION)	Sheffield Hallam University
PRINCIPAL INVESTIGATOR(S)	Dr Martin Culliney, Dr Karen Daniels
PROTOCOL AUTHOR(S)	Dr Martin Culliney, Dr Karen Daniels, Joanne Robson
TRIAL DESIGN	Two-arm cluster randomised controlled trial with random allocation at school level
TRIAL TYPE	Efficacy
PUPIL AGE RANGE AND KEY STAGE	Age 4-5, Early Years (Reception)
NUMBER OF SCHOOLS	180
NUMBER OF PUPILS	3600
PRIMARY OUTCOME MEASURE AND SOURCE	Number attainment, GL Sandwell Early Numeracy Test
SECONDARY OUTCOME MEASURE AND SOURCE	N/A

Protocol version history

VERSION	DATE	REASON FOR REVISION
1.1 [*latest*]	9 May 2024	Changes to the evaluation plan since original protocol: Pre-intervention survey dropped. Post-intervention survey still scheduled for summer 2024. Baseline assessments were extended into November 2023 due to assessor capacity. Biographies for the evaluation fieldworkers have been added.
1.0 [*original*]	20 July 2023	

[1] Make sure that the project title here matches the title of the document. Please ensure that there is an identification as a randomised trial in the title as per CONSORT requirements.

Table of contents

Study rationale and background

Pupils start school with different experiences of number which affects how they access opportunities to learn number, through play and through teacher-directed activities, creating a gap in early number attainment. Number sense is crucial to mathematics attainment and particularly important for children from disadvantaged communities who are at increased risk of lower mathematics attainment.

Counting Collections is a hands-on early mathematics approach to develop children's number sense (understanding of number and quantity). The Counting Collections approach supports counting, subitising, comparing numbers and composition of numbers. It involves children using containers of objects (manipulatives) to find out how many are in the collection. This requires the practitioner to have the confidence and knowledge of maths as an interconnected subject to 'ensure every interaction with a child is an opportunity for teaching and learning' (Ofsted 2021). Interactions between pupils, and between pupils and adults, are both central to Counting Collections.

The intervention includes opportunities for learners to engage in number interactions through play with teachers and peers. Effective learning in the early years comes through quality interactions with adults; this includes their interactions with children during planned and child-initiated play and activities (Ofsted 2022). The adult's understanding of intentional mathematical learning through play is therefore pivotal to pupil progress. Timely interactions, modelling, questioning, and reflective commentary are used to direct the young learner, requiring good knowledge of the complexities of number development. A pilot project of CC carried out in 2017-18 involved a two-day version of the PD programme in 21 early years settings with 600 2- to 5-year-olds across the East Midlands (Gripton and Pawluch, 2021). In this small-scale research study, Counting Collections supported development of teacher subject knowledge (rooted in developmental progressions in number), which led to increased progress and interest in number from the children as perceived by teachers. Overall, while the pilot study reported positive findings, the limited scope provided no opportunity to examine the transference of number skills into other learning situations or the retention of number skills over time.

Other relevant evidence includes Let's Count; a preschool mathematics intervention implemented from 2012 in "disadvantaged" communities across Australia. Findings indicated that learning and confidence increases when adults notice children's work in mathematics (Perry et al., 2016). In another relevant study, the Every Child Counts intervention was evaluated through a large-scale study of 8,000 low-achieving 6- and 7-year-old children. After an average of 43 half-hour, one-to-one lessons in three months, their number age test scores had risen by 14 months with an effect size of .85. Pupil attitudes towards learning mathematics, measured through teacher perceptions using a bespoke attitude survey rather than pupil self- reports because of the age of the children, also improved substantially, with an effect size of 0.7. (Henry et al., 2007). Children made strong progress irrespective of their background. It is suggested that the success of the intervention was due to its design, the teacher professional development programme, and the rigorous quality assurance (Dorwick, 2015). Despite these positive findings, the transference and retention of number skills were not studied.

This trial aims to identify the characteristics of the intervention which promote children's ability to retain and transfer the number skills gained and consider the sustained impact of the intervention through pupils' increased motivation to learn and confidence as a mathematical learner. The evidence of this would be seen through the child-initiated play and their

interaction with the CC provision in the classroom setting. In this efficacy trial, CC is scheduled for delivery in the 2023/24 school year. The evaluation will consist of two strands. The impact evaluation will draw on pre- and post-intervention pupil number assessments conducted in schools, to accurately assess pupil ability in number before and after the programme is delivered. Concurrently, the implementation and process evaluation (IPE) will include teacher surveys to gauge number confidence and to understand existing practice in schools, along with fieldwork visits and training observations to examine how the professional development is received by teachers and to gain insights into implementation.

Intervention

The intervention is described here using the TIDIER framework.

Name

Counting Collections

Why (theory/rationale)

Counting Collections is a practical approach that allows pupils to make connections between concepts through structured collaborative counting and recording sessions. The weekly sessions are based on a four-part routine, with pupils working in pairs. A key factor in the success of the intervention is the role of the adult as the facilitator and their knowledge of teaching and learning in number, including developmental progressions (learning trajectories). A small-scale pilot study (Gripton and Pawluch, 2021) indicated that Counting Collections supported the development of teacher subject knowledge (rooted in developmental progressions in number), which led to perceived increases in pupil attainment and interest in number. This requires further investigation, especially around the varying degrees of teacher knowledge and the training needed.

Who (recipients)

The intervention is designed for reception pupils (aged 4-5 years).

What (materials)

A typical Counting Collections session follows a teacher-led sequenced structure that encourages and guides children to work in pairs to plan, count and record their approach to counting a set of manipulatives (‘collections’ of everyday objects such as straws, small soft toys, pinecones). These collections form a counting library in the classroom. There are tools to aid counting (such as boxes, jars and number frames). The manipulatives and tools are provided by the delivery team so that all settings have the same resources available for the activities.

What (procedures)

A structured routine for counting is taught in four parts: choose, strategise, count, record. This routine represents one of the core components of the programme. Collection sizes are increased over time alongside teaching of more complex strategies including counting unitised groups.

Who (provider)

One teacher per school takes part in the professional development programme provided by the developer. This includes a dedicated online environment that provides ongoing support

throughout the programme. Teachers can revisit content, communicate with developers and discuss Counting Collections with other teachers.

How (format)

Trained teachers deliver the intervention to pupils in class. If TAs are allocated to support these sessions, it is expected that teachers will train them so that they understand the intervention. However, TA participation in Counting Collections is optional for schools.

Where (location)

The intervention is delivered in reception classrooms at schools in England. Any mainstream maintained school with at least 20 pupils in the 2023/24 reception cohort is eligible to take part in the trial.

When and how much (dosage)

The intervention will be delivered during the 2023/24 school year. It is expected that teacher training will take place immediately before the autumn half term, so the first sessions with pupils are likely to take place in early November. Delivery will run until the summer half term. Ideally sessions will be held weekly during this time. A minimum of 20 sessions must be completed over the study period for a school to be classed as compliant.

Tailoring (adaptation)

Control schools will continue teaching number to reception pupils using a 'business as usual' approach. They will be paid a total of £500 for participation in the trial. This is conditional upon complying with the terms of the Memorandum of Understanding, specifically by sharing pupil data and facilitating the completion of pre- and post-intervention pupil assessments in school.

Impact evaluation

Research questions

1. What is the impact of Counting Collections on reception pupil attainment in number as measured by the Sandwell Early Numeracy Test?
2. What is the impact of Counting Collections on disadvantaged reception pupil attainment in number as measured by the Sandwell Early Numeracy Test?

Design

This two-arm, two-level clustered efficacy trial is delivered over one school year to pupils in reception classes at intervention schools. Control schools continue with business as usual during the study period. Baseline and endline outcome testing will be carried out in the first and final half terms of the 2023/24 school year respectively. Further details on pupil testing can be found in the relevant section below.

Table 1: Trial design

Trial design, including number of arms		Two-arm, two-level cluster randomised trial
Unit of randomisation		School
Stratification variables (if applicable)		Geographical area Existing use of Mastering Number intervention
Primary outcome	variable	Number attainment
	measure (instrument, scale, source)	GL Sandwell Early Numeracy Test (B) score
Secondary outcome(s)	variable(s)	N/A
	measure(s) (instrument, scale, source)	N/A
Baseline for primary outcome	**variable**	Number attainment
	measure (instrument, scale, source)	GL Sandwell Early Numeracy Test (A) score
Baseline for secondary outcome	**variable**	N/A
	measure (instrument, scale, source)	N/A

Randomisation

Randomisation will be conducted at school level to minimise spillover risk. Geographical area (South West, East Midlands and South Yorkshire, North East) will be used as a stratification variable, along with whether the school is already using the Mastering Number intervention. This is to reduce the risk of allocation imbalance in certain geographical areas undermining the viability of the training, and to mitigate against the use of other relevant interventions confounding the results of this trial. The evaluators will remain blinded to group allocation at the time of randomisation, but it will not be possible to maintain this position once the schools have been allocated.

Randomisation was completed on 14 July 2023. Parent information sheets were sent to schools on 5 September. Collection of pupil data from schools began on 18 September. Baseline pupil assessments began on 25 September. Teacher training for intervention schools began on 9 October. Teacher training began while pupil data collection and baseline assessments were still underway. Schools were already aware of their treatment allocation by the time any pupil data was collected or baseline assessments completed. This was deemed unavoidable given the lack of time to complete all evaluation and delivery activities, including the postage of Counting Collections materials to intervention schools.

Intervention delivery began after Autumn half term (30 October or 6 November depending on location). Ideally assessments would have been completed prior to schools starting delivery.

Due to difficulties in assessor availability, additional time to complete baseline assessments was required. The challenges were compounded by the half term break in schools, which is two weeks in the Nottingham area, limiting the time when schools could be visited. All baseline assessments were completed by the end of November 2023.

Participants

Schools will be recruited by the developer. Participation is conditional upon abiding by the terms of the MoU but all mainstream schools with at least 20 pupils in the 2023/24 reception cohort are eligible to take part in the trial. As the trial is supported by the DfE Accelerator Fund, the aim is to recruit 50% of schools from local authorities that are classed as Education Investment Areas.

Only one class per school can take part in the trial. It would be burdensome for schools to source cover for more staff to participate in the training and substantial class level effects are not expected with this pupil age group. Due to resource and time constraints the number of pupils included in the evaluation will be limited to 20 per school. For larger classes, 20 pupils will be randomly selected by the evaluators prior to baseline testing. There are no constraints on pupil eligibility other than they must be in reception (age 4-5) throughout the study year.

Sample size calculations

The design is a 2-level clustered RCT. In calculating the Minimum Detectable Effect Size (MDES), the smallest effect size that could be detected as statistically significant (often set as $p<0.05$) with a statistical power of 80% or higher, our estimates are based on the following assumptions:

$\mathbf{M_{j\text{-}k\text{-}2}}$ - T-distribution multiplier assuming a two-tailed test with a statistical significance of 0.05, statistical power of =0.80 and J-K-2 (175) degrees of freedom
$\mathbf{R_i}$ – Participant (pupil) level pre/post-test correlation of 0.6 ($R_i^2 = 0.36$)
$\mathbf{R_c}$ – Cluster (school) level pre/post-test correlation of 0.2 ($R_c^2 = 0.04$)
ρ – Intracluster correlation (ICC) 0.17
j – number of schools (180)
m – number of pupils per school (20)
k – number of cluster level covariates (3)
P - Proportion of schools allocated to intervention group (P=0.5)

The ICC and participant correlation values are taken from the evaluation of Maths Champions (Robinson-Smith et al., 2018), the only early years maths trial published by EEF at the time of writing. The ICC reported at the analysis stage of that trial was 0.17, and the pupil pre/post-test correlation was 0.59. These figures come from an evaluation using a different outcome measure with slightly younger children than targeted here, but nevertheless this is recent evidence from a trial of a programme of the same duration and on the same subject. Cluster level correlations were not supplied in Robinson-Smith et al., (2018) but are conservatively estimated here as 0.20. Calculations were performed in Excel using the formula set out in Bloom et al (2007), which relates to two-level clustered randomised controlled trials. This allows covariates to be included at both individual (pupil) and cluster (school) level, which in turn increases sensitivity.

Equation 1:Minimum Detectable Effect Size in a two-level clustered randomised controlled trial

$$\mathrm{MDES} = M_{j-k-2}\sqrt{\left(\frac{\rho(1-R_c^2)}{\mathrm{P}(1-\mathrm{P})\mathrm{J}}\right)+\left(\frac{(1-\rho)(1-R_i^2)}{\mathrm{P}(1-\mathrm{P})\mathrm{Jm}}\right)}$$

Based on the assumptions above, an MDES of 0.20 standard deviations would require 150 schools with 20 pupils each. However, to mitigate against potential attrition, 180 schools were recruited. This is the sample size at randomisation and gives an overall MDES of 0.18. The MDES will not exceed 0.20 if school level attrition does not exceed 16%.

For the subgroup of pupils eligible for free school meals (FSM), (estimated at eight per school, as the number is likely to be higher in Education Investment Areas and the aim is to recruit at least 50% of schools from these districts), the MDES is 0.20 with 180 schools. With 150 schools, the FSM MDES is 0.22.

Table 2: Sample size calculations

		OVERALL	FSM
Minimum Detectable Effect Size (MDES)		0.18	0.20
Pre-test/ post-test correlations	level 1 (pupil)	0.60	0.60
	level 2 (school)	0.20	0.20
Intracluster correlations (ICCs)	level 2 (school)	0.17	0.17
Alpha[2]		0.05	0.05
Power		0.80	0.80
One-sided or two-sided?		2	2
Average cluster size		20	8
Number of schools[3]	Intervention	90	90
	Control	90	90
	Total	180	180
Number of pupils	Intervention	1800	720
	Control	1800	720
	Total	3600	1440

Outcome measures[4]

Baseline measures

The Sandwell Early Numeracy Test (SENT) will be used as the baseline and primary outcome measure. It has two components (A and B) and is suitable for gauging the impact of classroom interventions on a pre- and post-test basis. SENT A will be used as the baseline measure for this trial, with SENT B as the outcome measure. The publisher has confirmed that it is appropriate to use the assessment in this way, and it has been used similarly in previous research (Torgerson et al., 2011:49).

[2] Please adjust as necessary for trials with multiple primary outcomes, 3-arm trials, etc., when a Bonferroni correction is used to account for family-wise errors.
[3] Please adjust as necessary, e.g., for trials that are randomised at the class level.
[4] Please see the Statistical Analysis Guidance.

The assessment explores five strands of basic numeracy skills: identification, oral counting, value, object counting and language, but these are not validated for use as standalone measures. As Counting Collections aims to improve number attainment overall, all aspects of the SENT assessment are relevant to the intervention. Baseline testing was conducted in schools by researchers blinded to group allocation between 25 September and 24 November 2023. These researchers are students on education and psychology degrees recruited by the subcontractor appointed by the evaluators to administer pupil assessments in schools. Additional assessors were recruited toward the end of the assessment period to cover a shortfall in capacity in one of the study regions. These were teaching assistants hired through a supply agency and undertook the same training as the other assessors.

Primary outcome

As described above, SENT B will be used as the primary outcome measure. Data collection is planned for June/July 2024 and will again be conducted by the evaluators in the same way as the baseline testing.

Secondary outcomes

No secondary outcomes are to be measured.

Compliance

The developer has specified three criteria that schools must meet to achieve full compliance:

- Delivering at least 5 weekly Counting Collections sessions per half term
- Teacher attendance at all five professional development sessions (first session in person, then synchronous online sessions)
- Graded manipulatives and supporting tools must be present in the classroom

If one of the online professional development sessions is attended asynchronously, full compliance is still achieved provided that the other criteria are met. Partial compliance can be achieved through teachers completing the online PD sessions asynchronously. The unit of analysis for compliance is the school, although as mentioned above only one teacher per school is taking part in the trial.

Analysis

A multilevel approach will be adopted, with pupils clustered within schools. Multilevel linear regression models will be constructed for the SENT (B) primary outcome. For each model, the coefficient of the school-level dummy variable used to distinguish 'intervention group' pupils in schools who will receive the Counting Collections programme from 'control group' pupils will be converted into Hedges' g effect size statistics with 95% confidence intervals.

The first model will only include the school level group identifier (an outcome only model). The second model will also include the baseline test score as a covariate at the pupil and school levels[5]. SENT (A) will be used as the baseline covariate for analysis of the primary outcome. The final model will also include the variables used in the randomisation process (geographical area, whether the school is using the Mastering Number intervention) and will form the headline intention to treat (ITT) impact analysis for the SENT B primary outcome.

[5] These will be centred so that the school level will be centred on the mean for all schools and the pupil level will be centred around the school mean.

Follow-on ITT analyses will focus on the impact of Counting Collections on number attainment among disadvantaged pupils, as defined by the NPD variable FSM_EVER_6. The same three model stages used for the headline ITT analyses will be used.

Longitudinal follow-ups

None planned.

Implementation and process evaluation[6]

The implementation and process evaluation (IPE) will provide context to support the impact evaluation, produce evidence of what works for whom and under what circumstances, and explore fidelity of implementation.

Research questions

a) How effectively does the training equip teachers to deliver the intervention and improve their understanding of key number concepts?
b) Is the intervention associated with improvements in teacher knowledge of learning trajectories?
c) Do teachers believe that the intervention is associated with short-term improvements in pupil confidence, enjoyment and attainment in number?
d) What level of fidelity is observed during the trial and what influences fidelity?
e) What do the trial findings indicate about scalability?

Research methods

The IPE design is based on the agreed evidence-informed logic model and comprises:

1. **Further evidence review and discussion with stakeholders** to build an agreed evidence-informed logic model and inform data collection methods. The significance of teacher knowledge of number learning and how this may affect pupil knowledge, confidence and enjoyment of number will be explored.

2. **Observation of six training sessions** by early maths specialists, to examine the content and format of the training and materials. This will improve understanding of the programme and inform the school fieldwork and post-intervention surveys. Records of training attendance and other engagement with support will also be analysed to gauge fidelity.

3. **Pre- and post-intervention online teacher survey** looking at 'business as usual', school context, mathematics training, teacher experience and confidence. The post-intervention survey will be completed by all schools but will include additional questions for the intervention group. These will address contextual issues aiding or hindering effective implementation and will be informed by fieldwork findings to cover any emerging themes.

Protocol amendment May 2024: the baseline assessment period was demanding for schools with staff training in intervention schools along with the requirement to submit pupil data and undertake the assessments. The evaluation team did not think it was appropriate to send the teacher survey during this period. It was agreed with the delivery team that sending the post-intervention survey after the training had begun and the programme was being implemented in schools would produce misleading data that would be likely to understate any changes in teacher responses associated with the intervention. As such, the pre-intervention survey was

[6] Please follow the principles detailed in the Implementation and Process Evaluation Guidance (2019).

shelved, and it was agreed that the post-intervention survey should be designed to address the relevant research questions.

4. **School visits** to understand fidelity and ascertain influences on implementation and the extent to which the teacher practice is aligned with expectations. Our team of early years numeracy experts will visit ten schools with a mixture of locations and characteristics to:

- Observe at least one teacher-led whole class CC session
- Observe at least one pupil play number interaction
- Examine the resources available in the 'maths area' of the classroom
- Interview the teacher and the school leader responsible for CC at school level

Analysis

Table 3 (below) illustrates how the data collection methods map onto the research questions. Descriptive statistics will be summarised from the surveys to gauge participant perceptions of confidence in mathematics teaching and, for respondents in intervention schools, experiences of the training and of delivering the programme. The Chen et al. (2014) questionnaire on early years maths teacher beliefs and confidence will be included in the post-intervention survey. Pertinent themes arising from the fieldwork visits and interviews with teachers and school leaders can also be included in the post-intervention survey if relevant to the research questions.

Fieldwork interviews will be one-to-one and semi-structured. An interview schedule guided by the main research questions and pre-intervention survey findings will be designed. Interviews will enable the evaluation team to ask how Counting Collections is working, for whom and under what circumstances. Teacher perceptions on the impact of Counting Collections on children's progress in number and any attitudinal changes will also be explored, along with any barriers to its successful implementation. Interview data will be coded using NVivo. A *variable-orientated* approach to case analysis will be taken to enable broad patterns and recurring themes, across cases, to be identified (Miles et al., 2019).

An observation schedule will be prepared for lesson observations. This will involve examining fidelity to the Counting Collections lesson structure, content and key pedagogical principles. Descriptive fieldnotes will be made to describe the ways in which a mathematics area is organised and the materials it contains for children. Analysis will involve identifying key features of manipulative and tool provision in classrooms and examples of good practice.

Table 3: IPE methods overview linked to research questions

Research question	Data collection method	Participants
How effectively does the training equip teachers to deliver the intervention and improve their understanding of key number concepts?	Training observations Interviews and observations in school Post-intervention survey	Teachers, trainers Teachers in 10 intervention schools Teachers in all intervention schools
Is the intervention associated with improvements in teacher knowledge of learning trajectories?	Interviews and observations in school Post-intervention survey	Teachers in 10 intervention schools Teachers in all intervention schools

Do teachers believe that the intervention is associated with short-term improvements in pupil confidence, enjoyment and attainment in number?	Interviews and observations in school Post-intervention survey	Teachers in 10 intervention schools Teachers in all intervention schools
What level of fidelity is observed during the trial and what influences fidelity?	Interviews and observations in school Post-intervention survey	Teachers in 10 intervention schools All intervention teachers
What do the trial findings indicate about scalability?	Interviews and observations in school Post intervention online survey	All intervention teachers

Cost evaluation

Cost calculations will be based on data provided by the delivery team. This will be conducted in line with EEF cost evaluation guidance to produce per pupil costs over three years. Specific items to be considered will include time for teacher training and preparation (but not delivery as this takes place during normal lessons) and costs of learning materials including the materials and tools which are provided to each intervention school at the start of the trial.

Ethics and registration

Schools will receive a Participant Information Sheet and Consent Form that must be sent to the parents/carers of all pupils taking part in the trial before their data is shared with the evaluation team. This will be sent to schools once the pupils have started reception at the trial schools (in September 2023). It will contain details about the intervention and the evaluation, a statement on data protection along with links to further documentation on data sharing and other relevant matters, and a slip for parents/carers to return to the school if they wish for their child to not take part in the evaluation.

The evaluation was approved by the SHU ethics committee on 12 January 2023 (Ethic Review ID: ER50653704). The trial is publicly registered: ISRCTN96349771.

Data protection[7]

SHU and UoN are both Data Controllers for the Counting Collections evaluation. The processing of personal data through the Counting Collections evaluation is defined under GDPR as a specific task in the public interest. As data is being processed for the purpose of academic research, the main aim of which is to improve numeracy among school pupils, the legal basis for processing your personal data is as a 'Public Task' (Article 6 (1) (e)). https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/public-task/

Special category data, specifically pupil FSM status, will be accessed from the National Pupil Database using Unique Pupil Numbers provided by schools and processed for the purpose of scientific research as permitted under GDPR Article 9 (j). We are processing this data to determine if the Counting Collections programme has different effects on different pupil groups. EEF was established with a remit to break the link between family background and educational attainment, and all EEF projects conduct subgroup analysis on FSM pupils.

[7] Please see the Data Protection Statement for EEF Evaluations.

After the evaluation with EEF is complete, SHU and UoN will retain participants' data for research and knowledge-exchange purposes, including presentations at professional or academic conferences, or publications in professional or academic journals, for up to 25 years and for a period of no less than 7 years after the research project finishes. Data will be stored in the EEF data archive after publication of the final report. The project privacy notice contains details of all data to be processed: https://www.shu.ac.uk/-/media/home/research/sioe-rke/privacy-notices/eef-counting-collections-privacy-notice.docx

SHU's privacy notice provides full information of policies and procedures in relation to the personal data of research participants. https://www.shu.ac.uk/about-this-website/privacy-policy/privacy-notices/privacy-notice-for-research

Personnel

Evaluation team

Dr Martin Culliney, Senior Research Fellow at SHU, will act as principal investigator and impact evaluation lead. Martin is currently principal investigator for the EEF English Mastery trial and is lead author on the EEF evaluation reports Integrating English (2019) and REACH Primary (2021).

Dr Karen Daniels is Associate Professor at SHU and has significant experience of early years and primary pedagogy. Karen will act as IPE lead, drawing on her extensive experience of qualitative research and as IPE lead in REACH Primary and Ark English Mastery EEF projects.

Joanne Robson is a Senior Lecturer and Early Years Lead at SHU and has significant experience of early years and primary pedagogy, particularly early years mathematics. She will work on the IPE. Joanne was a primary teacher for 23 years, during which time she trained as a Numbers Count teacher leading the intervention across schools.

Sean Demack is Principal Research Fellow and Deputy Head of the SHU research centre. He has extensive experience of educational RCTs. He will act as statistical advisor.

Catherine Hathaway was a teacher for 17 years, including as Head of Early Years of a large school in China. Now a lecturer at SHU, Catherine teaches Early Maths and Early Years trainees as a maths specialist and will contribute to this trial as a fieldworker.

Jacqueline Launders taught in primary schools for 12 years, with roles as Maths Lead and senior leader. Now a Senior Lecturer in Primary and Early Years Education, she teaches on undergraduate and postgraduate QTS courses and will be a fieldworker on this evaluation.

Fiona Storey taught for 15 years, with experience ranging from EYFS to Year 5, and as KS1 maths lead. She is now Senior Lecturer in Primary and Early Years Education specialising in maths. Her role on this study is as a fieldworker.

Development team

Dr Catherine Gripton is an Associate Professor in the School of Education and the Principal Investigator leading the University of Nottingham team. She is an early childhood mathematics education researcher with extensive experience designing and developing mathematics professional development programmes.

Professor Shaaron Ainsworth leads Research and Knowledge Exchange in the School of Education at Nottingham and is the Convenor for the Learning Sciences Research Institute. She has led numerous interventions drawing on theories from cognitive science and has specific expertise in representations and number sense.

Professor Geoffrey Wake, Convenor for the Centre for Research in Mathematics Education, has led many large-scale mathematics education research projects in the UK, including Maths For Life, Centres for Excellence and LeMaPS, as well as conducting research in the learning of mathematics internationally.

Dr Marie Joubert is a Research Fellow in the Centre for Research in Mathematics Education. She has extensive research experience in various aspects of mathematics education and professional development and has led professional development for teachers of mathematics in a range of national and international contexts.

Risks

Risk	Mitigation	Adjusted risk
Evaluator staff departure	Very low turnover.	Low
Recruitment problems	Approach enough schools. Control schools offered payment as is usual in EEF trials.	Low
Difficulties in baseline/outcome test administration	Sufficient number of dedicated administrative staff in regular contact with schools and test administrators. Subcontractor for pupil assessments preparing for high workload by recruiting more testers.	Low
Intervention not delivered	Monitored through compliance data. Teachers encouraged to engage with full training programme, regular contact with developers during delivery period.	Low
School closures/lockdowns	Not expected but if so, schools will be last thing to close as in past. If external visitors are restricted, or schools are closed, fieldwork can be conducted remotely.	Medium
Pupil/teacher attrition	Whole classes take part, so statistical sensitivity more affected by schools than individuals. The delivery period is relatively short so staff departures also lower risk as a result.	Low

Timeline

Table 4: Timeline

Dates	Activity	Staff responsible/ leading
Oct 22	Set-up meetings and IDEA workshop	All
Nov-Dec 22	Ethical approval Draft MoU, consent and information forms Design IPE instruments Evidence review	SHU
Feb 23	Protocol Trial registration	SHU
Feb-Jul 23	Recruitment Data collection from schools	UoN
Jul 22	Randomisation	SHU
Sep-Nov 23	Collect pupil data Baseline testing	SHU
Nov 23	Day 1 teacher training/observations	SHU/UoN/schools
Nov 23- May 24	Intervention delivery	Schools
Nov 23- May 24	Conduct IPE school visits	SHU/schools
Nov 23- May 24	Day 2-5 teacher training/observations	SHU/UoN/schools
Jan 23	SAP	SHU
Mar 23	NPD application	SHU
Jun-Jul 24	Outcome testing Post-intervention teacher survey	SHU/schools
Sep 24	Data analysis	SHU
Jan 25	Report first draft	SHU
Apr 25	EEF to receive the final report	SHU

References

Bloom, H, S., Richburg-Hayes, L. and Black, A. R. (2007) 'Using Covariates to Improve Precision for Studies That Randomize Schools to Evaluate Educational Interventions' *Educational Evaluation and Policy Analysis* 29(1), 30–59

Bruner, J. S. (1966) *Toward a Theory of Instruction*. Cambridge: Harvard University Press.

Chen, J., McCray, J., Adams, M. and Leow, C. (2014) 'A Survey Study of Early Childhood Teachers' Beliefs and Confidence about Teaching Early Math' *Early Childhood Education Journal* 42, 367–377

Coldwell, M. and Maxwell, B. (2018) 'Using evidence-informed logic models to bridge methods in educational evaluation', *Review of Education*, 6(3), 267–300

Dorwick, N. (2015) 'Numbers Count: a large-scale intervention for young children who struggle with mathematics' *The Oxford Handbook of Numerical Cognition*, 1099–1117

Gripton, C. and Pawluch, D. (2021) 'Counting Collections in the Early Years'. *Mathematics Teaching* 275, 6-10

Henry, G. Mashbur, A. and Konold, T. (2007) 'Developing and Evaluating a Measure of Young Children's Attitudes Towards School and Learning' *Journal of Psychoeducational Assessment* 25 (3), 271-284

Miles, M. B., Huberman, A. M. and Saldana, J. (2019) *Qualitative Data Analysis: A Methods Source Book* (4th ed.), London: Sage

Ofsted (2021) *Research review series: mathematics.* Available online at: https://www.gov.uk/government/publications/research-review-series-mathematics/research-review-series-mathematics

Ofsted (2022) *Early Years Research Review: Best start in life Part 1.* Available online at: https://www.gov.uk/government/publications/best-start-in-life-a-research-review-for-early-years/best-start-in-life-part-1-setting-the-scene#teaching

Perry, B., Gervasoni, A., Hampshire, A., and O'Neill, W. (2016) 'Let's Count: Improving Community Approaches to Early Years Mathematics Learning, Teaching and Dispositions through Noticing, Exploring and Talking about Mathematics' in B. White, M. Chinnappan, and S. Trenholm (Eds.) *Opening up mathematics education research: Proceedings of the 39th annual conference of the Mathematics Education Research Group of Australasia* (75-82). Mathematics Education Research Group of Australasia.

Torgerson, C., Wiggins, A., Torgerson, D.J., Ainsworth, H., Barmby, P., Hewitt, C., Jones, K., Hendry, V., Askew, M., Bland, M., Coe, R., Higgins, S., Hodgen, J., Hulme, C., Tymms, P. (2011) Every Child Counts: The independent evaluation, Technical Report.

Appendix

Figure 1: Counting Collections Logic Model