Department for Education

# Adult numeracy randomised controlled trials: Programme report

May 2026

Authors: Susan Mackay, Naomi Williamson, Jamie Roberts, Madeleine Thornton, Dimitris Raidos, Akshay Choudhary, Megan Gardiner, Nicola De Waal

Government Social Research

# Contents

# List of figures

# List of tables

# Acknowledgements

This report would not have been possible without the valuable contributions of many individuals and organisations.

First and foremost, we extend our heartfelt gratitude to the trial participants who generously gave their time and openly shared their experiences. Your insights have been instrumental in shaping our understanding and have enriched the findings immeasurably.

We are deeply appreciative of the adult education providers and tutors who facilitated the delivery of interventions, collected and shared learner data, and offered valuable professional perspectives. Their dedication and cooperation were fundamental to the success of the programme.

Our sincere thanks go to the product developers whose sector expertise and support in intervention development and tutor training was invaluable: the Education and Training Foundation, Mathematics in Education, National Numeracy, Campaign for Learning and the Shannon Trust.

We gratefully acknowledge Etio, the Managed Service Supplier, for their work in recruiting and managing providers and schools for the trials. We also thank the Kicktag team for developing the Ipsos Data Portal, which streamlined our data collection processes.

This work has been greatly enhanced through collaboration with our evaluation consortium partners – the Institute for Employment Studies, Learning and Work Institute, King's College London, and RAND Europe – whose work on the individual trials has been exemplary.

We would also like to thank Professor Alex Sutherland as Independent Technical Advisor, members of the Technical Steering Group and the Cabinet Office Evaluation Taskforce for their advice, support and challenge throughout.

Finally, we express our appreciation to the project team at the Department for Education team, whose thoughtful guidance and constructive feedback have been invaluable throughout this programme of work.

# Executive summary

In 2023, Ipsos UK was appointed by the Department for Education (DfE) to lead a consortium[1] to deliver the evaluation elements of the Adult Numeracy Trials, delivered as part of the government's programme for improving adult numeracy (Multiply).

The aims of the Adult Numeracy Trials were to:

- generate robust, high-quality evidence on the impact of specific interventions designed to engage, motivate and teach essential maths skills to adults
- understand the feasibility, opportunities and challenges of implementing trials within the adult education sector
- support broader efforts to ensure value for money in adult education, by identifying what works.

The adult education sector in England is relatively small, with around 2,000 providers compared to 17,000 primary schools and 3,500 secondary schools and has limited prior experience of Randomised Controlled Trials (RCTs). Despite this, the programme successfully delivered 1 communication-based intervention as a full-scale effectiveness trial (Encouraging Progression) and recruited 96 providers, 72 schools and over 2,700 learners to 4 smaller pilot RCTs (Family Numeracy, Adapted Mastery Approach, Contextualised Approach and Preparation for Maths GCSE) and one Implementation and Process Evaluation (Embedded Maths).

This report presents programme-level findings, detailing the trial delivery processes and synthesising key findings. It is designed to inform any policymakers and researchers who may be planning future large-scale research and trials. The Multiply programme was subject to a separate evaluation, so this report does not assess the overall performance of the Multiply programme.

## Key findings

The Adult Numeracy Trials demonstrated that RCTs are feasible in the adult education sector, with providers willing to participate despite their limited prior experience and knowledge of RCTs. The ambitious programme successfully delivered 1 full-scale effectiveness trial and recruited large numbers of providers, schools and learners across 4 pilot trials and one Implementation and Process Evaluation (IPE).

[1] The consortium comprised Ipsos UK, the Institute for Employment Studies, Learning and Work Institute, King's College London and RAND Europe.

The pilot trials were only powered to identify medium to large impacts of an intervention, should these exist. Nevertheless, across all 6 trials, statistically significant positive impacts were identified for 2 interventions: Family Numeracy and Encouraging Progression. Family Numeracy was designed to encourage adult learners to take their first steps into formal numeracy learning through numeracy-based Family Learning. Encouraging Progression aimed to help learners progress from a short Multiply course to a regulated numeracy course, with positive impacts detected specifically for learners aged 50+[2].

Of the remaining 4 course-based trials, which tested different approaches to teaching maths to adults, one showed promise. Adapted Mastery had a positive impact, although results were not statistically significant at the 95% level[3]. Two of the trials were not able to identify signs of positive impact (Contextualised Approach and Preparation for Maths GCSE), with results comparable to business-as-usual teaching methods for maths Functional Skills Qualification (FSQ) Level 1 and GCSE courses respectively. This suggests either no positive impact exists for these interventions, or that any impact was too small to detect in the pilot trials. The fourth trial, Embedded Maths, faced significant implementation challenges. High tutor attrition after randomisation was attributed to tutor anxiety about teaching maths content, alongside the volume of training required. This intervention would therefore require changes to resolve these issues.

The trials also generated valuable methodological learning. The programme provides evidence that interventions can be implemented with good fidelity and positive learner and tutor experiences in the sector. In 5 of the 6 trials (excluding Embedded Maths), tutors and providers were willing and able to work within the constraints of an RCT. They collected and submitted substantial data in accordance with Data Processing Agreements (DPAs) and did not change their behaviour as a result of taking part in a trial. However, the Embedded Maths trial revealed different provider behaviour. Providers enrolled learners in additional maths courses alongside the trial course; 1 provider indicated this was to maximise gains in learners' maths confidence and competence. Although this contaminated the trial results, it provided valuable insight into provider motivations and behaviour.

---

[2] Two letters were tested and both significantly improved progression to qualification-bearing maths courses for learners aged 50+, compared to minimal or negative effects for younger learners. The differential impact between age groups was statistically significant at the 95% confidence level (p=0.038 for peer support letter; p=0.006 for urgent letter), suggesting that encouragement strategies are particularly effective for older adult learners.

[3] For the composite measure of confidence in maths, the estimated treatment effect was positive, and approached statistical significance at the 95% level (d=.27, p=.066).The results of the individual questions in the composite score show that the positive effect was entirely driven by responses to 2 of the 6 questions The small sample size at endline significantly reduced the effective sample size used in the analysis model. This reduction in sample size suggests a substantial loss of statistical power, meaning that the results are highly sensitive to variation.

Whilst RCTs are used in evaluation to understand the impact of specific interventions, they can be challenging to design and deliver and are not always the most appropriate methodology. Central government guidance acknowledges that RCTs work best where "a) reasonable sample sizes can be constructed to allow for tests to be carried out on data which have sufficient statistical power; b) the randomisation for the RCT can be feasibly and practically integrated into the intervention design before it is implemented; c) it can be confidently assumed that the intervention has no impact on the control group and d) restricting the intervention is appropriate and does not cause undue ethical risks"[4]

The Adult Numeracy Trials also highlighted delivery challenges that needed to be understood and overcome to run successful RCTs in the adult education sector. These challenges included: provider recruitment and retention to ensure trials were sufficiently powered to detect smaller impacts; provider engagement to ensure understanding of the purpose and requirements of RCTs and to improve fidelity and avoid contamination; and a range of data quality and matching issues when using the DfE Individual Learning Record (ILR)[5] administrative dataset for research purposes. The lessons learned from addressing these challenges, detailed in this report, provide a foundation for future trials in this sector.

[4] The Magenta Book - GOV.UK The Magenta Book, authored by HM Treasury, provides guidance on what to consider when designing an evaluation. Further information on when an RCT could be a feasible evaluation method can be found in Table 2.3 of the Magenta Book

[5] The ILR is an on-going collection of data about learners from training providers in the Further Education (FE) and Skills sector in England.

## Overview of adult numeracy trials

### Figure 1: Overview of Adult Numeracy Trials

| | Trial Name | Intervention | Treatment/Control Group | Findings | Cost | Evidence rating |
|---|---|---|---|---|---|---|
|  | **Encouraging Progression** *(full-scale effectiveness RCT)* | DfE sent 2 different letters encouraging sign up onto a regulated (qualification-bearing) course | **Treatment: Multiply learners who completed a non-regulated course, randomly selected to receive either letter A or letter B.** Control: As target audience, randomly selected to receive no letter. *All participants, whether on treatment or control, were subject to adult education advertising as usual.* | No statistically significant impact on learner progression overall. However, it showed positive and statistically significant effects for learners aged 50 and over | £ | ★★★ |
|  | **Family Numeracy** *(Small, pilot RCT)* | 6 x 60 minute sessions taught to parents alone (first 30 minutes) and then parents + their Year 2 child (30 minutes). The courses were held in primary schools and delivered by local adult education tutors. | **Treatment: parents of Year 2 children, who either do not have a Level 2 maths qualification or who expressed low confidence in maths.** Randomly selected (at school level) to receive the intervention. Control: As target audience, randomly selected (at school level) to **not** receive the intervention at the same time as those in the treatment group. They received the intervention once the trial had concluded. *All participants, whether on treatment or control, were subject to adult education advertising as usual.* | Statistically significant positive effect found on parents' and carers' progression to further maths learning | £££ | ★★☆ |
|  | **Adapted Mastery** *(Small, pilot RCT)* | 22 hours of maths Functional Skills Qualification (FSQ) L1 course content, taught using maths mastery approaches to teaching, adapted to adult learners. | **Treatment: maths FSQ Level 1 learners**, randomly selected (at provider level) to have their course taught using mastery approaches. Control: Same as the target audience, randomly selected (at provider level) to have their course taught using 'business as usual' methods | No statistically significant improvements in maths attainment, so if this exists it is likely to be small. Improvement in maths confidence were inconclusive, but showed promise. | ££ | ★★☆ |
|  | **Contextualised Approach** *(Small, pilot RCT)* | 18 hours of maths Functional Skills Qualification (FSQ) L1 course content, taught using or contextualised approaches to teaching | **Treatment: Maths FSQ Level 1 learners**, randomly selected (at provider level) to have their course taught using contextualised approaches. Control: Same as the target audience, randomly selected (at provider level) to have their course taught using 'business as usual' methods | No statistically significant improvements in maths attainment or confidence identified, meaning if these exist they are likely to be small. | ££ | ★★☆ |
|  | **Preparation for Maths GCSE** *(Small, pilot RCT)* | 7-hour maths study skills course, aligned with Maths GCSE curriculum content and taught at the start (5 hours) and end (2 hours) of the course | **Treatment: Maths GCSE learners**, randomly selected (at provider level) to have study skills incorporated into their course Control: same as the target audience, randomly selected (at provider level) to have their course taught using 'business as usual' methods | No statistically significant improvements in maths attainment identified, meaning if these exist they are likely to be small. | £ | ★★☆ |
|  | **Embedded Maths** *(Implementation and Process Evaluation)* | Re-working of the Level 2 Health and Social Care (H&SC) curriculum to also incorporate 18 hours of H&SC relevant FSQ Level 1 maths content (approx. 60%) | **Treatment: Health and Social Care Level 2 learners,** randomly selected (at provider level) to receive extra maths content in their course Control: same as the target audience, randomly selected (at provider level) to have their course taught using 'business as usual' methods | No impact findings, given the qualitative nature of this trial | ££ | ★☆☆ |

Source: DfE summary of trial findings. Cost ratings are qualitative judgements, based on the number of training hours needed, administrative time required to support delivery and other costs eg printing and postage). Evidence ratings are judgements based on the Nesta Standards of Evidence. This summarises information found in Trial findings

## Lessons learned

### Programme governance

DfE held overall responsibility for delivering the programme of Adult Numeracy Trials and contracted Product Developers, Evaluators, and the Managed Service Supplier (MSS). Collaboration between the DfE, MSS, Evaluators, providers and schools was facilitated through clear roles and responsibilities, a collaboration schedule developed at the initial contracting stage[6], regular partner meetings (all-partner, joint delivery and evaluator meetings), and engagement with adult education providers and schools.

Programme governance was facilitated by effective, flexible and collaborative partnership working, which was reviewed through quarterly contract review meetings. It included careful risk management, drawing on consortium expertise to develop solutions to challenges that arose. Purposefully developed cloud-based secure project management tools were used, which ensured that all consortium members had access to the latest programme information. This enabled productive discussions and created an open environment for resolving challenges and complexities in trial delivery.

However, challenges arose from the programme's scale and complexity. This was illustrated by the challenges of deciding whether trials with lower than expected recruitment or higher than expected attrition should be scaled down or stopped. These decisions needed to consider the impact on all three programme objectives[7], and led to the Embedded Maths trial continuing as an IPE. Limited sector experience of running RCTs, and short delivery timescales also posed challenges. Short delivery timescales were exacerbated by programme funding being based on financial years, which did not align with the academic year cycle for providers of administrative data collection. This meant that fieldwork delivery for 5 of the trials (all except Encouraging Progression), data collection, analysis and reporting all had to be completed in the final financial year of the programme. Running these project stages in parallel increased complexity and introduced additional risks to manage.

### Intervention development

The DfE, through the Multiply Programme Board, oversaw intervention identification and selection, with some oversight from HM Treasury and the Cabinet Office Evaluation Taskforce. A call for ideas was issued to the sector and DfE policy teams, while a

[6] The collaboration schedule formally set out how Ipsos and the MSS would work together. It detailed key principles and KPIs, as well as practical commitments like attending regular meetings, using common terminology, and sharing information openly and quickly.

[7] These were to: 1) generate robust, high-quality evidence on the impact of specific interventions designed to engage, motivate and teach essential maths skills to adults; 2) understand the feasibility, opportunities and challenges of implementing trials within the adult education sector and 3) support broader efforts to ensure value for money in adult education, by identifying what works.

systematic review[8] of the existing evidence base was undertaken to identify promising practice. The ideas submitted were developed further by DfE researchers using the PICO framework[9], in collaboration with those who suggested them. Those considered unsuitable for an RCT were ruled out. Where needed, specialist product developers were commissioned to turn the selected ideas into tangible, ready to deliver, interventions. The evaluators then conducted Trial Feasibility Assessments on these ideas, developed, refined and tested (where appropriate) Theories of Change, and produced initial power calculations to ensure interventions were 'trial-ready'. Stop/go gateways were added during the intervention development stage to try and ensure that only promising interventions went through to fieldwork.

The sector-wide approach to identifying interventions generated a wide range of ideas across all stages of the learner journey, from initial engagement and motivation through to pedagogical approaches and learner progression. Specialist product developers enhanced trial credibility through their expertise and co-design input. The Trial Feasibility Assessment process, combined with strong collaboration between product developers, evaluators and DfE ensured the selection of suitable interventions.

However, generating a series of interventions from a limited evidence base and developing them to be trial ready was a significant challenge in the timeframe for the programme. Given the high value placed on being able to innovate to identify and evidence successful interventions, explore the viability of RCTs in adult education and inform future policy, the programme accepted this challenge, whilst recognising that the innovative nature of the interventions and the limitations of the evidence base supporting some of them would mean that the early stages of set up would need to be monitored closely, and options to scale back trials kept open.

This was a particular issue for interventions focussed on engaging and motivating potential learners through peer support. Theory-based evaluation was undertaken to test these interventions further, and none progressed to trial. Whilst the theory of change for Embedded Maths was valid, a theory-based evaluation before the trial might have provided additional useful information on the potential level of maths anxiety that non-maths teachers would face when teaching Level 1 maths. However, the remaining interventions performed well in the field, with good fidelity, minimal changes required, and positive feedback on materials from both tutors and learners. This shows that trial ready

[8] Review of the evidence on numeracy skills interventions for adults - GOV.UK

[9] The concept of PICO was introduced in 1995 by Richardson et al. to break down clinical questions into searchable keywords. It is a mnemonic formula originally developed to help investigators frame research questions when designing a study. P: Population/Patient/Problem; I: Intervention; C: Comparison/Control and O: Outcome

interventions can be developed within shorter timeframes, with the right level of support and a high risk appetite from funders, aligned to the innovative nature of the project.

There were also trade-offs between intervention testing and refinement and the timeframe for recruitment. In the course-based trials, the interventions were still being refined after recruitment had started, which meant providers received detailed descriptions of the intervention only after committing to the trial. It is possible that this dampened recruitment efforts and contributed to some of the attrition these trials faced.

### Trial design

The trial design stage incorporated sampling, methods selection, and the development of Trial Protocols and Statistical Analysis Plans (SAPs). Due to the relatively small number of adult education providers, the programme adopted a centrally managed, integrated sampling strategy. The Managed Service Supplier (MSS) maintained a sampling frame based on ILR data, tracked participation across trials, and capped provider involvement at 3 trials. This approach reduced the potential for contamination between trials while minimising provider burden.

For each feasible intervention, evaluators developed a range of design options: IPEs, Pilot RCTs, full RCTs and quasi-experimental designs (QEDs). Trial protocols underwent a 4-stage review process to prespecify design elements, and Statistical Analysis Plans (SAPs) were finalised before any outcome data analysis began. The consortium established a common analytical framework covering descriptive statistics, ILR variables and coding, outcome standardisation, missing data handling, subgroup specification, effect size interpretation, and robustness checks.

The consortium's trial expertise combined with independent scrutiny from the Technical Steering Group and Technical Advisor enhanced technical rigour, refined power assumptions, and strengthened analytic plans. Preparing multiple design options upfront enabled the DfE to make rapid decisions on whether the scale of the trial should change based on real-time recruitment data. The consortium established consistent analytical approaches through collaborative development of a comprehensive analysis matrix, whilst maintaining flexibility to address unexpected findings (such as complete separation in the Family Numeracy trial) to ensure both rigour and applicability of results. Robust quality assurance was embedded throughout the analysis and reporting stages, with multiple review checkpoints, standardised reporting templates, and tools including summary presentations and an IPE Matrix to facilitate cross-trial comparisons and learning across the programme.

However, there were several challenges faced in trial design. The lack of prior benchmarks for recruitment, attrition and compliance in adult education trials meant

greater reliance on informed assumptions, though these trials now provide valuable reference points for future sector research.

Most significantly, for the course-based trials, the relatively small provider base (c. 2,000 providers, of varying sizes) combined with the desire to run multiple trials simultaneously covering different aspects of the learner journey limited trial scale. This led to pilot RCTs being chosen over full efficacy trials for the course-based trials. The challenge was compounded by cluster randomisation designs, at either a year group or provider level, because individual randomisation was not practical – given it is not possible to teach half a class one way and the other half another with only one teacher and one scheduled class. Many providers only offered one class per course for adult learners, making class-level randomisation feasible only with the largest providers, which would introduce high levels of sample bias and potentially exacerbate the impact of cluster randomisation[10].

Both individual and class-level randomisation within the same provider also carried risks of contamination bias, with learners and teachers sharing information about the treatment intervention. These issues have traditionally meant that RCTs have not been considered possible in the adult education sector. The programme of Adult Numeracy Trials demonstrates otherwise: robust trials designed to detect medium-large effects can be delivered through class and provider level randomisation.

### Recruitment

The programme of Adult Numeracy Trials involved recruitment of providers, schools and adult learners across the pilot course-based trials and the pilot Family Numeracy trial. The Managed Service Supplier (MSS) led recruitment of providers and schools on behalf of the DfE, using a range of strategies. A total of 289 expressions of interest (including sign-ups) were received from 166 separate providers for the course-based trials[11]. Of these, 96 providers signed contracts and were formally recruited, a conversion rate of 58%. For the Family Numeracy trial, 293 primary state schools expressed an interest, with 72 participating in the trial, a conversion rate of 25%. The need to geographically match adult education providers with primary schools contributed to the lower conversion rate for the Family Numeracy trial.

A total of 2,760 eligible learners were recruited by the providers and schools taking part in the course-based trials, representing additional enrolments funded through the RCT

[10] The Intracluster Correlation Coefficient (ICC) measures how similar participants are within the same cluster, compared to participants in different clusters. An ICC greater than 0.10 indicates strong clusters effects. For example, an ICC of 0.07 means that 7% of the variance in outcomes is due to differences between clusters, and 93% is due to differences between individuals. The Adult Numeracy Trials that relied on provider and school level randomisation (Adapted Mastery, Contextualised Approach, Preparation for Maths GCSE and Family Numeracy) were designed with ICCs ranging from 0.06 – 0.2, but ended up with ICCs in the range of 0.13 – 0.42

[11] (FSQ [143], Preparation for Maths GCSE [91] and Embedded Maths [55]).

programme. A further 13,647 existing Multiply learners took part in the Encouraging Progression trial.

Providers and tutors expressed interest in taking part in future trials and research, recognising the importance of participation to improving delivery of courses across the sector. However, providers emphasised that trials must be cost-effective, with funding covering delivery costs and staff time, and interventions fitting within existing curricula and course cycles. Tutors were willing to participate provided additional time requirements were minimal – a particular concern for the sector's predominantly part-time workforce.

Learners who knew about their trial participation were generally positive, but awareness varied considerably. Some learners were unaware they were in a trial, potentially due to the opt-out consent process, which likely also contributed to low survey response rates. Future trials should ensure clear communication about participation and consider how consent approaches affect learner awareness and engagement.

Recruitment to the course-based trials benefited from a multi-channel communication strategy: initial outreach through local authorities and sector bodies, followed by targeted MSS communications. Incentives were critical to converting interest into sign-ups. However, there were barriers to recruitment including compressed timelines and providers' limited awareness of trials leading to hesitancy to take part. Provider recruitment was further hampered by the relatively small size of the adult education sector, limited staff capacity, and a need for more in-depth 'market warming' to raise awareness of what taking part in trials would involve.

### Attrition

Most withdrawals occurred before trials started, with fewer instances after randomisation and during delivery. Nonetheless, the Intention to Treat (ITT)[12] design of the Adult Numeracy Trials meant that attrition post-randomisation reduced the overall power of the trials. This was a particular issue for the Embedded Maths trial, which had 80% post-randomisation attrition and was thus completed as an IPE. The 2 trials looking at the impact of interventions on Level 1 maths FSQ pass rates also had relatively high levels of post-randomisation attrition, with 42% attrition for the Contextualised Approach and 34% attrition for Adapted Mastery Approach. The remainder of the trials had attrition rates of around 20%. This will be valuable information for those undertaking trials in this sector in the future, as it will allow estimates for the level of post-randomisation attrition to be built into the trial design.

[12] Intention-to-treat (ITT) analysis includes all participants in the groups to which they were randomly assigned, regardless of whether they completed the intervention, dropped out, or deviated from the protocol. This approach preserves the benefits of randomisation and provides a pragmatic estimate of intervention effectiveness under real-world conditions.

Attrition occurred due to both cross-trial and trial-specific factors. Research was undertaken with providers who withdrew at different stages to better understand the barriers to undertaking RCTs in this sector. Internal provider challenges included insufficient teaching capacity and high tutor turnover. These issues were compounded by sector-wide maths tutor shortages and the prevalence of part-time work patterns, which significantly hindered trial participation. This was particularly problematic for trials requiring ongoing tutor training, which tutors had to attend in addition to their existing teaching commitments. External challenges included the sector's limited experience with RCTs as a research method and limited time in which to undertake recruitment, especially for Family Numeracy, where recruitment was a multi-stage process. To mitigate these issues in future trials, providers should receive trial information earlier and more clearly. Additionally, incentives for trial participation should be sufficient to make it financially viable for providers.

### Primary data collection

Whilst primary outcomes were measured using DfE administrative data from the ILR where possible, the Adult Numeracy Trials also required primary data collection to supplement this and gather information for the IPE. The primary data collection included management information (MI) on tutor training and learner attendance, plus surveys, interviews, maths assessments and observations, with methods varying based on intervention type and practical constraints.

Key success factors for ensuring complete and high-quality data included automated quality checks in the online data portal where data was entered, flexible and multi-mode survey administration (online and telephone), provider support for data submission, and strong collaboration between the evaluators, MSS and the DfE to identify and address data gaps. The trials relied on public task as the lawful basis for the sharing of personal data[13], with learners able to continue with their courses even if they did not want their data to be used in the trials. Course-based trials used an 'opt-out' approach, while Family Numeracy used opt-in consent, and participation in all primary research was optional.

Data collection challenges in the course-based trials included adult learners being able to join FSQ and GCSE courses up to 6 weeks after the start of the course. To complete baseline surveys and assessments, the evaluation team needed real-time enrolment information that was not available through the ILR data return cycle and instead relied on

[13] Personal data were processed under the UK GDPR lawful basis of public task (Article 6(1)(e)), as the research was necessary to support the Department's statutory functions and policy responsibilities. Consent was sought from participants via an opt-out process for the Adapted Mastery, Contextualised Approach, Preparation for GCSE and Embedded Maths trials, and via an opt in process for the Family Numeracy Trial. The seeking of consent was undertaken as an ethical and transparency measure to inform and respect participants' choices, rather than as the lawful basis for processing personal data.

providers updating trial-specific management information promptly. Other challenges included late data submissions from providers and confusion about which data to submit. To address these issues, the evaluation team provided one-to-one support to providers facing data-related challenges and adopted a flexible approach to baseline survey distribution and interview scheduling.

### Data management and linking

Ipsos, as the central data coordinator, was responsible for data management and linking. The trials required linking learner data from 3 sources: MI from providers, baseline and endline survey data, and ILR data.

Success factors included early access to ILR data extracts, which enabled preparation for linking, effective data standardisation and a hierarchical matching process to maximise the match rate. However, challenges arose from inconsistent ILR data submitted by providers, the use of temporary or inconsistent learner IDs and duplicate records. There were also significant levels of missing FSQ outcome data in the ILR dataset, which affected the analysis for the Adapted Mastery and Contextualised Approach trials.

## Recommendations

This report highlights key areas of success for the programme, as well as areas requiring further development or improvement. The following recommendations draw on lessons identified throughout the report and aim to support future RCT delivery in the adult education sector.

### Recommendations for trialists

- To maximise recruitment, future research and trials within the adult education sector should use a multi-channel approach: broad outreach via local authorities and sector representative bodies, followed by direct recruitment with individual providers and schools. This requires sufficient lead time (ideally a full year), adequate financial and non-financial incentives, and clear, regular, communication of trial requirements. The Adult Numeracy Trials found success in using videos and animations to explain what RCTs are and how they work.
- The data portal used to collect and quality assure management information in real-time successfully delivered high quality data with minimal gaps. To reduce burden and maximise data quality in future trials using a split evaluator/Managed Service Supplier commissioning model, primary data collection should use a single data portal aligned with intervention milestones. Real-time monitoring, automated data checks, and provider support should be available throughout. This centralised

approach eliminates the need for individual data sharing agreements with providers, while generating higher-quality, consistent data.

- Future RCTs using data portals should include a testing period before launch to ensure functionality and verify that automated checks are working properly.
- Intervention selection for future trials should balance the potential to detect impact with delivery complexity. Training requirements should be kept manageable, ideally under 10 hours, whilst ensuring the intervention differs sufficiently from standard practice to enable meaningful impact detection.
- Administrative data should take priority over survey data for measuring trial outcomes, despite its limitations. Interim admin data can test the data matching and analysis process and address issues where possible.
- Trial implementation support should include regular contact between evaluators and delivery staff, enabling teams to monitor fidelity and address challenges early.

### Recommendations for DfE

- Complex, multi-partner trials require flexible and collaborative governance. Clear contractual and working arrangements between the DfE, evaluators, product developers and the MSS should establish strong communication channels. These enable productive discussions and create an open environment for resolving delivery challenges and complexities.
- The processes and materials used to identify and select interventions were found to be valuable, including the Template for Intervention Description and Replication (TIDieR) checklist and feasibility assessment criteria. These tools facilitated decision making and could be adapted for research beyond RCTs. The selection process took longer than anticipated because the interventions needed significant development before they were ready to test. For future trials with limited timeframes, it is recommended to either use interventions that are already trial-ready, allowing more time for recruitment and trial design, or building in additional time for intervention development, including theory-based evaluation where required.
- This programme successfully brought together teams with diverse trial experience and incorporated independent experts to review trial designs, provide guidance on complex challenges and ensure quality. This approach proved beneficial and should be adopted in future adult education trials. During the design phase of future trials within the adult education sector, initial scoping of provider circumstances should ensure trial requirements do not create unnecessary burden. Scoping should explore delivery approaches, teaching capacity, training requirements, and cost estimations that include dedicated staff resources.

- Running multiple trials simultaneously prevented providers from participating in all studies, should they have wanted to, which reduced sample sizes and limited statistical power. Where feasible, future trials should be staggered to allow broader provider participation and improve statistical power.

## Conclusions

The Adult Numeracy Trials demonstrated that RCTs can be conducted at scale in the sector, generating valuable insights into a range of interventions that address some of the evidence gaps identified in the systematic review of what works to improve adult numeracy. The programme has increased both capacity and awareness of trials within the adult education sector. Providers and tutors now better understand what trial participation involves and the majority of participating providers (76%) said they would consider taking part in an RCT in the future. The programme successfully built the collaborative and data infrastructure required for delivery at this scale, strengthening sector readiness for this type of research. It established the design, recruitment, and data practices needed for well-powered trials in future, generating practical evidence on implementation presented in this report. Finally, DfE created exemplar governance structures and effectively formalised collaboration between evaluators, the MSS, and product developers, enabling successful trial delivery.

# 1. Introduction

## Introduction to the consortium and trials

Ipsos UK was appointed by the DfE to lead a consortium to deliver the evaluation elements of the Adult Numeracy trials, delivered as part of the government's programme for improving adult numeracy (Multiply). The consortium for this highly ambitious programme of work included the Institute for Employment Studies, Learning and Work Institute, King's College London and RAND Europe. The programme ran from May 2023 until March 2026. During this period, the consortium successfully delivered 6 trials: 1 communication-based intervention as a full-scale effectiveness trial, 4 trials as smaller pilot RCTs and 1 as an Implementation and Process Evaluation (IPE), with course-based interventions delivered during the 2024/25 academic year. This represented a significant achievement given the timeline typically required for an intervention pipeline to be developed, robust trial design, recruitment, delivery and evaluation. Further details are provided in Annex 2.

## Context and background to the trials

The Adult Numeracy Trials were delivered as part of the government's Multiply programme. Through this programme, adult learners across England attended free courses to improve their numeracy skills. This report provides an overview of the RCT programme. This is distinct from the evaluation of the Multiply programme.

At the beginning of the RCT programme, the OECD found that 9 million working age adults in England and Northern Ireland had low basic skills, placing them in the lower half of OECD numeracy rankings, behind comparator countries such as Japan, Germany and Canada (OECD, 2013).[14] Updated data from the Programme for the International Assessment of Adult Competencies (PIACC, OECD) published in 2023 (fieldwork took place between September 2022 and June 2023) found an estimated 8.5 million working age adults in England with low proficiency in literacy, or numeracy, or both. The same report found that England ranked above average in basic skills, but behind Finland, Japan, Sweden, Norway, Netherlands, Estonia, Flemish Community (Belgium), Denmark, Switzerland, Singapore and Germany (OECD, 2024)[15].

[14] OECD (2013), OECD Skills Outlook 2013: Survey of Adult Skills, OECD Skills Studies, OECD Publishing, Paris, https://doi.org/10.1787/9789264204256-en

[15] Updated data from the Programme for the International Assessment of Adult Competencies (PIACC) published in 2023 (fieldwork took place between September 2022 and June 2023) found an estimated 8.5 million working age adults in England with low proficiency in literacy, or numeracy, or both. The same report found that England ranked above average in basic skills, but behind Finland, Japan, Sweden, Norway, Netherlands, Estonia, Flemish Community (Belgium), Denmark, Switzerland, Singapore and

The aim of the Adut Numeracy Trials was to address evidence gaps across the learner journey: from initial engagement, through support for motivation to sign up to, and stay on, a course, to teaching approaches and progression to higher level courses. The core objectives of the RCT programme were to:

- generate robust, high-quality evidence on the impact of specific interventions designed to engage, motivate and teach essential maths skills to adults
- understand the feasibility, opportunities and challenges of implementing trials within the adult education sector
- support broader efforts to ensure value for money in adult education, by identifying what works.

At the design stage, it was decided that the 4 course-based trials would be delivered as smaller pilot RCTs due to insufficient provider recruitment to enable full-scale RCTs.

## Purpose and rationale for the programme report

This programme-level report details how the programme of Adult Numeracy Trials was designed and delivered. It reflects on how the trials were conducted including facilitators and barriers and offers recommendations for future research in adult education. The report draws on a wide range of sources including:

- DfE-generated documentation (such as internal decision-making documentation);
- consortium-generated documentation (such as Trial Feasibility Assessments and research tools for data collection)
- standard trial and research guidance (such as the Template for Intervention Description and Replication (TIDieR) checklist)
- participant-generated qualitative and quantitative data and evidence (from delivery partners and trial participants)
- trial-documentation (such as Protocols and trial reports)

This report has been written as a standalone document, with separate trial-specific reports providing detailed findings and outcomes for each trial. It aims to provide researchers and policymakers with information on both the processes and findings from the programme of work. The trial findings may be most relevant to those working in adult education, whilst learnings about the processes for delivering the trials will be useful for those commissioning and delivering large-scale programmes of research in future.

---

Germany. OECD (2024), Do Adults Have the Skills They Need to Thrive in a Changing World?: Survey of Adult Skills 2023, OECD Skills Studies, OECD Publishing, Paris, https://doi.org/10.1787/b263dc5d-en.

## Structure of the document

The remainder of this document is structured as follows:

### Section 2 Programme governance

This section outlines the consortium structure for delivery and the responsibilities of all organisations involved. The section includes reflections on the strengths and limitations of this governance approach, and implications for future management of large-scale trials and research.

### Section 3 Interventions

This section describes the processes for developing a pipeline of potential interventions to be tested and assessing the feasibility of these for an RCT in adult numeracy settings. The section explores intervention development and costs and reflects on the successes and challenges of selecting and preparing adult numeracy interventions for trialling.

### Section 4 Trial design

This section describes the stages and roles involved in trial design, including the development of trial protocols, analysis and reporting. The section reflects on what worked well and what could have been improved during the trial design stage, with implications for future trials in adult education settings.

### Section 5 Recruitment

This section outlines the strategy for recruiting providers and schools to participate in the trials, including participant numbers. The section reflects on the facilitators, barriers, and implications for recruitment to future trials.

### Section 6 Attrition

This section details the scale and key stages of trial attrition, exploring reasons for provider withdrawal. It provides recommendations for mitigating attrition in future trials.

### Section 7 Primary data collection

This section outlines the types of data collected across the trials, including management information, surveys, interviews and observations. The section discusses successes and areas for improvement in data collection, with recommendations for future trials.

## Section 8 Data management and linking

This section outlines the approach to data management and the process for linking Individualised Learner Record (ILR) data, management information and survey data. It identifies successes, challenges and recommendations for future research involving linked data sources.

## Section 9 Trial findings

This section summarises key findings across the trials, examining the impact findings as well as the implementation and process findings generated by adult numeracy interventions.

## Section 10 Conclusions and recommendations

This section draws together the findings of the report with broader reflections on the programme of work. It outlines lessons learned and recommendations for policy and future trials in adult education and beyond.

A glossary of key terms is included in Annex 1 and further details on the research and trial selection processes, ethics procedures, responsibilities, interventions, data, analysis and costs are included in annexes 2 to 14.

# 2. Programme governance

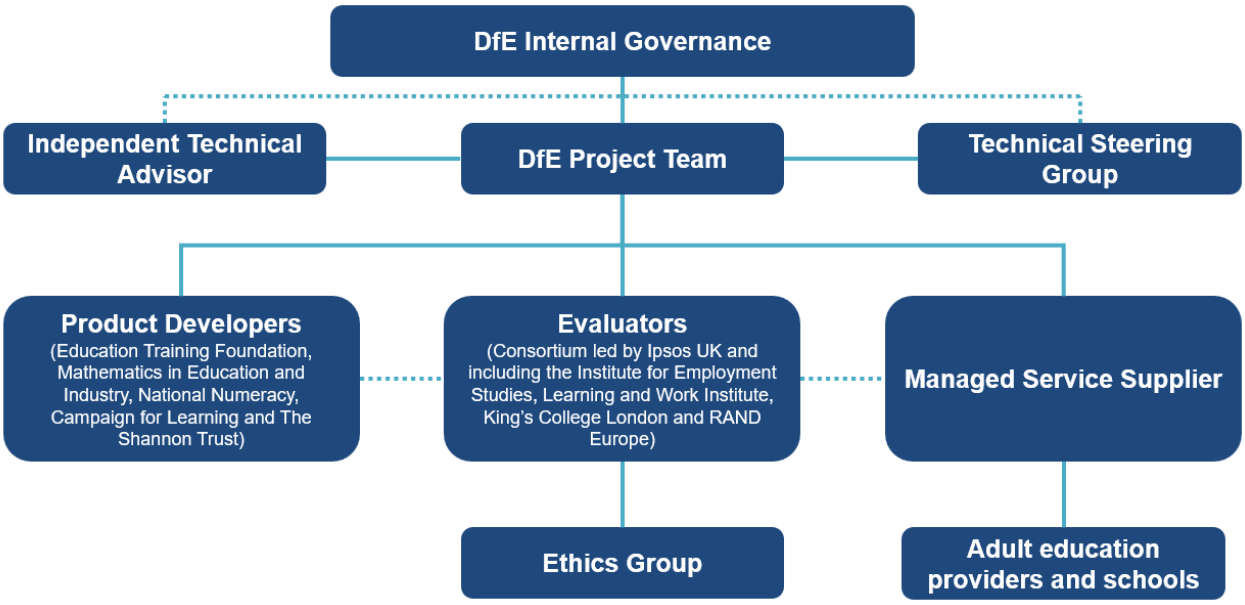
## Summary of chapter

This section reports on the governance arrangements for management and delivery of the Adult Numeracy Trials. It covers internal governance processes within DfE, the respective roles of product developers, Evaluators, the Managed Service Supplier (MSS) and the independent Technical Advisor, as well as wider stakeholders including the Technical Steering Group (TSG) and Ethics Panel. It reflects on how well the programme governance and management arrangements worked and the lessons learned.

## Structures and frameworks

### Overview

**Figure 2: Overview of governance structure for Adult Numeracy Trials**




Source: Ipsos UK

DfE was responsible for overall delivery of the programme of work. They established contracts with:

- **Product developers:** to develop and refine the interventions to be tested through the trials. They included the Education Training Foundation (ETF), Mathematics in

Education and Industry (MEI), National Numeracy, Campaign for Learning (CfL) and the Shannon Trust.

- **Evaluators:** a consortium led by Ipsos UK and including the Institute for Employment Studies, Learning and Work Institute, King's College London and RAND Europe. Evaluators were responsible for the design of the trials, including development and publication of Trial Protocols, data collection, analysis, synthesis and reporting.

- **Managed Service Supplier:** Etio (formerly known as Tribal) was appointed to lead recruitment of adult education providers and schools to the trials, contracting, monitoring of delivery and issuing of payments.

- **Independent technical advisor**: procured from the Cabinet Office Evaluation Taskforce Expert Advisory Panel.

DfE also established a Technical Steering Group (TSG) to provide advice, guidance and oversight of key design elements of the trials. The TSG also had ownership of technical risks and oversight of final outputs and ethical considerations. All trials were subject to ethical review by a dedicated Ethics Panel that was established by Ipsos UK. This included a representative from each organisation within the consortium who had expertise in the ethical issues associated with trials, but no direct involvement in delivery of these.[16]

## DfE internal governance

The Adult Numeracy Trials were overseen by 3 levels of governance: a Delivery Board, a Senior Responsible Officer (SRO) led Steering Group and a TSG. In addition, DfE appointed an independent Technical Advisor to support them in providing sufficient oversight and quality assurance of the trials both formally (through attendance at TSG meetings) and informally (through ad hoc meetings with DfE in an advisory capacity).

These are summarised in Table 1.

[16] The lead organisation was not revealed to the Ethics Panel at the review stage to mitigate any possible vested interest amongst panel members in trials delivered by their organisation.

**Table 1: DfE governance arrangements for Adult Numeracy Trials**

| Level of governance | Remit | Membership | Frequency of meetings |
|---|---|---|---|
| **Delivery board** | Overall responsibility for delivery of the trials against the programme aims. Report into wider DfE and UKSPF governance structures | Board led by DfE and including representation from across Government[17] | Monthly |
| **SRO steering group** | Policy and research-led decision-making body for operational delivery of the programme, including monitoring of issues raised by the Evaluator and Managed Service Supplier. Work with TSG and DfE research team to inform day-to-day decision making and escalate relevant issues to the Delivery Board. | DfE research, policy, finance and commercial teams | Fortnightly |
| **Technical steering group** | Provide scrutiny of trial design and sign off Trial Protocols; own technical risks and issues within the programme; oversight of peer review, analysis and reporting and final outputs. Review ethical considerations associated with medium and high-risk projects. | DfE (adult skills research, adult skills policy and DfE Evaluation Lead[18]), independent Technical Advisor, UKSPF analysts; Cabinet Office Evaluation Taskforce | Varied depending on delivery requirements |

Source: Ipsos and DfE

The governance structures for the Adult Numeracy Trials were more extensive than the standard DfE approach to overseeing research and evaluation projects, reflecting the ambition, technical complexity and scale of the work. They were designed to be flexible to ensure appropriate oversight at each stage (e.g., membership of the SRO Steering Group included commercial and finance representatives during the contracting and set

[17] Membership included representation from the Ministry of Housing, Communities and Local Government (MHCLG), His Majesty's Treasury (HMT), Department for Work and Pensions (DWP), the Ministry of Justice (MoJ) and the Cabinet Office Evaluation Taskforce (ETF) as well as representatives from numerous teams within DfE, including Finance, Commercial, Analytical and the Major Projects Team

[18] The DfE Evaluation Lead was the formal link into the DfE Research Board, the DfE Chief Social Researcher and the DfE Chief Analyst.

up phase, but not during trial feasibility, design and delivery) and to ensure that DfE had access to the required technical expertise to effectively oversee quality of delivery (via the Technical Steering Group and Technical Advisor).These groups made stop/go decisions at gateway stages for each potential trial, agreed thresholds for escalation to the Multiply Delivery Board and ensured programme risks were embedded within the wider Multiply risk register. The RCT programme was also subject to regular reviews by the DfE Major Projects Team.

## Ethics panel

DfE has an established ethics approval process for all research and evaluation studies. Given the scale, profile and complexity of the Adult Numeracy Trials, additional processes were required. The Adult Numeracy Trials involved withholding or delaying access to potentially effective numeracy interventions for some participants. Careful consideration therefore had to be given to ensuring that they were conducted in an ethical manner, both for the protection of participants and the validity of the findings. The evaluation consortium established a dedicated Ethics Panel to ensure a consistent, rigorous and independent approach to reviewing the ethical issues associated with each of the trials. The Terms of Reference for the Ethics Panel is provided in Annex 2.

### Process for ethics review

The role of the Ethics Panel was to review trial-level ethics forms (completed by lead triallists) and provide comments, advice and guidance on ethical issues associated with each trial. The assessment was based on the 6 ethical principles set out in the Government Social Research (GSR) Ethics Guidance[19]. The Ethics Panel reported their feedback from the review process in a template developed for the trials (see Annex 3).

The Panel provided a Red / Amber / Green (RAG) rating to each trial based on whether they considered it to be high, medium or low risk. Trials rated as medium or high risk underwent an additional ethical review by the TSG and DfE.[20] An overview of the steps involved is provided in Annex 3.

[19] GSR 6 ethical principles are: 1. Clear and defined public benefit, 2. Sound application, conduct and interpretation, 3. Data protection regulations, 4. Specific and informed consent, 5. Enabling participation, 6. Minimising personal and social harm.

[20] Two trials underwent an additional ethical review by the Ethics Panel: Family Numeracy and Embedded Maths. Count Me In was also subject to further ethical review by the Ministry of Justice's ethics board due to primary research taking place in prisons.

# Collaboration between suppliers, providers and DfE

This section provides an overview of the approaches used to support collaboration during design and delivery of the Adult Numeracy Trials. It covers collaboration between DfE, the Evaluators and MSS, as well as with adult education providers and schools.

## Clarity on roles and responsibilities

Roles and responsibilities were jointly developed by the Evaluator and MSS and signed off by DfE. The Responsible, Accountable, Consulted, Informed (RACI) framework (see Annex 5) details these roles for each stage of trial delivery from intervention development through to analysis, synthesis and reporting and exit planning. It covered the responsibilities of DfE, Product Developers, Evaluators, the MSS and participating adult education providers. This was useful in clarifying the respective roles and responsibilities of the key stakeholders at each stage of delivery. It was revisited and updated during delivery, for example with exit plans of the Evaluator and MSS.

Collaboration between the DfE, MSS, Evaluators, providers and schools was further facilitated through a collaboration schedule developed at the initial contracting stage. This formally set out how Ipsos and the MSS would work together. It detailed key principles and KPIs, as well as practical commitments like attending regular meetings, using common terminology, and sharing information openly and quickly.

## Programme management

### Tasks and timetable management

Statements of Work were issued by DfE to the Evaluators at approximately every 3-6 months throughout delivery. This approach provided flexibility in delivering the programme of work, which was particularly important during the early trial feasibility and design stages, when it was not yet clear which interventions would be taken forward to trial. It also allowed new activities to be incorporated as they arose.

Smartsheets, an online cloud-based platform, was used to support programme management. It was favoured over other options (e.g., Word / Excel updates, MS Project) because it facilitated collaboration with multiple users across organisations, ensured password protected access and provided real-time updates (for example, the number of providers recruited per trial and the number of providers contracted). The platform supported progress updates between the Evaluators, DfE and the MSS, while also helping to manage tasks, timelines and risks.

### Risk management

The DfE, Evaluators and the MSS undertook an active risk management approach, which included:

- Running detailed risk workshops to identify possible risks and mitigations. This included running a ‘pre-mortem’ for the RCT programme with DfE teams (including policy, research, programme management and commercial) to help develop understanding of the activities to be undertaken, judgements required, dependencies, and the impacts of decisions made in the design stage.
- Aligning programme risks with the wider Multiply programme’s risk and issue management processes, with clearly identified risk owners and escalation routes.
- Developing and agreeing a formal risk appetite with the Multiply Delivery Board that sought to reduce risk aversion.
- Holding weekly meetings with the Evaluators and MSS to review risks and agree mitigating actions, plus quarterly meetings to horizon-scan for upcoming risks.

## Partner meetings

### All-partner workshops

DfE met with the Evaluators for a full-day in-person workshop five times during the three-year delivery of the programme. MSS attended these workshops where relevant. These all-partner workshops facilitated collaboration and extended discussion at key points in scoping, design, delivery and reporting of the trials.

### Joint delivery meetings

DfE met with the Evaluator and the MSS fortnightly throughout trial delivery to discuss progress and issues arising. In the early stages, these meetings focussed on recruitment and contracting, while during delivery they shifted to monitoring compliance with trial requirements, particularly around data collection.

### Evaluator meetings

The DfE held separate weekly meetings with Ipsos, reducing to fortnightly once the trials moved into the analysis and reporting phase. These meetings focussed on progress in delivering individual trials as well as programme-level tasks and activities, such as managing the Ethics Panel and developing templates for use across trials. Quarterly contract review meetings between DfE and Ipsos provided opportunities to reflect on the programme of work, review progress against milestones and KPIs, conduct detailed risk reviews, and develop longer-term plans.

Ipsos coordinated regular consortium meetings to share updates and address emerging issues.

### MSS meetings

DfE additionally met weekly with Etio during key recruitment periods to understand progress and inform decisions about the size and types of trials that would be feasible to deliver.

## Engagement with adult education providers and schools

The MSS was responsible for engagement and recruitment of adult education providers and schools to the Adult Numeracy Trials via a range of activities including attendance at conferences, social media campaigns, articles, phone calls and publishing online articles and social media posts[21]. Ipsos established an exclusive online Learning Platform for providers and schools taking part in the trials. Evaluation partners delivered topical webinars on research and evaluation within adult education settings as an incentive for trial participation.

The MSS was also responsible for contract management of adult education providers. This included Keeping-in-Touch (KIT) calls with each provider once per term, which served as compliance checks to ensure that trial delivery was underway, providers had uploaded data in line with trial requirements, and they had read and disseminated Trial-Readiness Packs (see Annex 6). The MSS also conducted site visits to 10% of providers, which had a similar agenda to the KIT calls but were in person.

# What worked well?

## Partnership working and range of expertise

**Collaboration among all stakeholders involved:** there was effective collaboration between all organisations involved in delivering the trials. This can be attributed in part to the systems established (e.g., Smartsheets, regular meetings, collaboration schedule, flexibility of contracts to quickly expand to include new tasks) but also to the open, transparent and collaborative culture that was set by DfE from the outset, and which guided interactions and ways-of-working between partners.

**Expertise of partners**: the product developers were highly experienced, knowledgeable and well-networked within the adult education sector. This brought credibility to the work and was a key selling point for getting providers and schools on board. The evaluation consortium was made up of leading RCT experts who were

[21] Developed by Ipsos UK and available here: Promotional video

highly experienced in designing and delivering trials in complex settings. This ensured that the trials were informed by deep sector and evaluation expertise.

**Embedding and recognition of the 'test and learn' nature of the programme[22]:** All stakeholders involved were supportive of efforts to capture learning from delivery of the trials. This resulted in open communication, including when faced with challenges in delivery, and collaborative efforts to identify solutions.

### Project management tools and approaches

**Regular communication:** the regular meetings between DfE, Evaluators and the MSS were well structured and planned to make the best use of attendees' time and were revised (duration / frequency) according to need.

**All-partner meetings:** provided valuable touch points to bring key partners together to discuss critical issues and to reach decisions, as well as helping to develop trust and collegiality for the duration of the programme.

**Pre-mortems:** identifying risks and addressing them before issues arose, and establishing an agreed risk appetite with the Delivery Board prior to trial commencement.

**Smartsheets:** proved sufficiently flexible to adapt to the programme's evolving needs and served effectively as a central source of up-to-date information on delivery progress.

## What were the challenges?

- **Scale and complexity:** it was an ambitious programme of work to deliver multiple trials within a single academic year. The wide range of partners involved necessitated comprehensive programme management tools and approaches. The technical requirements contributed to a resource-intensive design stage, with multiple trial options developed for different recruitment scenarios and check-and-challenge provided by the TSG prior to trial designs being finalised and DfE gateway reviews.

- **Sector experience of trialling**: these RCTs were some of the first to be delivered in the adult education sector in the UK. Participating adult education providers therefore needed significant guidance and support to enable them to meet the trial requirements including information webinars, detailed written guidance and ongoing support through dedicated mailboxes for queries.

[22] This is also a reflection of the dual objectives of the Trials to (i) generate high-quality causal evidence on the impact of different interventions on adult numeracy to inform policy in adult skills and (ii) to capture learning on designing and delivering trials within Government and across the adult education sector.

- **Timescales for delivery:** the work had to be delivered in relatively short timescales. This was exacerbated by programme funding being based on financial years, which did not align with the academic year cycle for administrative data collection. This meant that fieldwork delivery for 5 of the trials (all except Encouraging Progression), data collection, analysis and reporting all had to be completed in the final financial year of the programme. Running these project stages in parallel increased complexity and introduced additional risks to manage. This heightened the need for good communication and collaboration between partners.

- **Real-time updates:** the rapid pace and multiple workstreams made it challenging to provide consistent, real-time updates to all partners. The speed of developments, particularly in the recruitment phase, often outpaced reporting. Greater consideration could have been given to the formatting of early progress reports to ensure they were clear and accessible, including to enable updates to senior stakeholders.

# 3. Intervention development


## Summary of chapter

This section reports on the process for identifying interventions for the Adult Numeracy Trials, including those selected for trialling and those that were not. It outlines the selection process, from initial idea generation through to assessment and preparation for trialling. It details the roles and responsibilities of DfE, the Evaluators and product developers at each stage. Finally, it reflects on how effectively the identification and selection process worked and lessons learned.


## Identifying interventions for trialling

A total of 6 interventions were taken forward (Figure 3 provides an overview of the key features of each). The process for identifying potential interventions for trialling comprised 3 strands of activity: conducting a systematic review of what works in adult maths education; engaging with the adult education sector; and gathering insights from DfE policy teams.

### Systematic review of what works[23]

The systematic review published by DfE in 2023 revealed limited evidence on what works in adult numeracy regarding learner engagement, teaching methods, and improving outcomes. Most studies drew inferences from stakeholders' perceptions or completion rates of adult numeracy programmes. They typically related to interventions which were not standardised, meaning they could not be considered for trialling. There were very few examples of robust impact assessments, particularly Randomised Controlled Trials (RCTs).

[23] Alma Economics (2023) Numeracy skills interventions for adults (19+): A systematic review of the evidence. https://dera.ioe.ac.uk/id/eprint/39678/1/Numeracy_skills_interventions_for_adults__19%2B__A_systematic_review_of_the_evidence.pdf

## Figure 3: Interventions tested through the Adult Numeracy Trials

| | Trial Name | Product developer | Stage of the learner journey | Intervention description | Training requirements | Target Population | Primary Outcome | Secondary Outcome | Cost |
|---|---|---|---|---|---|---|---|---|---|
|  | **Encouraging Progression** *(full-scale effectiveness RCT)* | DfE | *After a course:* post-course progression to more maths learning | DfE sent 2 different letters encouraging sign up onto a regulated (qualification-bearing) course | N/A | Multiply learners who completed a non-regulated course | Progression to a qualify-cation-bearing course | Progression to other courses | **£** |
|  | **Family Numeracy** *(Small, pilot RCT)* | Campaign for Learning | *Before a course:* engagement and motivation for maths learning | 6 x 60 minute sessions taught to parents alone (first 30 minutes) and then parents + their Year 2 child (30 minutes). The courses were held in primary schools and delivered by local adult education tutors. | 2.5 hours | Parents of Year 2 children, who either do not have a Level 2 maths qualification or who expressed low confidence in maths. | Parent progression to further maths courses | Progression to other course, maths confidence and child maths attainment | **£££** |
|  | **Adapted Mastery** *(Small, pilot RCT)* | Education & Training Foundation | *During a course:* pedagogy | 22 hours of maths Functional Skills Qualification (FSQ) L1 course content, taught using maths mastery approaches to teaching, adapted to adult learners. | 10 hours initial training plus 24 hours weekly lesson study sessions | Maths FSQ Level 1 learners | FSQ maths pa ss rates | Learner attendance levels, changes in maths confidence and difference in learner maths ability. | **££** |
|  | **Contextualised Approach** *(Small, pilot RCT)* | Education & Training Foundation | *During a course:* pedagogy | 18 hours of maths Functional Skills Qualification (FSQ) L1 course content, taught using or contextualised approaches to teaching | 10 hours initial training plus 19 hours weekly lesson study sessions | Maths FSQ Level 1 learners | FSQ maths pass rates | Maths confidence | **££** |
|  | **Preparation for Maths GCSE** *(Small, pilot RCT)* | MEI | *During a course:* pedagogy | 7-hour maths study skills course, aligned with Maths GCSE curriculum content and taught at the start (5 hours) and end (2 hours) of the course | 2 hours | Maths GCSE learners | Improvement in GCSE maths grades | Pass rates (4+), course completion and confidence in maths. | **£** |
|  | **Embedded Maths** *(Implementation and Process Evaluation)* | Education & Training Foundation | *Before a maths course:* motivation for maths learning | Delivery of 12 x 1.5 hour Level 2 Health and Social Care (H&SC) lessons, re-worked to also incorporate 18 hours of H&SC relevant FSQ Level 1 maths content (approx. 60%) | 10 hours | Health and Social Care Level 2 learners | Maths confidence (learners) | Maths confidence (tutors); learner progression to other maths courses | **££** |

Source: DfE analysis of Adult Numeracy Trial Reports

### Engaging the adult education sector

To identify existing localised interventions that could be tested through RCTs to address these evidence gaps, DfE issued a call to the sector and academia. Engagement was undertaken through existing channels, including Multiply Leads[24] who helped to stimulate demand amongst their network of local providers, DfE's network of Innovator Colleges[25], and relationships with stakeholder bodies in the adult education sector. The DfE research team developed a simple application pro forma ('ideas form') to help practitioners, product developers and others structure their responses. They also offered informal discussions (both in person and online) for those with less developed ideas.

### Gathering insights from DfE policy teams

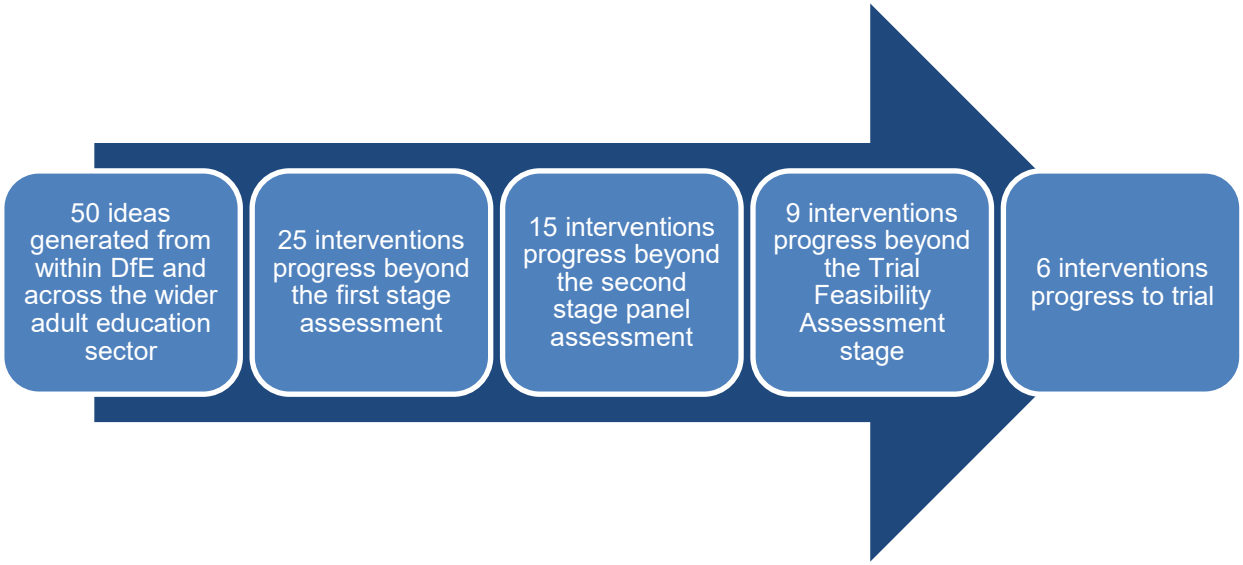
Prior to the launch of the Adult Numeracy Trials, DfE did not have a dedicated policy team focused on pedagogy in adult maths education that could suggest interventions for testing. Therefore, alongside sector engagement, the DfE team issued a call for ideas to policy and research colleagues across the department, drawing on their considerable expertise to build a long list of potential interventions for testing.

## Intervention selection

Figure 4 provides an overview of the key stages involved in the intervention selection process from idea generation through to trialling. The sections that follow provide further detail on what was involved at each stage.

[24] There was a DfE Multiply Lead within all 82 local government areas.
[25] DfE, in collaboration with Innovate UK, funded 33 Innovator Colleges through the Further Education Innovation Fund (FEIF) pilot. They established Local Innovation Centres to catalyse local business growth, particularly by helping SMEs adopt new technologies, business models, and processes.

**Figure 4 Overview of intervention selection**




## Applications and first stage assessment

DfE received 50 ideas from various sources, most commonly from local government and from within the department.[26] This was fewer than anticipated, reflecting the adult education sector's limited experience in researching adult maths pedagogy and conducting experimental research.

These initial ideas underwent a preliminary assessment by the DfE team to determine their suitability for experimental trialling. A total of 25 ideas were taken forward to the second stage[27]. The reasons for not progressing ideas varied and included: research questions being unsuitable for RCTs or already addressed through the national evaluation of Multiply; concepts overlapping with topics already researched via the Adult Education Budget (AEB); and feasibility issues (e.g., target population too small, intervention too long for delivery timelines).

Prior to the second stage assessment, DfE researchers worked with those who had submitted the 25 ideas to develop comprehensive Proofs of Concept, including key design details (e.g., hypotheses to be tested, trial population, primary and secondary outcomes, estimated sample sizes), technical details (e.g., method of randomisation) and estimated costs. DfE also mapped the 25 ideas against different stages in the learner journey (e.g. from communications and engagement interventions through to those aimed at improving maths attainment or supporting learners to progress to further

[26] 50 ideas were submitted: 20 from local areas, 10 from DfE colleagues, 5 from a review of local investment plans, 5 from the Social Policy Association, 2 from charities and 2 from other Government departments.
[27] A total of 36 ideas made it through the first stage assessment, but several were suggested by more than one source or combined with other ideas, resulting in 25 being taken forward.

learning), target cohort (focused on a specific group or all learners) and type of course or delivery model.

## Second stage assessment

The second stage assessment was led by DfE with oversight from HM Treasury and the Cabinet Office. Each idea was assessed by a pool of assessors against a set of criteria covering: the need for the trial, management and delivery, value for money (VfM) and risks (see Annex 7). Each idea was scored on a five-point scale and classified as demonstrating either limited, acceptable or strong evidence.

Over half of the ideas (14 out of 25) were put forward to Trial Feasibility Assessment (discussed below) without further deliberation.[28] A further 4 ideas were automatically ruled out because the mean score across assessors was below 3. The remaining 7 were moderated at panel. A total of 15 were taken forward following conclusion of this process.

At this point, for those interventions that had not been submitted by product developers, DfE issued a call for proposals from developers to refine the 15 ideas into deliverable interventions.

## Cost assessments

DfE conducted cost assessments for each intervention being considered for trialling, using the department's cost modelling approach. This involved working with Adult Education Budget (AEB) colleagues to refine per learner estimates for different types of activity. The cost modelling considered various sample sizes within the agreed budget ceiling. These assessments informed first and second stage decisions, with a small number of high-cost ideas being rejected on financial grounds.

## Trial feasibility assessment

DfE was unable to secure product developers for all 16 ideas, meaning only 9 progressed to the formal Trial Feasibility Assessment (TFA) process (detailed in Annex 8), which was led by the Evaluators.

The Evaluators developed a TFA template and worked with product developers to complete it for each intervention, informing judgements on feasibility (see Annex 9). The assessment examined the operational, technical, scheduling and economic feasibility of trialling each intervention.

[28] Ideas were passed without further deliberation if all assessors gave a passing score (mean of 3 or above in all categories) or all but one assessor gave a passing score and the other assessor's score was at least 2.5 or above.

Each intervention received Red, Amber, Green (RAG) ratings across the 4 assessment categories, which informed an overall rating. The TFA process identified potential challenges or evidence gaps that needed addressing before interventions were ready for trialling. Interventions progressed with amber or green overall ratings progressed to the next stage.

## Developing interventions for trialling

The 9 interventions that were developed for trialling were:

- **Preparation for Maths GCSE:** 7 one-hour lessons for adult learners on Maths GCSE courses covering theories of resilience and growth mindset, study skills, revision, and exam skills (developer: MEI (subcontracted by National Numeracy))
- **Encouraging Progression:** 2 different messaging approaches to motivate adults who had completed a Multiply course that did not lead to a qualification to progress to a maths course that could lead to a qualification (developer: DfE)
- **Embedding Maths in Health and Social Care**: 18 guided learning hours for Health and Social Care students covering approximately 60% of the FSQ Level 1 maths curriculum embedded in their course content to help address barriers to learning (such as elevated levels of maths anxiety) in a supportive and engaging environment (developer: ETF)
- **Contextualised Approach for Functional Skills Qualification (FSQ) Maths Level 1:** 12 weekly 1.5-hour contextualised sessions (testing a Realistic Mathematics Education (RME) approach) delivered to adult Functional Skills Level 1 maths learners by functional skills tutors (developer: ETF)
- **Adapted Mastery Approach (AMA) for FSQ Maths Level 1:** adapting the AMA approach, originally developed for primary and secondary schools, to adult learners (GCSE resit and FSQ Levels 1 and 2 learners aged 16-19-year-old years in further education settings) (developer: ETF)
- **Family Numeracy:** structured, in-person group sessions designed for adults and their children to strengthen numeracy skills through shared family learning (developer: CfL)
- **Count Me In:** Sessions of approx. 20-30 minutes around 3 times a week to enhance numerical literacy of prisoners (developer: Shannon Trust)

- **Personalised Numeracy Assessments:** 1-hour initial assessment of participants' numeracy level and needs and a 45-minute follow up session including signposting to courses (developer: National Numeracy)
- **Confidence Building Workshops:** 1 hour group sessions that aim to improve participants' confidence in maths and signposting to courses (developer: National Numeracy)

## Developing theories of change

The Evaluators and DfE worked with product developers to develop or refine Theories of Change (ToC) for the 9 interventions. These articulated how inputs, activities and outputs would generate outcomes and impacts for learners, how outcomes could be measured, and identified key assumptions, mechanisms and risks. The process involved document reviews, coding interventions using the Template for Intervention Description and Replication (TIDieR) framework, and workshops with the Evaluators to test and strengthen each ToC.

In 3 instances where existing evidence was limited, DfE commissioned additional research to test interventions with learners, examining how they worked in practice and their potential outputs.[29] Despite showing promise for further research, these interventions proved too personalised for trialling (for example, their duration varied depending on the individual learner).

## Power calculations

At this stage, the Evaluators also produced initial power calculations to determine the levels of recruitment (of providers and learners) that would be required to assess the impact of the interventions on intended outcomes (see Section 4: Trial Design for further details). This included an assessment of the total population of eligible providers for each trial, including where these overlapped (i.e. where trials would be recruiting from the same pool of providers).

## Development of intervention delivery materials

Following the trial feasibility assessment stage, DfE commissioned product developers for the 6 interventions progressing to trial to develop delivery materials. These included lessons plans and accompanying resources such as slide decks for use by tutors, along with training programmes to prepare tutors for delivery.

[29] Count Me In, Confidence Building Workshops and Personalised Numeracy Assessments.

## Interventions progressed to trial

The 6 interventions that were progressed to trial (Annex 10) were:

- Preparation for Maths GCSE
- Encouraging Progression
- Embedding Maths in Health and Social Care
- Contextualised Approach for Functional Skills Qualification (FSQ) Maths Level 1
- Adapted Mastery Approach (AMA) for FSQ Maths Level 1
- Family Numeracy

## What worked well?

**The systematic evidence review ensured that the Adult Numeracy Trials did not duplicate existing evidence,** highlighting both ideas for trialling but also gaps in the existing evidence.

**The call to the adult education sector for suggestions of interventions for trialling was successful in generating a range of ideas to be considered and refined.** It also provided DfE with useful insight into the extent to which the sector was developing and experimenting with pedagogical approaches, and its readiness for experimental research.

**Holistic review of interventions** against classic research criteria (e.g., strength of existing evidence base), as well as policy interests (e.g., mapping of the submitted ideas against stages in the learner journey), cost of delivery and ease of implementation criteria ensured a range of relevant interventions with low barriers to entry for the sector were taken forward.

**The involvement of product developers brought relevant skills, specialist knowledge and expertise** to intervention development, bolstering the credibility of the trials within the sector.

**Joint Intellectual Property (IP) arrangements** for interventions co-developed by DfE and product developers meant that DfE had shared ownership of these.

**Collaboration between product developers, the Evaluators and DfE at the TFA stage, as well as the TFA process itself, supported sufficient check and challenge of intervention ideas.** The Theory of Change workshops were particularly successful,

bringing together key stakeholders to reach collective agreement on the intended outcomes of the interventions to be tested through trialling.

## What were the challenges?

**The period from initial engagement with the adult education sector to the submission deadline was relatively short**[30] and a bolder communications approach may have generated more idea submissions.

**Similarly, some product developers were reluctant to submit their interventions for trialling**, particularly when these formed a core part of their business offer. This highlights the importance of clearly communicating at the call for ideas stage what participation in trials entails, including the possibility that their interventions might not demonstrate effectiveness or impact.

**Some interventions were already being delivered nationally,** which made it difficult to establish trial conditions – specifically, ensuring that control group learners were not exposed to the intervention.

**Decision-making on which ideas and interventions to take forward was sometimes slow.** There was a tendency to seek certainty before discounting ideas, when in hindsight some could have been ruled out earlier. Faster decision-making at earlier stages would have allowed more time for developing the interventions that were taken forward.

**Some interventions were similar to approaches already gaining traction in the sector.** For example, Maths Mastery as a pedagogy was already being tested and rolled out in some 16-19 educational settings. This created a risk of contamination, where control group tutors may have been exposed to similar approaches through wider sector training or professional development, making it potentially difficult to isolate the effect of the intervention.[31] However, the trial was considered suitable to progress despite this risk, which was explored in detail through the IPE.

[30] The period for idea generation was approximately 10 weeks.
[31] This is common in educational research and was mitigated as far as possible by additional IPE activity with the control group to ensure that a clear understanding of what constituted 'business as usual' was gathered during the research.

# 4. Trial design


### Summary of chapter

This section reports on the design process for the Adult Numeracy Trials. It covers the approach to sampling, methods selection, development of Trial Protocols and Statistical Analysis Plans (SAPs) and analysis and reporting. It concludes with reflections on how well the trial design process worked and lessons learned.


## Overview

Figure 5 provides an overview of the key design features of each of the Adult Numeracy Trials.

## Sampling strategy

A key challenge for the course-based Adult Numeracy Trials related to available capacity within the adult education system to deliver the scale of trials that were planned. An integrated approach to sampling was therefore required to ensure efficiency in the allocation of providers to trials and to maximise the number of trials that would be possible with the sample available. This required central management and a structured approach to sample design, recruitment, and randomisation. It also required close working between the Evaluators and Managed Service Supplier (MSS).

The MSS was responsible for maintaining a comprehensive sampling frame of eligible providers built from Individual Learner Record (ILR) data. This central database tracked provider participation across trials, ensuring that no single provider would be asked to participate in more than 3 trials. Whilst individual randomisation would have been preferred to maximise statistical power, cluster randomisation at the provider level was necessary to prevent contamination between treatment and control groups and to accommodate practical delivery constraints. For example, for interventions requiring tutors to adopt new teaching methods, the same tutor could not simultaneously deliver both standard and new approaches to learners within the same class.

## Figure 5: Overview of the design of the Adult Numeracy Trials

| | Trial Name | Type of trial | No. of providers | | No. of schools | | No. of learners | | MDES* | | | ICC** | | Primary outcome P value*** | Attrition | Evidence rating |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Protocol (post attrition) | Achieved | Protocol (post attrition) | Achieved | Protocol (post attrition) | Achieved | Protocol (post attrition) | Ex-post | Achieved effect size (primary outcome) | Protocol (post attrition) | Achieved | | | |
|  | **Encouraging Progression** *(full-scale effectiveness RCT)* | 3-arm trial, randomised at the individual level | N/A | | N/A | | 4,500 | 13,647 *T: 5,557 C: 5,558* | 0.09 | 0.06 | 0.004 | N/A | | 0.120 *(peer letter)* 0.590 *(urgent letter)* | 18% *(16,672 learners at random-isation)* | ★★★ |
|  | **Family Numeracy** *(Small, pilot RCT)* | 2-armed parallel cluster randomised pilot trial, with random allocation at school level, with waitlist | N/A | 24 | 80 | 60 *T: 31 C: 29* | 800 | 421 | 0.4 | 0.6 | 1.16 | 0.2 | 0.34 | 0.003 | 15% *(71 schools at random-isation)* | ★★☆ |
|  | **Adapted Mastery** *(Small, pilot RCT)* | 3-armed parallel cluster randomised pilot trial, with random allocation at provider level | 30 | 25 *T: 13 C: 12* | N/A | | 900 | 914 *T: 470 C: 444* | 0.34 | 0.51 | 0.001 | 0.075 | 0.19 | 0.782 | 34% *(38 providers at random-isation)* | ★★☆ |
|  | **Contextualised Approach** *(Small, pilot RCT)* | 3-armed parallel cluster randomised pilot trial, with random allocation at provider level | 30 | 29 *T: 14 C:15* | N/A | | 900 | 719 *T: 204 C: 515* | 0.34 | 0.7 | -0.02 | 0.075 | 0.42 | 0.977 | 42% *(50 at random-isation)* | ★★☆ |
|  | **Preparation for Maths GCSE** *(Small, pilot RCT)* | 2-armed parallel cluster randomised pilot trial, with random allocation at provider level | 22 | 32 *T: 16 C: 16* | N/A | | 1,760 | 1,021 *T: 716 C: 305* | 0.28 | 0.43 | 0.01 | 0.06 | 0.13 | 0.949 | 22% *(39 at randomis-ation)* | ★★☆ |
|  | **Embedded Maths** *(Implementation and Process Evaluation)* | Small-scale, non-randomised descriptive implementation and process evaluation (IPE) | 10 | 2 | N/A | | 200 | 68 | N/A | N/A | N/A | N/A | N/A | N/A | 80% | ★☆☆ |

Source: DfE analysis of Adult Numeracy Trial Reports

*The Minimum Detectable Effect Size (MDES) is calculated using either Cohen's *d* (for continuous outcome measures) or Cohen's *h* (for binary outcome measures). Both are standardised effect size measures, with Cohen's *d* expressed in standard deviation units and Cohen's *h* calculated using an arcsine transformation to stabilise variance. Interpretation thresholds are similar for both, with 0.2 showing a small effect size, 0.5 a medium one and 0.8 a large one. For example, a Cohen's *h* of 0.20 may reflect, roughly, a difference of 40% versus 47% passing an exam (exact mapping depends on the baseline rate).

** The Intracluster Correlation Coefficient (ICC) measures how similar participants are within the same cluster, compared to participants in different clusters. An ICC greater than 0.10 indicates strong clusters effects. For example, an ICC of 0.07 means that 7% of the variance in outcomes is due to differences between clusters, and 93% is due to differences between individuals

***The p-value is the probability that a result occurred by chance. A small p-value (usually 0.05 or less) suggests the result is 'statistically significant', meaning it is unlikely to have occurred by chance.

To determine the feasibility of running multiple trials with the available provider base, a step-by-step process was developed. This involved estimating required sample sizes for each trial and collectively across the programme, accounting for anticipated attrition and opt-out rates (estimated at 30% and 10% respectively) and calculating the total number of providers required under various participation scenarios. This systematic approach helped identify opportunities to combine trials with similar target populations and outcome measures, thereby reducing the overall demand on the provider network[32]. The strategy also addressed the risk of contamination across trials targeting overlapping populations through development of a shared participant database to track involvement across multiple interventions (facilitated through the Ipsos Data Portal, discussed in detail in Section 5 Data Collection).

## Methods selection

Once interventions were considered feasible for taking forward to trial[33], the Evaluators developed a series of design options for different recruitment scenarios. These ranged from Implementation and Process Evaluations (IPEs) through to Full Efficacy Trials (RCTs), including Pilot RCTs and various Quasi-Experimental Designs (QEDs). For each design option, the Evaluators estimated the minimum number of providers that would need to be recruited. They also specified:

- the type of study (such as experimental, quasi-experimental or observational)
- its level on the Maryland Scientific Methods Scale[34]
- the strength of evidence it would generate, including an assessment of the risk of bias, expected internal and external validity
- the counterfactual approach
- data requirements including sample sizes, timings (baseline and endline), and coverage (learner, cohort, institution)
- power estimation assumptions including effective sample size at analysis, allocation between treatment and control (if applicable) and attrition assumptions
- the research questions it would address

[32] The Adapted Mastery Approach and Contextualised Approach interventions both targeted FSQ Level 1 Maths learners and shared similar outcomes. These were therefore combined into a 3-arm cluster randomised trial with a shared control group, reducing the number of providers needed compared to running 2 separate trials. Whilst the interventions had a shared control group, they maintained separate Trial Protocols and were reported independently.
[33] Section 3: Interventions provides a detailed description of the Trial Feasibility Assessment process.
[34] The Maryland Scientific Methods Scale is a five-level framework for assessing evaluation rigour, ranging from Level 1 (simple correlations without comparison groups) through to Level 5 (RCTs). Higher levels indicate stronger causal evidence, with Levels 4-5 (QEDs with well-matched comparisons and RCTs) considered most robust for determining intervention effectiveness.

After provider recruitment concluded, DfE selected the final trial designs based on the information in the options papers and advice from the Technical Advisor and Technical Steering Group. Table A12 in Annex 10 details the final designs for each trial.

## Trial protocols and statistical analysis plans

### Drafting

After the trial designs were agreed, the Evaluators developed a series of Trial Protocols and accompanying Statistical Analysis Plans (SAPs) using templates created by DfE. These templates drew on best practice and publicly available examples from organisations such as the Education Endowment Foundation[35].

The Trial Protocols detailed the research questions and hypotheses to be tested, along with eligibility criteria for participants (such as age and prior maths attainment) and recruitment procedures for both providers and learners. They specified the intervention's content, duration, delivery methods and who would deliver it, as well as what the control group would receive. The protocols outlined randomisation procedures, primary and secondary outcome measures, sample size calculations with power assumptions and planned statistical analyses. They also addressed ethical considerations including consent procedures and data protection arrangements, strategies for minimising attrition and contamination between groups, implementation fidelity monitoring, and procedures for handling protocol deviations or adverse events. Finally, the protocols included timelines for all trial activities, data management plans, and governance arrangements specifying the roles and responsibilities of all parties involved in trial delivery.

The accompanying SAPs provided detailed specifications for all quantitative analyses to be conducted. They included descriptions of the study population and analysis samples (such as intention-to-treat and per-protocol populations), definitions of compliance and protocol adherence, and procedures for handling missing data including any imputation methods. The SAPs specified the statistical methods for primary and secondary analyses, including the models to be fitted, covariates to be adjusted for, and methods for handling clustered data. They detailed subgroup analyses to be conducted, sensitivity analyses to test the robustness of findings, and any interim analyses if planned. The documents also set out rules for dealing with multiple comparisons, significance levels and confidence intervals to be reported. Additionally, they included procedures for assessing baseline balance between groups, checking statistical assumptions, handling outliers, and reporting effect sizes. The SAPs were finalised before any outcome data

[35] https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluation/evaluation-guidance-and-resources/protocol-study-plan-and-sap-templates

were analysed to maintain the integrity of the trial and prevent data-driven decision-making that could bias results.

## Review process

The Trial Protocols and SAPs underwent a 4-stage review process, described below.

**1**. **Peer review:** All draft Trial Protocols and SAPs underwent peer review, with lead trialists within the Evaluation consortium reviewing each other's protocols and providing feedback before submission to DfE. This process provided quality assurance whilst also enabling learning across the consortium, as trialists could adopt effective approaches from other trials to ensure greater consistency across the programme.

**2**. **DfE review**: The second stage involved the DfE team reviewing and providing feedback to lead triallists on the first draft Trial Protocols submitted. This review assessed Protocols against programme requirements, checking for alignment with policy objectives, feasibility within available timescales and budgets, and compliance with ethical and data protection standards. The DfE team provided written feedback on each Protocol, highlighting areas requiring clarification or amendment, and where necessary, arranging follow-up discussions with trialists to resolve complex issues.

**3. Technical Steering Group review:** After incorporating DfE feedback, lead trialists submitted their revised Trial Protocols and SAPs to the TSG and presented their methodology and underpinning rationale. The TSG provided scrutiny and challenge during the meeting and written feedback afterwards, focussing particularly on ensuring the trials met the highest quality standards for technical rigour.

**4. Ethics review**: Following TSG review, lead triallists submitted Ethics Forms summarising their trials to the Ethics Panel[36], who reviewed ethical issues with triallists present for part of the meeting to address questions. The Panel provided written advisory feedback, which triallists addressed before updating their Trial Protocols. The Panel's report and revised Protocols were then shared with DfE and the TSG for review.

Trial Leads provided written responses to the TSG and Ethics panel explaining how they had addressed their feedback and, where applicable, justifying any recommendations not implemented. DfE retained overarching responsibility for signing off the Trial Protocols and determining the extent to which TSG and Ethics Panel feedback should be actioned.

## Analysis and reporting

**The consortium meetings proved valuable in establishing detailed analytical approaches across all trials**. Through regular collaborative sessions, a comprehensive

[36] Section 2: Governance provides further details on the Ethics Panel.

analysis matrix was developed that mapped methodological decisions to specific trial contexts. The lead trialists from the evaluation consortium agreed the approach to descriptive statistics, ILR data variables and coding, outcome analysis and data standardisation, missing data, subgroup analysis, effect size interpretation and robustness checks. The aim was to ensure consistency across trials while allowing for justified variations where necessary. The agreed approaches were documented specifying both the common methodology and any trial-specific variations with their rationale (see Annex 11). The document was then shared with the DfE and TSG for review and sign off.

**For most trials, analysis adhered closely to the pre-specified Statistical Analysis Plans.** However, there was a need for adaptive approaches when encountering unexpected findings. A notable example occurred in the Family Numeracy trial, where complete separation was observed (zero outcomes observed in the control group). The pre-specified model could not accommodate this scenario, necessitating a systematic exploration of alternative modelling approaches. The team tested several robust alternatives, drawing on guidance from the Technical Steering Group, ultimately selecting the most appropriate solution based on statistical validity and clarity. This balance between maintaining analytical rigour and responding flexibly to unanticipated findings ensured both the integrity applicability of the trial results.

**Quality assurance was embedded throughout the analysis stage of the Adult Numeracy Trials.** During the data linking phase, extensive standardisation and checks were conducted, with early ILR data snapshots used to test processes and assess completeness before final analysis. The analytical framework, developed by the consortium and reviewed by the independent Technical Advisor, documented agreed approaches for handling analytical challenges such as missing data. All analyses underwent peer review within the evaluation consortium, including code verification to ensure reproducibility. Statistical outputs were systematically checked against pre-specified Statistical Analysis Plans, with any deviations documented and justified within individual trial reports.

**The DfE provided a standardised template for trial reporting to ensure consistency across trials.** Reports underwent multiple stages of quality assurance: lead trialists conducted internal reviews before circulating drafts for peer review within the wider consortium. A structured peer review protocol ensured consistent scrutiny at this stage. Draft reports were submitted to the DfE and Technical Steering Group for review and feedback. Finally, Ipsos' Quality Director conducted editorial reviews using a consistency checklist developed in collaboration with the DfE, ensuring a level of consistency across trial reports in terms of structure, terminology and presentation of findings.

**Tools were developed to facilitate cross-trial learning and enable meaningful comparisons.** Summary presentation templates were created using standardised

PowerPoint slides that captured key metrics, effect sizes and confidence intervals in consistent visual formats. Additionally, an IPE Matrix was developed to enable systematic comparison of implementation factors, process measures and outcomes across diverse trial contexts (see Figure 9). These tools aimed to support stakeholders to quickly identify patterns and draw insights across the programme of trials.

### Publication

The Trial Protocols and accompanying SAPs were published on the Open Science Framework (OSF)[37] to ensure transparency and enable future replication. Prior to trial commencement, trial leads were required to upload final PDF versions of both the Trial Protocols and SAPs to the OSF. This pre-registration helped maintain scientific integrity by making the planned analyses public before data collection began. Trial leads were responsible for maintaining these documents throughout the trial period, updating Trial Protocols to reflect actual delivery timelines, any protocol deviations, and final recruitment volumes. Where substantial changes occurred during trial implementation, updated versions were uploaded with clear version control and change logs to maintain a transparent audit trail. This approach aligned with best practice in open science, enabling other researchers to learn from both planned and actual trial implementation.

## What worked well?

**The consortium comprised triallists with extensive expertise in designing and delivering trials across a range of contexts**. This proved invaluable when designing trials in a relatively new field and sector with limited prior experience, particularly for establishing realistic assumptions to underpin power calculations and estimating likely attrition rates. It also provided an additional level of quality assurance to the review process.

**The TSG and independent Technical Advisor brought considerable methodological expertise and independent scrutiny to the trial design process, serving as a critical quality assurance mechanism for the programme.** Their focus on technical rigour helped ensure that trials met the highest methodological standards, identifying potential threats to validity, questioning assumptions underpinning power calculations, and suggesting refinements to analytical approaches. This external scrutiny was particularly valuable given the ambitious scale of the programme and the challenges of conducting trials in adult education settings where robust evaluation evidence was limited. The TSG's input not only strengthened individual trial designs but also helped maintain

[37] Annex 9 contains weblinks to each of the published Trial Protocols.

consistency in quality across the portfolio of trials, ultimately enhancing the credibility and usefulness of the evidence generated for policy and practice.

**The development of multiple design options prior to provider recruitment meant that all possible methodological scenarios were thoroughly explored upfront**. This approach enabled DfE and Trial Leads to make robust design decisions without needing to revisit the entire review process, as might have been necessary had only one design been considered. It also allowed DfE to make informed decisions quickly once recruitment concluded, which was particularly important given the tight timescales between the end of recruitment and the start of trial delivery.

## What were the challenges?

**Delivering multiple trials simultaneously in a sector with a relatively small provider base (compared to the school sector, for example), necessitated careful sampling.** This constraint also led to prioritising Pilot RCTs over full efficacy trials, as the latter would have required the vast majority of adult education providers to have signed up – an unrealistic expectation for a sector with limited previous experience of trialling.

**Due to project timelines and term start dates, the trial design process and provider recruitment ran concurrently rather than sequentially** (ideally, trial design would be finalised after knowing how many providers had been recruited). This meant that multiple drafts of Trial Protocols were needed to accommodate changing methodological requirements, rather than simply refining a single agreed methodology. It also meant that providers were asked to commit before receiving full trial details, with some trial designs changing after recruitment.

**The lack of previous trialling in the adult education sector meant that key information to inform trial designs was missing**. This included likely recruitment numbers, attrition rates of both providers and learners, and compliance with data requirements (such as survey response rates). This meant that the trial designs were based on informed estimates from triallists with limited or no benchmarks to guide their assumptions, particularly for power calculations where evidence was limited. However, these trials have now established valuable benchmarks for future researchers conducting trials in adult education.

# 5. Recruitment


## Summary of chapter

This section reports on the recruitment of providers, schools and adult learners (aged 19 or above) to the 4 course-based adult numeracy, and Family Numeracy, trials. Recruitment refers specifically to providers or schools that formally signed contracts to participate, as distinct from expressions of interest, which indicated initial interest but did not always result in participation. Only those who signed contracts are included in the recruitment figures. This section draws on management information, the stakeholder engagement strategy, surveys with providers and schools, interviews with the managed service supplier (MSS), and lessons learned during programme delivery.

The recruitment of providers and schools was led by the MSS on behalf of the Department for Education (DfE). Decisions relating to the level of randomisation and the target number of providers to be recruited to each of the trials are explained in section 4: Trial Design.


## Recruitment of providers and schools

### Expressions of interest

A total of 289 expressions of interest in the course-based trials were received from 166 separate providers (FSQ [143], Preparation for Maths GCSE [91] and Embedded Maths [55]). Of these, 96 providers went on to sign a contract and were formally recruited while 70 did not (a conversion rate of 58%). A further 293 primary state schools expressed an interest in the Family Numeracy trial and 72 of these went on to participate in the trial (25% of the total). The reasons why some providers and schools did not progress in the trial are discussed in section 6: Attrition.

### Number and profile of providers / schools recruited

As noted, 96 separate providers and 72 schools were recruited to the course-based trials, defined as having signed a contract at the point of randomisation (if the trial involved randomisation). Of the providers, 8 took part in 3 trials (from among Family Numeracy, Contextualised Approach for FSQ Maths Level 1 and Adapted Mastery Approach for FSQ Maths Level 1 and Preparation for Maths GCSE), 28 took part in 2 trials and the remaining 60 took part in one trial. There were a range of different provider types recruited to the trials. Local authority-run adult learning services formed the largest group (37), followed by private, third-sector or independent providers (32) and Further

Education (FE) colleges (27). All schools recruited to the Family Numeracy trial were state primary schools. Recruitment targets for each trial and the actual numbers recruited are set out in Table 2.

**Table 2: Target and actual number of providers/schools recruited per trial**

| Trial name (recruitment of providers or schools) | Target no. for full RCT | Target no. for pilot RCT | No. recruited[38] | Final trial design |
|---|---|---|---|---|
| Embedded Maths (providers) | [z] | 60 | 10 | Implementation and Process Evaluation |
| Family numeracy (schools) | 188 | 36 | 72 | Pilot RCT |
| FSQ Maths Level 1 (providers)[39] | 177 | 66 | 69 | Pilot RCT |
| Preparation for Maths GCSE (providers) | 206 | 30 | 32 | Pilot RCT |

Source: Management information and analysis undertaken by Evaluators to identify target numbers of providers/schools for different trial design options. Totals not provided as some providers signed up to more than one trial. [z] Full RCT not possible due to not enough providers delivering Health and Social Care Level 2 to adults in England

Recruitment targets to power a full RCT were not met but preparations were made in advance to enable a pivot to alternative methods requiring lower numbers of providers. The evaluators prepared a series of papers setting out step-down options for each trial for different recruitment scenarios (see section 4: Trial Design for more details). These were used to support discussions and check-and-challenge meetings with DfE and the Technical Steering Group on alternative options and decisions on the final methods.

[38] Defined as having signed a contract for Embedded Maths) and at the point of randomisation (for the Pilot RCTs) (Not all providers and schools remained in the trials – retention and attrition are discussed in Section 6: Attrition.

[39] Included 2 trials Contextualised Approach for FSQ Maths Level 1 and Adapted Mastery Approach for FSQ Maths Level 1

# Recruitment strategy and its evolution

## ‘Top-down’ recruitment

The recruitment strategy was initially conceived as a ‘top-down’ model[40], whereby representative bodies were engaged first and encouraged to recruit adult learning providers from their networks. The primary targets were Mayoral Strategic Authorities (MSAs)[41], local authorities (LAs), and sector representative bodies such as the Association of Colleges (AoC), Holex and the Association of Employment and Learning Providers (AELP). The MSS anticipated a single agreement with an MSA who could potentially recruit in hundreds of providers[42]. However, according to MSS feedback, this top-down approach was less effective than expected, and ultimately no such agreements were secured with these organisations.

In the early stages of ‘market warming’ (January/February 2024), the MSS spent considerable time engaging with MSAs and LAs. The MSS engagement with MSAs and LAs did not yield the numbers expected. A key barrier noted by the MSS was that the funding offered was insufficient to incentivise large MSAs and LAs to take on the administration associated with managing the requirements of the trials on behalf of providers. During this time DfE also wrote to MSAs and posted on their social media channels. Social media channels raised awareness of the trials but based on interviews with providers, this did not directly translate to signups.

Although this recruitment strategy did not yield the expected numbers it set the foundations for further recruitment raising awareness of the trials within the sector.

## Direct-to-provider recruitment

The recruitment strategy pivoted to a more effective approach, directly contacting providers instead of MSAs and LAs, combining personalised direct contact with targeted, multi-channel campaigns. The MSS reported that the most effective methods to stimulate sign-ups were personalised emails and direct calls from the MSS, DfE and messaging disseminated by local authorities.

[40] A top-down recruitment model in this context refers to a strategic approach that begins by engaging sector-wide bodies, national representatives, and large provider networks before reaching individual organisations. This typically involves starting with umbrella organisations, sector leadership bodies, and national networks who then facilitate connections with their member institutions and local providers.
[41] There was a transition from Mayoral Combined Authorities to MSAs as part of the 2025 Devolution Bill, reflecting a broader shift in UK devolution, moving from bespoke arrangements to a uniformed statutory framework.
[42] The MSS Stakeholder Engagement Strategy (February 2024) noted there are: 10 MCAs (covering 84 LAs), 317 LAs, 569 providers delivering Adult Education courses, 158 Further Education Colleges, 267 providers delivering GCSE maths, 545 providers delivering FSQ Maths level 1 and 55 providers delivering Health and Social Care level 2 to adult-only classes.

> "Doing direct outreach to the providers and colleges that we had contact details for was the best. I think just speaking to them and talking them through the trial" – *Managed Service Supplier (MSS)*

Providers and schools that were recruited to the trials were invited to complete a short survey to share feedback on their experiences. This included those who signed up to one or more trials but subsequently withdrew. Table 3 shows that recruited providers and schools were most likely to have heard about the adult numeracy trials from their local authority (31%) or DfE (18%). This suggests that whilst the initial 'top-down' strategy of recruiting MSAs or LAs was unsuccessful, engagement of LAs was an effective approach to raising awareness of the trials amongst eligible providers.

**Table 3: How did you hear about the Multiply Education Research Trials (adult numeracy trials)?**

| Ways of hearing about the Multiply Education Research Trials (adult numeracy trials) | Percentage of providers and schools |
|---|---|
| Local authority (incl. Family Learning Team, Multiply Local Areas Team) | 31% |
| Department for Education | 18% |
| Word of mouth | 9% |
| Direct approach from Managed Service Supplier (MSS) | 8% |
| Direct approach from course content / training provider | 6% |
| Sector / industry bodies | 5% |
| Attendance at an event / conference | 5% |
| MSS trial website | 4% |
| Social media e.g. LinkedIn post | 2% |
| From my Multi-Academy Trust | 2% |
| FE Week article | 0% |
| Other[43] | 13% |
| Don't know / can't remember | 9% |

QHEAR: How did you hear about the Multiply Education Research Trials (adult numeracy trials)? This question was asked to both providers and schools who took part in the trials (n=54) and those who withdrew (n=73). Percentages do not add up to 100% as this was a multicode question so participants could select more than one response.

The MSS also reported that webinars (co-delivered by evaluators, product developers, and DfE) and in-person conferences were effective at generating sign-ups. These activities provided an opportunity to explain the purpose of the trials and the benefits of

[43] Other responses included adult education organisations such as Step2Skills and CALAT adult learning and Multiply contact.

participation, which was beneficial given the adult learning sector's limited previous experience of RCTs.

## Trial-specific recruitment approaches

The initial recruitment strategy for the Family Numeracy trial relied on providers with family learning experience to recruit schools. However, this approach proved less successful than anticipated, leading to a change partway through due to low provider sign-ups. The contact list available to the MSS (provided by DfE) contained providers delivering Multiply provision rather family learning specialists. This prompted them to refocus their efforts on engaging local authority family learning teams directly.

The MSS also shifted to recruiting schools directly (rather than through providers). Once schools were recruited, the MSS manually matched them with local providers, effectively bringing the demand to providers rather than expecting them to generate it through school engagement. Although time-consuming, this approach proved more successful than the original provider-led strategy.

The MSS delivered a social media campaign via LinkedIn and X to support recruitment. They found LinkedIn to be the most effective in generating interest and signups. The MSS identified the early availability of the Trial Readiness Pack[44] as being a key success factor, which meant that providers and schools had full information on the trial requirements prior to signing up (which was not the case for some of the other trials).

## Incentives

Financial and non-financial incentives were offered to providers, schools and learners that participated in the course-based trials and research activities. Table 4 details the financial incentives that were provided.

[44] See section 7: Primary Data Collection for further information on Trial Readiness Packs.

**Table 4: Financial incentives**

| Providers / Schools | Treatment | Control |
|---|---|---|
| Providers | £1000 participation payment per provider | £1000 participation payment per provider |
| Providers | 10% admin premium | 10% admin premium |
| Providers | £7.20 per additional guided learning hour | Not applicable |
| Schools | £500 for schools that recruited 8 or more (eligible) parents | £500 for schools that recruited 8 or more (eligible) parents |
| Learners[45] [46] | £30 shopping vouchers for participation in in-depth interviews | £30 shopping vouchers for participation in in-depth interviews |

Source: MI data / Trial Readiness Packs

To help overcome initial recruitment challenges, a £1,000 participation incentive was offered to all providers for each trial several months after recruitment began. This was offered to all providers, regardless of when they signed up or whether they were randomised to the treatment or control group. The MSS reported that there was a notable increase in the number of providers expressing an interest in taking part and signed up after the incentive was introduced. This learning was taken forward in the Family Numeracy trial (which recruited later than the other trials) and MSS offered incentives from the outset for providers (£1,000) and schools (£500) which the MSS reported supported with recruitment.

Alongside monetary incentives, a range of non-financial incentives were highlighted in the recruitment materials as key benefits to providers and schools for taking part. This feedback came from tutors, providers and the MSS. These included:

- free intervention-relevant training for tutors in the treatment groups, delivered by the product developers of each of the interventions being tested
- access to new teaching resources and materials for those in the treatment groups, equipping staff with new pedagogies
- exclusive access to research and findings via a dedicated online learning resources platform and webinars

[45] Incentives for learners were only available for participating in research activities, not for participating in the learning. Not all learners took part in an interview. Those invited were selected by evaluators to ensure a mix by different characteristics. Participation was voluntary.
[46] No incentives were offered to learners in the Encouraging Progression trial.

- involvement in a large-scale, government-funded research project. This was framed as a chance to be an ‘innovator’ and ‘sector ambassador’ which aligned to some providers' strategic goals
- for Family Numeracy, a ‘waitlist’ design was offered as an incentive to schools that were randomised to the control group, this meant they could still receive the family learning programme (in the following academic year, outside of the trial period)

# Recruitment of learners

## Number and profile of learners recruited

In total, 2,760 eligible learners took part in the course-based trials.[47] Table 5 shows that the Preparation for Maths GCSE, Contextualised Approach and Adapted Mastery Approach for FSQ Maths Level 1 trials accounted for the highest shares of participating learners (combined accounting for 82% of the total). The Family Numeracy trial accounted for around 1 in 7 (15%) and Embedded Maths had the smallest overall share of participating learners (3%).

The table also shows the average number of learners per provider / school. This could be useful in informing the design of future trials in the sector, particularly in relation to the recruitment assumptions underpinning power calculations.

All trials aimed to target learners with low skill and/or confidence in maths. The eligibility criteria for all trials was adult learners (aged 19 or over during the academic year). For Preparation for Maths GCSE and FSQ Maths Level 1, the eligibility criteria was adult learners who did not have a Level 2 maths qualification. For Family Numeracy, learners who did have Level 2 maths were still eligible to take part if they were not confident in using maths in everyday life. For Embedded Maths, eligibility was based on eligibility for the Level 2 Health and Social Care course, rather than their existing maths ability and qualifications. Participants for the Encouraging Progression trial were selected from the ILR as having completed a non-qualification-bearing Multiply course.

[47] This excludes learners who dropped out, did not consent to take part, consented to take part but the provider subsequently dropped out, whose data was submitted erroneously or were otherwise ineligible.

**Table 5: Number of learners who took part in the trials**

| Trial name (unit of randomisation) | Total number of learners[48] | Average number of learners per provider / school[49] |
|---|---|---|
| Preparation for Maths GCSE (providers) | 1,058 | 42 |
| FSQ Maths Level 1 (providers)[50] | 1,213 | 28 |
| Embedded Maths (providers) | 68 | 34 |
| Family Numeracy (schools) | 421 | 7 |
| Total | 2,760 | - |

Source: Learner data

The profile of target learners differed for each trial and recruitment, and retention approaches reflected this.

**Preparation for Maths GCSE Trial**: Target learners were adults who signed up to a Maths GCSE course. The MSS noted that because these learners signed up to a year-long course, their predicted numbers were more reliable, and there were fewer drop-outs compared to other trials that involved shorter courses and more flexible learning approaches.

**FSQ Maths Level 1 Trial:** Target learners were adults studying for a Level 1 Functional Skills Qualification (FSQ) in Maths. Providers often had mixed-level classes, teaching Level 1 and Level 2 learners together. Providers' predictions for FSQ Maths Level 1 learner sign-ups were less certain and forecast numbers often varied considerably to actual learners recruited.

**Embedded Maths IPE:** Target learners were adults studying for a Level 2 Health and Social Care vocational qualification. A key characteristic of this group was that they were not actively seeking to engage in a maths course as the intervention was designed to embed numeracy into their vocational course.

**Family Numeracy Trial:** Target learners were parents or carers of Year 2 primary school children. This group was found to be challenging to engage, with specific difficulties faced in reaching working parents. The success of recruitment depended heavily on the school's ability to get eligible parents to sign up to the programme and opt in to the trial.

[48] This is the final number of eligible learners who took part in the trials.
[49] Average is based on the final number of learners divided by the final number of providers or schools (Family Numeracy) after attrition/drop out.
[50] Included 2 trials Contextualised Approach for FSQ Maths Level 1 and Adapted Mastery Approach for FSQ Maths Level 1

## Learner recruitment strategy

### Recruitment strategy for FSQ Maths Level 1, Preparation for Maths GCSE and Embedded Maths

For FSQ Maths Level 1 (Contextualised Approach and Adapted Mastery Approach), Preparation for Maths GCSE and Embedded Maths, providers were responsible for recruiting learners to the trial including providing them with the relevant information about what taking part would involve. Responsibility for engaging the learners and getting them to participate in early-stage research (such as the baseline surveys) also mainly fell to the provider. The MSS and the evaluation team were responsible for providing the relevant materials and explaining the process to providers. Once learner contact details had been shared with the evaluation team, recruitment to interviews and surveys could be done directly.

Providers and schools were provided with information about the trials (Trial Readiness Packs) to share with their learners to support recruitment. The Trial Readiness Packs contained learner information sheets and links to the baseline learner surveys. For FSQ Maths Level 1 (Contextualised Approach and Adapted Mastery Approach), Preparation for Maths GCSE and Embedded Maths, consent was not required under the legal basis 'Public Task' but learners who wished to opt out from participating could inform their tutor, provider or the evaluation team directly.

### Recruitment strategy for Family Numeracy

For the Family Numeracy trial, a more hands-on support strategy was implemented due to the need for learners to opt into this trial, and to minimise burden on participating schools given that this was a new area of activity for them. The MSS, product developer and evaluation team were all involved in supporting recruitment. Methods to support schools in encouraging parents to enrol in the programme included providing a detailed recruitment pack with materials developed by the product developer, including an 'invitation' from the children asking parents if they would like to participate and school posters. As learners were recruited to the Family Numeracy trial via schools (rather than independently signing up to adult maths courses as per the other course-based trials), they had to provide explicit opt-in consent to taking part in the trial. The parental enrolment form included details of what opting into the trial would involve[51]. Other recruitment activities included sharing Trial Readiness Packs (for providers, schools, parents and tutors), producing a video explaining the benefits of the programme to parents, and hosting webinars with schools and providers explaining the evaluation requirements.

[51] Learners could still participate in the programme if they did not opt in to the trial.

### Information shared with learners recruited

Providers and schools were responsible for recruiting learners to the trials. This included ensuring each learner received full information about the trial (provided in the relevant sections of the Trial-Readiness Packs), how their data would be used and the opportunity to opt out. For the Family Numeracy trial, learners were required to provide opt-in consent for participation.

## What worked well?

**The MSS said the multi-channel communication strategy maximised reach and impact.** DfE sent personalised communications to providers and schools to encourage take up and engaged their Local Area teams to help raise awareness. Alongside, the MSS were in direct communication with providers to address any questions and sign them up for the trial. This was particularly important for a sector where there has been limited previous experience of RCTs.

**Providers highlighted the offer of high-quality Continuing Professional Development (CPD) to be a key selling point** as the adult learning sector had limited access to free training opportunities for tutors.

**The MSS said that, once introduced, the £1,000 incentive payment for providers was critical to converting Expressions of Interest (EOIs) into sign-ups**. Whilst providers could see the range of potential benefits to taking part, it was only viable for them if the funding model covered their costs.

> "There was a delay in the agreement of the funding model which kind of hindered the recruitment in the beginning but once it was agreed and there was a £1,000 incentive attached to it, that got a lot more people interested" – *Managed Service Supplier (MSS)*

**Credible partners helped to build interest, assure providers of the potential benefits and maintain momentum**. Leveraging the reputation of the DfE, product developers such as Mathematics in Education and Industry (MEI), Campaign for Learning (CfL) and The Education Training Foundation (ETF), helped build trust and facilitate engagement.

**The MSS found that emphasising the potential positive impact for future learners and contribution to the evidence base motivated some providers to express interest.**

**The MSS said recruitment was less challenging for trials where the intervention was easily understood,** embedded within business-as-usual teaching and where the

tutor training requirements were perceived as manageable. The Preparation for Maths GCSE trial was consistently described as the 'easiest' to recruit to for these reasons.

**Flexibility from the MSS, DfE and the evaluation team in adapting the approach to recruitment was essential.** Pivoting quickly and engaging the evaluation consortium and support from DfE to respond rapidly to challenges was essential in supporting recruitment efforts within a compressed timeframe.

## What were the challenges?

**The tight timeline for delivery meant that the products and trial requirements were still being finalised once recruitment had started**. The MSS reported that this lack of clarity on what would be involved prevented some providers from signing up.

**Launching multiple trials at the same time to an adult education sector that is relatively small** (there are around 2,000 providers across England compared to, for instance, around 17,000 primary schools and 3,500 secondary schools) – limited further by not all adult education providers offering relevant courses. This meant providers did not always have capacity if already taking part in one of the other trials.

**In some cases, providers said the demands of the trial were too high for the incentives offered, leading to disengagement**. The initial funding model (before the £1,000 incentive was introduced) was not seen as attractive for many providers as did not cover the costs involved in taking part in the trial.

**On some trials, providers found the tutor training requirements to be burdensome and a potential barrier to participation**, including for Embedded Maths and FSQ Maths Level 1 (Contextualised Approach and Adapted Mastery Approach) which had more intensive requirements than the others.

**The lack of experience with RCTs within the adult education sector meant that significant time and resources had to be invested by the MSS and Evaluators in 'market warming'**. This included explaining evaluation concepts such as randomisation and control groups. Schools were typically more familiar with these research methods due to the higher prevalence of trials in their sector.

**More direct contact with tutors may have been beneficial in supporting recruitment of learners**. Those in the control group who did not receive training or direct contact with the evaluation team before the trials started were not as well equipped to explain the benefits of participating in the trials to learners.

## Capacity of the adult education sector to deliver trials

A key constraint to delivering trials in adult education is the capacity of the sector. Adult education providers were not resistant to the idea of trials; they were often excited by the opportunity to contribute to the evidence base and bring new approaches to their learners. However, the relatively small pool of adult education providers (around 2,000 in total) was a constraint to recruiting enough eligible providers to deliver fully powered RCTs (based on randomisation at the provider level). This was compounded by recruiting for several trials at the same time.

Further work is required to better understand barriers to engagement of MSAs. For those who do participate, the administrative burden must be minimised and clearly communicated, with financial incentives sufficient to compensate for the effort involved. Finally, while CPD for tutors is highly valued, careful consideration needs to be given to the time investment required to take part so as not to constrain their already limited capacity.

# 6. Attrition


## Summary of chapter

This section reports on the scale, timing and reasons for attrition from the trials as well as possible mitigations to prevent attrition from future trials. Attrition refers to providers and schools who withdrew at different stages in the process including when a trial contract was issued, after contract signed, after randomisation and after delivery starting (this section does not report in detail on providers withdrawing at the expression of interest stage, which is discussed in Section 5: recruitment). Information about attrition was drawn from data recorded by the Managed Service Supplier (MSS), interviews (n=10) and surveys with providers who withdrew from trials (n=54), an interview with the Managed Service Supplier and lessons learned during delivery of the programme.


## Scale of attrition

This section examines attrition separately for the provider-based trials (Contextualised Approach for FSQ Maths Level 1 and Adapted Mastery Approach for FSQ Maths Level 1, Prep for Maths GCSE and Embedded Maths) and the school-based Family Numeracy trial, as they involved different recruitment and delivery processes.

### FSQ Maths Level 1 (Contextualised Approach and Adapted Mastery Approach), Prep for Maths GCSE and Embedded Maths

The MSS attrition report shows a degree of provider attrition across all trials and no notable differences by type of provider (further education colleges, local authorities or other types of adult education providers). Attrition tended to happen before the trials started and there were fewer instances of attrition after randomisation and once delivery was underway.

Table 6 provides an overview of the scale of attrition of providers across trials. It shows the numbers of providers who were issued a contract (but may not have signed the contract), the total number who withdrew and the resultant numbers that were included in the final analysis.

**Table 6: Scale of attrition for providers across trials**

| Trial | Number of providers issued a contract | Total who withdrew (at any stage after being issued a contract) | Number of providers included in final analysis |
|---|---|---|---|
| FSQ Maths Level 1[52] | 85 | 43 | 42 |
| Prep for Maths GCSE | 46 | 21 | 25 |
| Embedded Maths | 21 | 19 | 2 |

Source: Number of contracts issued and providers withdrawn from the MSS's attrition report, number of providers included in the final analysis from final matched datasets.

### Family Numeracy

As noted in Section 5: Recruitment, a total of 293 schools expressed an interest in taking part in the Family Numeracy trial. Interested schools were required to recruit a minimum of 5 eligible learners to progress to randomisation. Many schools struggled to reach this target and were therefore unable to progress in the trial. The final number of schools recruited was 72 and, following further attrition, the number included in the final analysis was 61.

## Timing of attrition in trial delivery

Figure 6 shows the attrition points for learning providers, FSQ Maths Level 1 (Contextualised Approach and Adapted Mastery Approach), Preparation for Maths GCSE and Embedded Maths, and schools, Family Numeracy, at each stage in trial delivery.

[52] Included 2 trials Contextualised Approach for FSQ Maths Level 1 and Adapted Mastery Approach for FSQ Maths Level 1

**Figure 6: Attrition map**

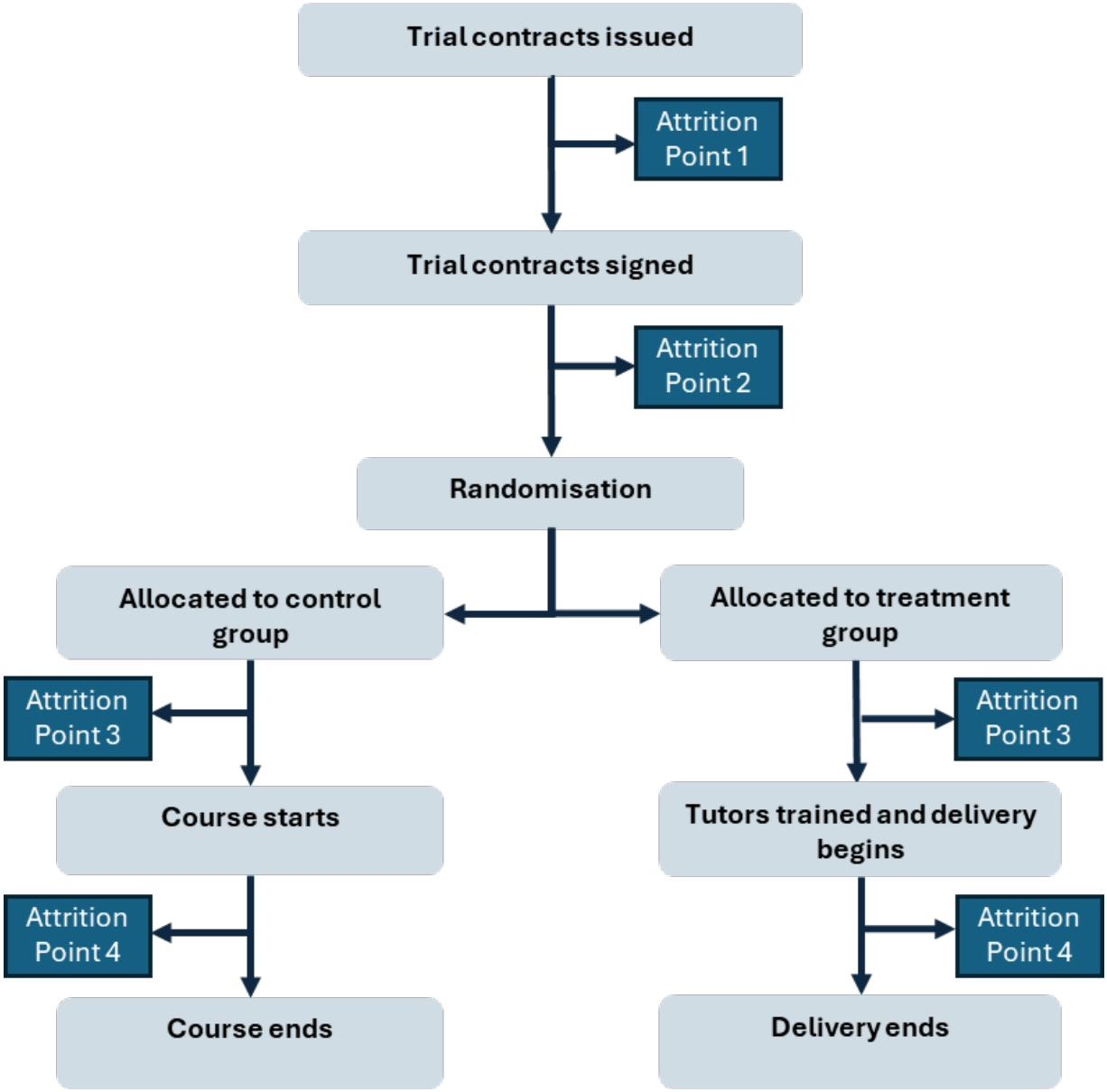



Table 7 shows the number of providers/schools that withdrew at each Attrition Point (AP) by trial based on information from the MSS attrition report. The sections that follow provide further detail on the reasons for attrition at each point.

**Table 7: Number of providers/schools that withdrew by stage**

| Trial (providers / schools) | Before contract signed – AP1 | Before randomisation – AP2 | After randomisation – AP3 (Treatment) | After randomisation – AP3 (Control) | After trial start – AP4 (Treatment) | After trial start – AP4 (Control) | Total |
|---|---|---|---|---|---|---|---|
| FSQ Maths Level 1[53] (providers) | 16 | 0 | 13 | 6 | 5 | 3 | **43** |
| Prep for Maths GCSE (providers) | 8 | 6 | 1 | 0 | 3 | 3 | **21** |
| Embedded Maths (providers) | 8 | 3 | 4 | [z] | 4 | [z] | **19** |
| Family Numeracy (schools) | 96 | 125 | 6 | 5 | 0 | 0 | **232** |

Source: MSS's attrition report. Note: The total number of unique providers who withdrew is lower than the sum of these figures, as providers could have withdrawn from multiple trials. The Embedded Maths trial did not have a control group, which is why this is marked as NA.

## Attrition point 1: Before contract signed

Across trials there was some degree of attrition before contracts were signed. Providers either disengaged with communications at this point and were withdrawn (following agreement with DfE after several attempts to communicate with the provider) or voluntarily withdrew having reconsidered what was required.

## Attrition point 2: Before randomisation

Attrition at this point varied notably across trials. The attrition was particularly substantial for Family Numeracy at this stage as it was when schools were expected to recruit parents to the programme. The majority (n=79/125) of Family Numeracy school withdrawals at this point were a result of not recruiting enough parents. The FSQ Maths Level 1, Prep for Maths GCSE and Embedded Maths did not have to recruit learners in the same way, with learners already signed up to the courses, and given the option to opt out of the trial.

[53] Included 2 trials Contextualised Approach for FSQ Maths Level 1 and Adapted Mastery Approach for FSQ Maths Level 1

### Attrition point 3: after randomisation

For FSQ Maths Level 1 (Contextualised Approach and Adapted Mastery Approach), Prep for Maths GCSE and Embedded Maths, post-randomisation attrition was typically unrelated to treatment or control group allocation. Providers who withdrew after randomisation cited other reasons (see Section 6.3: Reasons for attrition). However, group allocation did affect a small number of Family Numeracy schools, as some of those assigned to the control group could not delay delivery until the autumn term as the waitlist trial design required. Some schools were also affected when their associated provider withdrew at this stage (n=4 Family Numeracy providers withdrew).

### Attrition point 4: after trial start

A small number of providers (n=9) across trials did not participate in the tutor training required of the treatment group and subsequently did not deliver the intervention. Attrition at this point typically occurred when providers stopped engaging with communications from the MSS. Some providers who withdrew at this point cited lack of resource (admin support, tutors, learners) to continue with the trial and meet the associated data requirements. Those that withdrew at this stage had not submitted any learner data to the Ipsos Data Portal.

## Reasons for attrition

This section explores reasons for attrition from the course-base trials, based on feedback from providers who withdrew (gathered via interviews and surveys). It first discusses reasons common across all trials, then examines trial-specific reasons.

### Cross-trial reasons

#### Internal challenges

A key driver of attrition amongst trials, and a barrier to meeting trial requirements, was a lack of tutor and teaching capacity within providers. Engagement with the adult education sector indicated that there was (and remains) widespread tutor shortages, high levels of tutor turnover between academic years, and persistent barriers to recruitment and retention of teaching staff. Amongst providers surveyed, over half (57%) stated that capacity was a reason for their withdrawal, specifically a lack of tutor capacity to deliver the course (32%) and insufficient time to plan, adapt and deliver the new approach (25%). The lack of capacity was particularly acute for trials that involved more time-intensive interventions and tutor training requirements, though capacity was considered a broader problem at the sector level. Part-time and evening work patterns across the adult education sector also meant that some providers were unable to dedicate sufficient resource to trial delivery and data upload requirements.

Challenges with capacity were compounded for providers that were taking part in multiple trials, particularly those allocated to more than one treatment group.

The varied approach to teaching across the adult education sector presented challenges for intervention delivery and compliance to trial requirements, contributing to attrition. For example, the course-based trials assumed that delivery would align with typical academic year timings. This meant that providers that did not follow term dates (e.g. compressed learning approaches) were sometimes unable to deliver the interventions as intended and thus withdrew voluntarily. Qualitative evidence indicated that evening and weekend teaching also meant that some trials were incompatible with 'business as usual' delivery.

> "Our timetable for the functional skills courses is very tight. Some of them are in the evening, so we couldn't add on extra hours at the end of the night or before it, because people are coming from work." – *Withdrawn provider (Treatment group)*

### External challenges

A lack of understanding of requirements, insufficient guidance, and mid-trial changes also contributed to attrition. Almost 1 in 5 (18%) surveyed providers who withdrew reported that participation requirements were not clearly explained. Interview participants indicated that training and Trial Readiness Packs were delivered too close to the start of the academic year, leaving insufficient preparation time. This issue was compounded when trial parameters changed, such as the revised delivery timeframes for Family Numeracy. The delayed distribution of Trial Readiness Packs further contributed to confusion around data requirements. Additionally, the sector's limited awareness and experience of RCTs meant providers were often unprepared for the data requirements and administrative burden involved.

The resource required to take part in a trial was another reason for attrition amongst some providers, despite the additional funding provided upon contract signing. Almost 1 in 5 (18%) surveyed providers that withdrew from at least one trial cited insufficient funding[54] to cover the costs of delivery as a reason for withdrawing. Providers also referenced the fact that trial funding did not align with the timescales for the Multiply programme funding, which impacted staff and teaching capacity. Some providers stated that their organisation was at a 'loss' from trial participation, despite some of the non-financial benefits.

> "The monetary value of the contracts is almost not worth the hassle, not in a bad way, but it has got to make a profit for the company and be worth it financially. We

[54] Funding for providers was £1,000 participation payment per provider and 10% admin premium. Additionally for those in the treatment group, £7.20 per additional guided learning hour.

would put more time and effort into something, but not for such a little return." – *Withdrawn provider (Treatment group)*

Surveyed providers who remained in the trials were asked what additional support would have been helpful. Of these, 22% wanted more information about trial requirements, while 30% requested more lead-in time for preparation. For example, interviewed providers noted that whilst the Trial Readiness Packs were useful resources, they would have benefited from receiving them earlier.

### Allocation to treatment or control groups

Whilst allocation to treatment or control groups contributed to attrition, it was not a primary reason for withdrawal. Rather it exacerbated existing capacity and resource challenges reported by providers who withdrew.

Providers allocated to the control group that subsequently withdrew did not consider this a factor that led to withdrawal, but did express disappointment that they were unable to benefit from access to professional development. Conversely, interview evidence indicated that some providers allocated to the treatment group withdrew due to tutor capacity challenges, which were exacerbated as a result of having been allocated to the treatment group across multiple trials.

## Trial-based reasons[55]

### Family Numeracy trial

The school-based nature of the Family Numeracy trial created distinct challenges in recruiting primary schools and learners (as discussed in section 5: Recruitment) and was also a cause of attrition at both the provider and school level. The most common reason for attrition for Family Numeracy providers that were surveyed was insufficient time to plan, adapt and deliver the new approach. Providers were also reliant on schools to recruit learners, a variation from how providers usually recruit learners onto their courses. In cases where insufficient learners were recruited, providers had to withdraw. The challenges associated with this approach to recruitment were recognised by the MSS as a trial-specific cause of attrition.

Providers that were interviewed also noted the challenges associated with recruiting and engaging schools, particularly in the busy Autumn term. Though there were other reasons to explain attrition from the Family Numeracy trial (related to delayed allocation and changes to delivery timescales), the additional, two-tiered demands around both school and learner recruitment directly contributed to attrition from this trial.

[55] Prep for Maths GCSE is not reported as there were no reasons for attrition specific to Prep for Maths GCSE.

For primary schools surveyed, the most common reason for attrition was an inability to recruit enough learners to take part (64%). In particular, schools were unable to provide learners with clarity around delivery timings (as the waitlist design for the control group was still being finalised), which exacerbated existing challenges to recruitment and contributed towards attrition:

> "One of the main reasons [for withdrawing] was the lack of clarity around the date that parents/guardians would take part...If parents had known the date at this point and been given the forms, we may have got more to be interested." – *Withdrawn primary school (Family Numeracy)*

### FSQ Maths Level 1 trial (Contextualised Approach and Adapted Mastery Approach)

The main reason for withdrawal of providers from the FSQ Maths Level 1 (Contextualised Approach and Adapted Mastery Approach) trial related to challenges meeting trial requirements, particularly in relation to tutor training. For treatment tutors on both the Adapted Mastery and Contextualised arms of the trial, training consisted of an initial session and compulsory ongoing 'lesson study' for at least 80% of the 12 to 15 weekly online sessions (totalling 2.5 days or 10 hours per tutor). Survey and interview findings indicated that these requirements were considered too much of a burden for some providers, especially where tutors were employed on a part-time basis. More generally across the trials, 18% of surveyed withdrawn providers also reflected that the 'onerous' nature of tutor training contributed to their withdrawal.

### Embedded Maths

For Embedded Maths, lack of confidence to deliver the intervention was a key reason for withdrawal amongst providers. The learning log indicated that health and social care tutors (who were not specialised in maths) lacked the confidence required to teach the intervention. One provider that was interviewed also reflected on the challenges for health and social care tutors to understand and teach the maths content as part of usual course delivery. This highlights potential challenges for future interventions that embed maths within vocational courses, and implications this may have on delivery and trial attrition.

## Lessons learned: how to mitigate attrition on future trials

### Providing more information earlier

Future trials should focus on providing information as far in advance as possible before the start of the trial. Almost a quarter (23%) of providers and schools who had withdrawn from at least one trial said more lead-in time to prepare may have enabled them to continue in the trial. This was also reported in interviews where providers said they

needed more time to ensure they had sufficient resource in place to manage the trial requirements and deliver the intervention.

> "I think that everything that [the Managed Service Supplier] produced has been really good. It just is all a little bit too late. And I think that's the biggest issue. I would love to be involved in all of it." – *Withdrawn provider (Prep for Maths GCSE)*

Providers would have welcomed more information about payments, the resource commitment involved and how the trial differed from core Multiply provision. The MSS reported that some providers were reluctant to progress to sign up to trials until the payment amount had been confirmed. Future trials should ensure early confirmation of the funding model to avoid early attrition.

## Adapting trials to the adult education sector

Trials involving interventions that could be delivered within the existing curriculum experienced lower levels of attrition once delivery started. For example, Preparation for Maths GCSE involved 7 sessions embedded within existing curriculum hours. Feedback from providers indicated that this made it easy to deliver and the MSS observed that it contributed to less attrition from this trial. Where possible, future trials should consider trialling interventions that can be adapted to fit within the existing curriculum.

Adult education providers do not always follow a standard structure in terms of class times, length of lessons and mode of delivery (online or in person). Attrition from future trials could potentially be minimised by offering a higher degree of flexibility in delivery, while still maintaining fidelity to the intervention. Where possible, future trials should allow a degree of flexibility to ensure interventions can be delivered within the structures of adult education.

> “As I said, we've done research before. We're doing it at the moment. It's something we do look to do, but obviously it just has to work for us." – *Withdrawn provider (Prep for maths GCSE and Functional Skills Qualification)*

## Ensuring sufficient funding is available

Although funding was not the main reason for providers taking part in the trials, it was required to ensure that participation was financially viable. The available funding did not always cover provider costs to take part in the trial, particularly those in the treatment group. This was especially true for providers with small class sizes because the funding model was partly based on the number of learners.

# 7. Primary data collection


## Summary of chapter

This section details the approach to primary data collection across the Multiply trials. It provides an overview of the primary data that was used in the trials, and the tools and processes used to collect this. It includes consideration of what worked well in data collection, as well as challenges faced and approaches to overcoming these. It is based on analysis of data collected through the online data portal, technical reports, trial reports and lessons log.


## Overview

Primary data for the trials was collected through management information (MI) submitted by providers and schools, surveys, interviews, maths assessments and observations. A summary of the primary data used across the trials is presented in Figure 7 and further detail on the data used in each of the individual trials is presented in Figure 8.

**Figure 7: Primary data used in the Multiply trials**

## Figure 8: Overview of data used in each trial

| | Trial Name | Baseline survey | | Endline survey | | Maths assessment | | IPE interviews | | | Missing data |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Treatment *N* *Response rate** | Control *N* *Response rate** | Treatment *N* *Response rate* | Control *N* *Response rate* | Adults ** | Children | Learners | Tutors/ teachers | Providers/ schools | Missing ILR data (primary obj.) |
|  | **Encouraging Progression** *(full-scale effectiveness RCT)* | N/A | N/A | Peer: 247 (5.2%) Urgent: 263 (5.7%) | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
|  | **Family Numeracy** *(Small, pilot RCT)* | 68 31% | 70 35% | 128 58% | 127 64% | N/A | Baseline: 286 Endline: 283 | T: 30 C: 4 | Teacher: T: 7,C:7 Tutors: T:20, C:0 | T: 4 C:0 | NA*** |
|  | **Adapted Mastery** *(Small, pilot RCT)* | 107 23% (est.) | 114 26% (est.) | 62 13% | 137 31% | Treatment Endline: 9 Control Endline: 10 | N/A | T: 30 C:10 | T: 33 C: 10 | T: 17 C: 10 | 339 (37%) |
|  | **Contextualised Approach** *(Small, pilot RCT)* | 80 39% (est.) | 146 28% (est.) | 117 57% | 237 46% | Treatment Endline: 15 Control Endline: 10 | N/A | T: 32 C: 10 | T: 11 C: 10 | T: 10 C: 10 | 122 (17%) |
|  | **Preparation for Maths GCSE** *(Small, pilot RCT)* | 135 44% (est.) | 86 12% (est.) | 113 12% | 331 34% | N/A | N/A | T: 31 C: 10 | T: 15 C: 3 | T: 4 C: 4 | 37 (3.5%) |
|  | **Embedded Maths** *(Implementation and Process Evaluation)* | 30 44% | N/A | N/A | N/A | Treatment Baseline: 23 Endline: N/A | N/A | 14 | 2 | 2 | N/A |

Source: DfE analysis of Adult Numeracy Trial Reports

*Response rates usually reflect the number of learners who completed a survey divided by the number of learners who were invited to take part in the survey x 100. As the baseline surveys for Preparation for Maths GCSE, AMA and CA were distributed by providers via an open link, we cannot say with certainty the total number of learners who received the invitation. As such, the response rate is an estimate based on dividing the number of responses by the number of learners who were in the trial at the time of randomisation x100. The true response rate could be higher or lower.

**The trial protocol included an endline test of maths skills for control and treatment learners. However, this was not able to be delivered as planned. The maths assessments were originally intended to be administered at baseline and endline to measure progress over time. However, due to timing constraints (including

course duration) and low uptake of the optional assessments, this was not possible. The maths assessments were instead administered to examine how the tests could work in practice (“testing the test”) and to capture learning on the administration and validity of the tests as a research tool.

***The primary outcome in the FN trial is defined as progression to a maths related course. Therefore, by definition, this outcome cannot have missingness.

## Provider management information

An online data portal was set up to facilitate the secure transfer of learner and tutor data from participating providers and schools to the evaluation team. It was designed to be user-friendly to minimise burden, enable efficient data collation, perform quality assurance checks at the point of submission and ensure all data transfer was secure and compliant with data protection legislation. It was set up to collect essential personal data, contact details, consent (where relevant), and attendance data for each learner participating in the course-based trials. It also collected details about tutor delivery of sessions (for those in the treatment group). Dedicated fields in the data submission templates captured confirmation that learners had been given the relevant information about the trial and not opted out of their data being shared for the purposes of accessing their Individual Learner Record (ILR) and for being contacted for primary research (surveys and interviews). Table 8 provides an overview of the Management Information (MI) collected.

**Table 8: Overview of MI data**

| Type of records | FSQ L1 Maths Contextualised | FSQ L1 Maths Mastery | Embedded Maths | Family numeracy | Prep for Maths GCSE | Total |
|---|---|---|---|---|---|---|
| Learner records (treatment) | 204 | 494 | 68 | 302 | 321 | 1,389 |
| Learner records (control) | 515 | 515 | [z] | 271 | 737 | 1,523 |
| Tutor records (treatment) | 13 | 16 | 4 | 134 | 17 | 184 |
| Tutor records (control) | 38 | 38 | [z] | See footnote[56] | 47 | 85 |

Source: Ipsos matched data files

FSQ L1 Maths control learners and tutors are only counted once in the totals

## MSS data management

The MSS managed data on provider recruitment, participation, and withdrawals from the trials, whilst also monitoring contractual milestones and storing provider contracts.

As the MSS contract concluded before project completion, all relevant data (including provider recruitment, withdrawals, participation and tutor training attendance) was

[56] Family numeracy tutor data: Tutors were based in provider organisations. 134 tutors were trained in delivering the intervention prior to the randomisation of schools to treatment or control groups. Providers could have schools in both treatment and control groups. Family Numeracy had a waitlist design so that control schools received the intervention after the trial finished, therefore not all tutors who were trained delivered the intervention to the treatment group.

securely transferred to Ipsos for auditing purposes and to inform final analysis and reporting. The MSS Exit Strategy ensured data transfer occurred one month before contract end, allowing time for review.

# Surveys

## Learner surveys

### Baseline survey

A baseline survey was administered to learners on the course-based trials, FSQ Level 1 Maths Contextualised Approach and Adapted Mastery Approach), Family Numeracy and Prep for Maths GCSE, with routing to trial-specific questions. The survey was open between September 2024 and May 2025, with distribution aligned to course start dates. Fieldwork was scheduled to stop at least 8 weeks before course end dates to allow a sufficient window between the baseline and endline surveys.

Providers distributed the baseline survey to learners by giving them trial information sheets with QR codes and weblinks to the survey. These materials were embedded within the Trial Readiness Packs. Providers were asked to facilitate learners to complete the survey in classrooms in the first few sessions of the course. Ipsos followed up with learners who did not complete the survey online to offer them the opportunity to complete it by telephone (CATI). The sample for CATI was drawn from the learner data uploaded by providers to the Ipsos Data Portal, which was cross-checked against those who had completed the survey online. Annex 12 provides a breakdown of responses by trial and mode.

The proportion of completed surveys via each mode (online or CATI) were broadly in line with expectations. Most baseline survey responses were online, suggesting that the mode and distribution plan were the right choices. This also suggests that providers were successfully distributing the open link to the survey to learners in classrooms, and that learners were taking part.

The average time taken to complete the online baseline surveys was 17 minutes. This indicates that most learners likely took the survey in one sitting, which further suggests that providers had successfully asked learners to complete the survey in class.

The open link approach was taken due to the lack of available sample data during the baseline period (as providers had not yet uploaded learner data to the Ipsos Data Portal). When sample started to come through, telephone follow-ups began. This distribution strategy did have some limitations. First, researchers were unable to monitor response rates, identify which learners had completed the survey and take follow up action as needed (e.g. issuing reminders). Furthermore, the survey system could not automatically

integrate responses from the open link with the CATI sample. Instead, a matching process was required to work out who had responded and to create CATI samples, which required additional time. An additional question was needed during CATI fieldwork to confirm if participants had already responded online. Finally, researchers depended on education providers to share the survey link with learners. If the providers did not distribute the link and had not yet uploaded their learner data to the Ipsos Data Portal, researchers had no alternative means to reach those participants.

### Endline survey

The endline learner surveys were open between November 2024 and September 2025, again with variable end dates according to course dates for each trial. There were some variations in the mode of delivery of the endline survey.

The FSQ Level 1 Maths (Contextualised Approach and Adapted Mastery Approach) endline was initially administered as an online survey accessible by an open link. Providers were asked to share the survey with learners on, or close to, the end date of their course via QR codes and weblinks embedded in information sheets provided by Ipsos. Non-responders were followed up by Ipsos and invited to complete the survey via telephone (CATI). Once more sample became available, Ipsos administered the endline survey for learners finishing their courses in summer as a mixed-mode online and telephone survey with sample pre-loaded.

The Preparation for Maths GCSE endline survey could be completed online or by telephone. Unique links were emailed to trial participants with an invitation to take part online. Non-respondents were followed up by telephone and invited to complete a CATI interview.

The Family Numeracy endline survey was delivered in a multimode format (online and telephone). Unique links were emailed to participants to take part online and these were followed up with CATI dialling to non-responders. Ten telephone interviews were administered using live translation services from Language Line for non-English speakers. Two interviews were administered in Urdu, 2 in Kurdish, and one each in Soninke, Tamil, Arabic, Bengali, Romanian and Spanish.

An endline learner survey was conducted for the Encouraging Progression trial. Participants who had received a letter were invited to take part in a short (3 minute) online survey to inform the Implementation and Process Evaluation (IPE).

There were some benefits to using a mixed-mode email and telephone follow-up survey at endline, in contrast to the open link approach adopted at baseline. Firstly, response rates could be monitored in real time and reminder strategies adapted to maximise response rates. This approach also ensured full coverage of the sample, minimising the risk that some did not receive an invitation to complete the survey from providers (e.g. if

they did not attend class on the day it was distributed). Finally, there was less data processing required as the online and CATI data could be collected in the same data file without the need to manually match separate datasets.

Table 9 provides an overview of survey responses collected across the trials. Further details on the response rates can be found in Annex 12.

**Table 9: Learner surveys completed**

| Survey | Encouraging Progression | FSQ L1 Maths Contextualised | FSQ L1 Maths Mastery | Embedded Maths | Family Numeracy | Prep for Maths GCSE | Total |
|---|---|---|---|---|---|---|---|
| Baseline treatment | [z] | 80 | 140 | 35 | 68 | 143 | **466** |
| Baseline control | NA | 141 | 141 | NA | 71 | 101 | **313** |
| Endline treatment | 510 | 117 | 85 | NA | 179 | 143 | **1034** |
| Endline control | [z] | 237 | 237 | NA | 164 | 436 | **837** |

Source: Matched data
FSQ L1 Maths control learners are only counted once in the totals
NA = Not Applicable

## Surveys with providers and schools

Providers and schools who expressed interest in the Multiply trials or signed a contract but later withdrew were invited to take part in a short survey to share feedback on their experiences. The survey was administered between 27 March and 13 June 2025 and lasted approximately 5-10 minutes. Findings from the survey are presented in Section 6: Attrition.

Providers and schools who completed the trials were also invited to take part in a similar short survey about their experiences. This included their reasons for signing up, what they hoped to get out of it and any additional support they would have liked to have received. This survey was administered between 16 September and 9 October 2025 and also lasted approximately 5-10 minutes.

The findings from both surveys were skewed towards Family Numeracy as there was a higher number of responses from providers and schools who took part in that trial, relative to the others.

Due to the low base sizes, neither of the surveys were weighted.

**Table 10: Responses to surveys of providers and schools**

| Survey | FSQ L1 Maths[57] | Embedded Maths | Family numeracy | Prep for Maths GCSE | Total |
|---|---|---|---|---|---|
| Survey of providers / schools who withdrew | 3 | 1 | 65 | 1 | **73** |
| Survey of providers / schools who remained in the trials | 18 | 2 | 34 | 11 | **54** |

Source: Ipsos records

Providers who withdrew from more than one trial are counted once in the totals. Totals also include those who responded but did not specify which trial(s) they had signed up to.

# Interviews

## Interviews with learners, tutors, teachers and stakeholders

Interviews were conducted with learners, tutors, teachers (for Family Numeracy only) and wider stakeholders. Interviews took place via telephone or online, lasting between 20-60 minutes. The majority were completed during spring and summer 2025 when courses were complete or nearing completion. Interviews were mostly conducted individually apart from some stakeholder interviews which were conducted as paired interviews at the request of participants. Learners (only) who participated in an interview received a £30 shopping voucher in recognition of their time and contribution.

Learners were recruited via contact information uploaded to the online data portal. For Embedded Maths, all learners were invited to take part. For the other trials (FSQ Level 1 Maths, Preparation for Maths GCSE) with a larger sample, a purposive approach (i.e. targeted to ensure diversity of interviewee characteristics such as age, gender etc) was used to identify and recruit a diverse range of learners. Tutors were recruited via contact details shared by the MSS or via data submitted to the online data portal. In most cases, one tutor per provider was selected to take part, although in some cases invitations were extended to additional tutors within the same organisations to reach target interview numbers. Stakeholders were recruited via contact details shared by the MSS or via DfE. Table 11 provides an overview the number of interviews completed by trial.

[57] Included 2 trials Contextualised Approach for FSQ Maths Level 1 and Adapted Mastery Approach for FSQ Maths Level 1

**Table 11: Interviews completed**

| Interviewees | FSQ L1 Maths Mastery | FSQ L1 Maths Contextualised | Embedded Maths | Family Numeracy | Prep for Maths GCSE |
|---|---|---|---|---|---|
| Learners: Treatment | 30 | 32 | 14 | 30 | 31 |
| Learners: Control | 5 | 0 | - | 4 | 10 |
| **Learners: Total** | **35** | **32** | **14** | **34** | **41** |
| Tutors: Treatment | 33 | 23 | 2 | 20 | 10 |
| Tutors: Control | 5 | 10 | - | 0 | 3 |
| **Tutors: Total** | **38** | **33** | **2** | **20** | **13** |
| Teachers: Treatment | - | - | - | 7 | - |
| Teachers: Control | - | - | - | 7 | - |
| **Teachers: Total** | **-** | **-** | **-** | **14** | **-** |
| Stakeholders[58]: Treatment | 22 | 10 | 3 | 4 | 4 |
| Stakeholders: Control | 0 | 10 | - | 0 | 4 |
| Stakeholders: N/A | 0 | 4 | - | 11 | 3 |
| **Total** | **22** | **24** | **3** | **15** | **11** |

Source: Ipsos, IES and KCL records

## Interviews with providers who withdrew

Providers who withdrew from one or more trials were invited to take part in an interview about their experience of being recruited and reasons for withdrawing. A total of 10 interviews took place between 16 December 2024 and 17 January 2025, each lasting approximately 30 minutes. Table 12 sets out the number of interviews completed by trial withdrawn from.

[58] Stakeholders included participating adult education providers, representatives from the Managed Service Supplier and the intervention developers.

**Table 12: Interviews with providers who withdrew**

| Trial | Number of Interviews |
|---|---|
| FSQ Maths Level 1 | 7 |
| Embedded Maths | 2 |
| Family Numeracy | 1 |
| Preparation for Maths GCSE | 4 |
| **Total** | **10** |

Source: Ipsos records

*Providers who withdrew from more than one trial are counted once in the total

## Maths assessments

DfE developed a maths assessment based on research and guidance produced by NFER (National Foundation for Educational Research). This was aligned to the levels of learners on the FSQ Level 1 Maths and Embedded Maths trials (i.e. below Level 2). They were originally intended to be administered at baseline and endline to enable an assessment of progress over time. However, due to timings (including course duration) and low uptake of the assessment (which was optional), this was not possible. The maths assessments were instead administered to examine how the tests could work in practice ('testing the test') and to capture learning on the administration and viability of these as a research tool.

Maths assessments were conducted with 54 learners on the FSQ Level 1 Maths and Embedded Maths trials (some learners would have completed both baseline and endline assessments). For Embedded Maths, assessments were conducted with 2 providers in September and December 2024. For FSQ, they were conducted with learners in 3 providers in June and July 2025. Learners who took part in a maths assessment received a £30 voucher in recognition of their time.

## Observations

For Preparation for Maths GCSE, 5 in-person lesson observations were conducted in March and April 2025. For Embedded Maths, one observation of tutor training took place. Researchers who conducted the observations used the AEIOU Observation Framework – tailored to cover the research questions for this trial – to take notes for analysis purposes. For the FSQ Level 1 Maths trial, observations were planned of tutor training

sessions. However, the timing of the sessions and the small group format of these meant that this was not possible.

## What worked well?

Several aspects of the data collection process worked well across the Multiply trials. Based on Ipsos's experience key success factors included the robust technical design of the online data portal, a flexible and multi-channel approach to survey administration, hands-on support for providers, and strong collaboration between the evaluation team, MSS and DfE.

**The online data portal was established to collect learner data from providers and schools was effective in ensuring the secure transfer of high-quality data on learners participating in the trials.** Submission was via a sample template (Excel) which was downloaded by users from the portal, completed, and reuploaded. The portal was designed with a series of hard and soft rules that were applied at the point of data upload. This automated the quality assurance checks, reporting any data quality issues back to providers and schools immediately, and allowing most errors to be corrected at source. The portal also acted as a secure front-end, passing data directly to Ipsos' servers without storing it, ensuring full compliance with data protection rules.

**A flexible, mixed-mode approach to surveys was successful in maximising response rates**. The baseline and endline surveys used a sequential mixed-mode design, where online surveys were issued first, followed by telephone (CATI) calls to non-respondents. A range of techniques were used to boost responses such as offering language interpretation for telephone surveys, sending reminders via SMS text (as well as email) and adjusting survey windows to accommodate different course schedules. This was effective in ensuring as much data as possible could be gathered.

**Intensive and proactive support for providers was critical in overcoming challenges faced by some in sharing data**. The level of support required by providers to submit learner data to the online portal was considerably higher than originally anticipated. This was in part due to confidence and capability to use the portal, including using Excel. Support included a dedicated inbox and phone number, information webinars demonstrating the data upload process, and numerous one-to-one calls. The evaluation team also uploaded learner data on providers' behalf in some instances to overcome occasional issues with the portal or with capacity or capability.

**Strong collaboration and clear governance structures between the MSS, evaluation partners and DfE enabled effective decision-making**. The complexity of the programme required close collaboration between all stakeholders involved. This was consistently highlighted as a key strength of programme delivery, including at all-partner

workshops to reflect on progress and lessons learned. For example, a quantitative working group was established to facilitate collaborative working across evaluation partners and DfE on survey design and approaches to boosting response rates.

## What were the challenges?

**The lack of a complete sample of learners at the start of the trials, varying course start/end dates, and late data submissions created logistical challenges**. Providers were asked to upload learner data to the portal within 2 weeks of courses starting. However, this was not possible for some courses which started early in the academic year as the data portal had not yet been launched. Providers did not always submit learner data in a timely manner and some were still submitting data well into the academic year. This meant that the evaluation team had to take a flexible approach to administering baseline surveys e.g. using multiple channels and keeping the survey open for as long as possible to maximise returns.

**While the data submitted to the portal had to meet hard rules on format (e.g. to include the correct number of digits for telephone numbers), the validity of these details was sometimes variable**. There were higher than expected levels of wrong and invalid numbers leading to unsuccessful calls to learners. The reasons for this are unclear. In a small number of cases, providers submitted incorrect data (e.g. dates of birth made up of all the same characters, schools sometimes did this as they planned to go back and update when they had the information from parents) in order to bypass the hard rules and enable their data to be submitted.

**The compressed programme timeline created challenges for the launch and implementation of the data portal**. The online data portal was launched after some courses had already begun, leaving insufficient time for thorough testing; the original plan to pilot the system with 'friendly' providers using live data was not feasible within the tight timescales. This resulted in several teething issues, such as providers submitting data for ineligible learners, requiring additional validation rules to be added to the system after launch. The model, which placed responsibility for data entry with providers and schools, highlighted capacity issues with some who required additional support. Inconsistent naming conventions for providers and schools created challenges in matching records accurately. Finally, while the portal itself was secure, the process was not without risk; several data security incidents occurred, largely due to providers sending personal data via email rather than using the designated portal.

**Providers had to engage with 2 separate portals for learner data and contract management data, which sometimes created confusion**. Providers were required to use the Ipsos data portal for submission of data for the evaluation, such as learner details and attendance. They had to use a separate portal to submit information and data

required for contract management, compliance, and invoicing. This led to a high volume of communications, which was sometimes confusing and burdensome for providers. Future research should consider one portal for data collection from providers with a secure function to only give organisations access to the data they require.

**The timing of data collection, with long gaps between baseline and endline, created challenges with learner recall and engagement**. Some learners were confused when contacted to take part in an endline survey, especially if they did not recall completing the baseline (54% of learners who completed a CATI endline interview said they had not been asked to complete a baseline). This suggests that the time lag between the start and end of the courses diminished learners' recall of the trial or that communications about surveys failed to reach some learners. This was a particular issue for the Preparation for Maths GCSE trial, which ran for a full academic year.

**Some groups of interviewees were more challenging to arrange interviews with, which meant that additional resources were required to meet targets**. There was a higher-than-usual number of no-shows for learner interviews compared to evaluator experiences of similar trials and contact details for tutors were challenging to obtain (particularly for those in control groups). Alternative methods were used to reach tutors, including arranging these via strategic leads within participating providers. In addition, a relatively high proportion of learners did not have English as a first language (see Annex 12). This was sometimes found to be a barrier to them engaging with surveys and interviews (learner information sheets and surveys were translated or interpreted on request).

# 8. Data management


### Summary of chapter

This section details the approach to data management and linking across the adult numeracy trials. As Central Data Coordinator, Ipsos collected, managed and linked learner data from multiple sources for trial-level analysis. We examine the technical processes used for data matching, what worked well as well as challenges faced. The analysis draws on data collected from the Ipsos Data Portal, Individualised Learner Record (ILR) datasets (ongoing collection of data about learners and the learning undertaken by them from learning providers in the Further Education and Skills sector), survey responses, and technical documentation from trial delivery.


## Approach to data management

Ipsos was the Central Data Coordinator for the adult numeracy trials, overseeing the collection and management of learner data across all trials, working closely with consortium partners (listed in Section 2 Programme Governance) and the Managed Service Supplier (MSS). As Data Controller, DfE determined the purposes and means of processing learners' personal data. Ipsos, consortium partners, and the MSS acted as Data Processors, processing data according to DfE guidance.

DfE established Data Processing Agreements (DPAs) with Ipsos and the MSS. Additional DPAs were established between Ipsos and consortium partners to enable data sharing for trial analysis and reporting. The MSS, responsible for recruitment of providers, issued contracts to providers outlining data processing terms and conditions. Annex 13 shows the data flow process for learner data from providers.

The legal basis for collecting personal data in the adult numeracy trials was Public Task, as the work served the public interest, eliminating the need for individual consent. This meant that data sharing agreements with individual providers were not required. Learners could 'opt out' of having their ILR records linked to provider-submitted data. However, the legal basis for taking part in research activities, such as surveys and interviews, was consent requiring learners to provide explicit consent for these activities.

## Learner data sources used for linking

The trials required linking learner data from multiple sources including management information from providers, surveys and ILR data. An overview of these sources is

provided below, with more information on how the data was collected provided in section 7 data collection:

- provider management information transferred to Ipsos via the online Ipsos Data Portal. Providers shared learner contact details, demographic information and confirmation they had not opted-out of their data being shared
- Ipsos conducted baseline and endline surveys of learners via an open link or direct invites by text, email or telephone
- ILR is an on-going collection of data about learners from training providers in the Further Education (FE) and Skills sector in England ILR data

## Data structure and delivery cycle

Ipsos received ILR data from DfE monthly throughout the academic year (August to July), with each release providing a cumulative 'snapshot' rather than standalone data. Each month's submission therefore built upon and supplemented the previous returns.

The ILR contains detailed records for learners enrolled in further education and skills training. It captures both personal learner information (such as demographic characteristics and contact details) and course information (aims, start and end dates, completion status and grades achieved). DfE delivered this data monthly in the form of 5 Excel files, each capturing different aspects of learner information:

**Aims dataset:** Course enrolment details for each learner, including course name, start and end dates, completion status, and grade achieved. As learners can enrol in multiple courses, this dataset may contain several records per individual.

**Learner dataset:** Demographic and personal characteristics of each learner including postcode, contact permissions, maths grades.

**Learner Named dataset:** Personally Identifiable Information (PII) for each learner, such as phone number and email address.

**Prior Attainment:** Historical qualification levels achieved by each learner before their current course of study.

**Employment Status:** Current employment situation of each learner (for example in paid employment, seeking work) along with the date this status was recorded.

Ipsos combined the 5 ILR files using the unique learner numbers (ULNs) present in each dataset.

### Data management

Ipsos created trial-specific merged datasets by combining learner data from 3 sources: the Ipsos Data Portal, survey responses, and ILR records. Ipsos then shared relevant data extracts with each trial consortium partner, including only information for their trial participants, preventing any sharing of extraneous data.

## Data linkage

For the 4 course-based trials[59] and the Family Numeracy trial, Ipsos linked learner data collected from the Ipsos Data Portal with survey and ILR data. Data matching used a hierarchical ("waterfall") system to maximise learner matches:

- exact matches records linked through perfect matching of selected data keys
- fuzzy matches where exact matching failed, algorithms identified and linked highly similar but non-identical records (such as those with minor spelling errors). Fuzzy matching used name, postcode and date of birth. These probable matches were manually reviewed for accuracy before being accepted

Combining exact and fuzzy matches significantly improved the linkage rate between core learner records and survey data, overcoming issues of minor data entry variations.

The conceptual idea behind the data linkage was to create a data "warehouse" environment. This involved preparing and standardising disparate datasets (the Ipsos Data Portal, ILR and surveys) and then merging them to create a single, consolidated master record for each trial participant. This unified dataset enabled robust and efficient analysis by consolidating all learner information in one place. The final merge identified 2,663 out of 2,855 learners in the ILR dataset, achieving a 93.3% match rate. Further details are provided in Annex 13.

### Process for linking the data

The step-by-step process for linking the data is set out below.

1. Preparation and Normalisation: All data was cleaned and standardising, including alignment of column names across files, reformatting of birth dates into a consistent format (`YYYYMMDD`), and cleaning of personal information (postcodes, email addresses, and phone numbers) through removal of spaces and standardisation of capitals.

[59] These were Preparation for Maths GCSE, Contextualised Approach FSQ Maths Level 1 and Adapted Mastery Approach for FSQ Maths Level 1 and Embedded Maths

2. Creating an audit trail: Key matching columns were duplicated and labelled with their source (e.g., ilr_dateofbirth), ensuring each data point's origin could be traced after datasets were combined.

3. Finalising participant lists: Trial participant lists were refined through removal of test records, withdrawn learners, and duplicate entries, creating one definitive record per learner per trial.

4. Linking datasets: Data was linked separately for each trial, with cleaned participant lists linked with the ILR, survey, and achievement datasets using a systematic matching approach.

5. Matching process: Matched methods were prioritised from most to least reliable, ensuring the best matches were identified first. Multiple data points were combined to create unique identifiers for each learner.

The following describes the different types of keys used for matching records:

- highest confidence keys used the most robust and unique information available for direct matching (Unique Learner Number + Date of Birth + Postcode)
- fallback keys used personal information when a match could not be found with a higher-confidence key (Full Name + Date of Birth + Postcode)
- hybrid keys used for fuzzy matching as a last resort, incorporating outcomes from the fuzzy matching process to link survey data (combining either the original data (for exact matches) or the corrected data from verified fuzzy matches)

## Ipsos Data Portal match results to ILR

Table 13 shows matched and unmatched learners between the Ipsos Data Portal and ILR datasets. Of the 2,912 total trial participants, 57 took part in 2 trials, leaving 2,855 unique learners for ILR matching. As outlined above, matching occurred in 2 stages. First, high-confidence keys directly matched 2,099 learners. Second, fuzzy matching on name, postcode, and date of birth produced an additional 360 matches. This left 396 learners without ILR records.

Most unmatched learners (342) came from the Family Numeracy trial, including 204 control group learners not expected in the ILR dataset as they had not participated in learning. Excluding these, the final attrition rate was 6.7% (192 out of 2,855). Among these 192 unmatched learners, 138 were from the Family Numeracy treatment group, with 108 never attended any sessions. The high proportion of unmatched Family Numeracy participants (342 out of 573) likely reflects the trial design, parents enrolled on courses before delivery dates were confirmed, with many ultimately not attending.

**Table 13: Matched learners between Ipsos Data Portal and ILR datasets**

| Trial | Matched | Unmatched | Grand Total |
|---|---|---|---|
| Contextualised Curriculum | 202 | 2 | 204 |
| Embedded Maths | 67 | 1 | 68 |
| Family Numeracy | 231 | 342 | 573 |
| FSQ Level 1 | 503 | 12 | 515 |
| Maths Mastery | 489 | 5 | 494 |
| Preparation for Maths GCSE | 1,024 | 34 | 1,058 |
| Total | 2,516 | 396 | 2,912 |

Source: Data linking specification and the data matched file outputs

## Ipsos Data Portal match results to survey (including fuzzy matching)

Table 14 shows the total number of matched and unmatched learners between the Ipsos Data Portal and Survey datasets.

**Table 14: Matched learners between Ipsos Data Portal and survey datasets**

| Trial | Matched | Unmatched | Grand Total |
|---|---|---|---|
| Contextualised Curriculum | 153 | 51 | 204 |
| Embedded Maths | 42 | 26 | 68 |
| Family Numeracy | 376 | 197 | 573 |
| FSQ Level 1 | 304 | 211 | 515 |
| Maths Mastery | 175 | 319 | 494 |
| Preparation for Maths GCSE | 690 | 368 | 1,058 |
| Total | 1,740 | 1,172 | 2,912 |

Source: Data linking specification and the data matched file outputs

### Data quality, transparency and integrity

Several key decisions shaped the data management process to ensure quality, transparency, and integrity:

**Standardisation rules:** All data columns received standard names and uniform formatting for dates, postcodes, and phone numbers to enable accurate matching.

**Master identity resolution:** A master lookup list resolved inconsistent learner IDs in the source data. Each learner's definitive ID was taken from their most recent record.

**Data source prioritisation:** A priority system handled instances conflicting information when joining datasets. When learners had different values for the same field across sources, the system automatically selected values from the higher-ranked source and flagged these decisions.

**Traceability:** Prefixed copies of all key matching fields were created, ensuring the original of all data in the final merged records remained easily traceable.

## What worked well?

**The monthly provision of ILR data enabled familiarisation with the data structure.** Trial data linking across the 3 datasets using earlier ILR versions helped formulate queries and assess data quality and completeness. This saved considerable time for the final merge and code development when the final ILR dataset arrived.

**The DfE ILR team quickly resolved queries about the ILR data.** DfE's website[60] provides detailed information about each ILR variable, supporting data interpretation and analysis.

**The foundational work of data normalisation and standardisation was crucial for successful data linking**. Although resource-intensive, this preparatory phase prevented errors and ensured highly accurate automated matching.

**The hierarchical matching strategy, from high-confidence to fallback keys, provided a robust and thorough method for linking records,** achieving high learner match rates.

**Fuzzy matching effectively increased survey data matches** by resolving minor typos and formatting issues, ensuring the data could be fully utilised.

[60] ILR specification - https://guidance.submit-learner-data.service.gov.uk/25-26/ilr/overview

**Creating a clear audit trail by labelling data sources and flagging automated decisions** (such as prioritising one data source over another) ensured complete transparency.

## What were the challenges?

**ILR data quality depends on providers submitting accurate returns.** The DfE, MSS, and Ipsos invested significant resources in contacting course-based providers to encourage ILR completion and ensure the correct course identifiers were used.

**DfE introduced a new code (LDM394) to identify learners on the adult numeracy trials.** However, not all providers included this information in their submissions, limiting its effectiveness.

**Temporary Unique Learner Numbers (such as "9999999999") in the final data return (R14) present significant challenges for longitudinal analysis.** Learners with temporary ULNs cannot be reliably tracked across academic years, particularly when transferring to new providers (UKPRN) and receiving new Learner Reference Numbers.

**ILR data varies between funding streams and academic years,** meaning that subsequent ILR versions cannot be assumed to be consistent with earlier iterations.

**The Encouraging Progression trial analysis faced delays due to changes in adult maths provision coding within the ILR.** Additional interpretation work was needed once the data arrived, impacting reporting timelines.

**Varying column names and data formats** across sources complicated data handling.

**Duplicate records and varying IDs for the same learners** complicated data management.

## Data management and linking conclusions and recommendations

The data management and linking approach for the adult numeracy trials achieved considerable success. The hierarchical matching strategy, combining exact and fuzzy matching techniques, effectively overcome minor data entry variations and significantly increased linkage rates.

Several challenges remain for future trials. These include ensuring early provider engagement regarding data submission requirements and establishing standardised data formats across all sources from the outset. Evaluation timelines should incorporate contingency for potential ILR coding changes and data interpretation challenges.

Future trials should retain the successful elements from the current approach, including the hierarchical matching strategy, fuzzy matching, and clear documentation of data decisions, which all contributed to the high match rates achieved.

# 9. Trial findings


### Summary of chapter

This section reports on the findings from 6 adult numeracy intervention trials delivered through the Multiply programme. The trials comprised Encouraging Progression, 4 course-based interventions (Embedded Maths, Contextualised Approach, Adapted Mastery Approach and Preparation for Maths GCSE) and Family Numeracy. For each trial, impact evaluation findings relating to primary and secondary outcomes are reported, alongside implementation and process evaluation evidence examining delivery compliance, fidelity, feasibility and participant experiences. The section concludes with a synthesis comparing effectiveness and implementation factors across all 6 trials, drawing on individual trial reports produced by independent evaluators.


## Overview of interventions

**Encouraging Progression**: this communication intervention sent targeted letters to adult learners who had completed a non-qualification-bearing Multiply maths course. Participants received 1 of 2 letter types: a 'peer support' letter featuring a quote from another learner about increased confidence, or an 'urgent' letter emphasising the career benefits of improved numeracy skills. Both aimed to encourage their progression to qualification-bearing maths courses.

**Embedded Maths:** this intervention integrated 12 one-and-a-half-hour maths lessons into existing Level 2 Health and Social Care vocational courses. The curriculum covered contextualised Functional Skills Level 1 maths concepts including medication calculations, budgeting, measurements and data interpretation, designed to build learner confidence and competence in workplace-related numeracy tasks.

**Contextualised Approach for Functional Skills Qualification (FSQ) Level 1 Maths**: This intervention used Realistic Mathematics Education (RME) principles across 12 one-and-a-half-hour lessons within existing FSQ Level 1 maths courses. The RME approach introduces real-world contexts before mathematical concepts, covering number, measure, shape and space, and data. It aimed to improve learners' attitudes towards maths, their understanding of its practical relevance, and their attainment in FSQ Level 1 maths exams.

**Adapted Mastery Approach for Functional Skills Qualification (FSQ) Level 1 Maths**: this intervention delivered 15 one-and-a-half-hour lessons within existing FSQ Level 1 maths courses, applying mastery principles to develop deep understanding of

mathematical structures. The approach emphasised building on prior learning, curriculum coherence and connections, and developing fluency alongside understanding. It also fostered a culture where everyone could succeed and aimed to support learners' achievement in FSQ Level 1 maths examinations.

**Family Numeracy:** This intervention provided 6 one-hour sessions in primary schools for parents/carers aged 19+ with children in Key Stage 1 Year 2. Sessions covered number, shape/space/measure, and statistics through practical activities involving parents and children working together, supplemented by home learning tasks. The programme aimed to improve parents' maths confidence and their ability to support their children's maths learning whilst addressing barriers to adult maths participation.

**Preparation for Maths GCSE:** This intervention comprised 7 one-hour lessons integrated into the existing GCSE curriculum for adult learners. Content included theories of resilience and growth mindset, alongside study skills, revision techniques and exam strategies, all designed to support learners' preparation for GCSE Maths exams.

## Primary and secondary outcomes

Table 15 details the primary and secondary outcomes measured in each trial.

**Table 15: Primary and secondary outcomes**

| Trial | Primary outcomes | Secondary outcomes |
|---|---|---|
| **Encouraging Progression** | Started a qualification-bearing Maths course at any level within 7 months of letter dispatch | Started any learning likely to result in a qualification<br>Started a maths course at Level 2 or above |
| **Contextualised Approach** | Functional Skills Qualification (FSQ) Level 1 pass rate | Attendance<br>Confidence in maths<br>Maths skills |
| **Adapted Mastery Approach** | Functional Skills Qualification (FSQ) Level 1 pass rate | Attendance (not reported on)<br>Confidence in maths<br>Maths skills (not reported on) |

| Trial | Primary outcomes | Secondary outcomes |
|---|---|---|
| **Family Numeracy** | Progression to further maths courses | Confidence in maths<br>Progression to: any courses; any qualification bearing courses; any qualification bearing maths courses; any Level 2 qualification bearing maths courses<br>Child's maths attainment |
| **Preparation for Maths GCSE** | GCSE maths course grade | Level 2 Pass (Grade 4+) assessments<br>Course completion<br>Confidence in maths |

Source: Trial level reports

As Embedded Maths was an IPE only, it did not have primary and secondary outcomes. Instead, short and medium-term outcomes were identified. These are summarised in Table 16.

**Table 16: Short and medium-term outcomes**

| Trial | Short-term outcomes | Medium-term outcomes |
|---|---|---|
| **Embedded maths** | **Tutors:**<br>• Improved maths knowledge<br>• Adopt a maths lens, exposing learners to varied maths approaches<br>• Promote a positive mindset<br>**Learners:**<br>• Able to identify the maths in H&SC curriculum and workplace.<br>• Improved mindset and attitudes to maths<br>• Appropriate maths skills and knowledge to complete vocationally related maths tasks, including problem solving. | **Tutors:**<br>• Increased confidence in own ability<br>**Learners:**<br>• Able to apply maths in a H&SC vocationally relevant test to Level 1 FSQ<br>• Increased maths confidence<br>• Feel more capable to apply maths in a H&SC work context. |

Source: Embedded Maths Report

# Trial evaluation findings

## Encouraging Progression

### Impact evaluation findings

The Encouraging Progression intervention did not have a statistically significant[61] effect on the primary outcome – whether learners started a qualification-bearing maths course within 7 months of receiving a letter. The peer support letter showed a very small positive effect on course enrolment, while the urgent letter showed a slight negative effect, but neither difference was large enough to rule out chance when compared to the control group who received no letter. Although this trial had a much larger sample as it was a full trial (13,638 learners analysed), communication interventions typically produce very small effects. Even this substantial sample may not have been sufficient to detect modest but meaningful impacts.

The intervention also showed no statistically significant effects on the secondary outcomes. Neither letter increased the likelihood of learners starting any qualification-bearing course; both intervention groups actually showed slightly lower progression rates than the control group. For progression to Level 2 or above maths courses specifically, the peer support letter showed a small positive effect while the urgent letter showed a negative effect, but neither difference was statistically significant.

However, exploratory subgroup analysis revealed an important finding: both letters had a statistically significant positive impact for learners aged 50 or more. This older age group showed notably higher progression rates after receiving either letter, suggesting the intervention may be effective for specific populations even if not universally successful.

### Implementation and process evaluation findings

The implementation of the letter intervention was straightforward and cost-effective. Letters were successfully dispatched to all 11,114 learners in the intervention arms. However, the intervention faced significant challenges with learner engagement and recall. Only around 5.5% of learners with email addresses responded to the follow-up survey, and of those who responded, just over one-third (36-37%) remembered receiving the letter they had been sent 8 months earlier.

Among the small proportion of learners who recalled receiving the letters, engagement was reasonably positive. The majority (86%) of those who remembered the letter reported opening it, and nearly all who opened it read the contents. However, the letters had minimal influence on learners' subsequent behaviour. Of those who remembered receiving a letter and later enrolled in a further course, only a handful felt the letter had

[61] For all trials when referring to statistical difference this is at the 95% confidence interval.

positively influenced their enrolment decision. This low recall rate and limited perceived influence aligns with the lack of statistically significant impacts found in the trial.

The trial successfully demonstrated the feasibility of conducting randomised experiments to test different communication approaches with adult learners. The intervention required no ongoing delivery or provider involvement beyond initial data sharing, making it highly scalable. However, the findings suggest that single postal letters may be insufficient to change adult learner behaviour. With nearly two-thirds of survey respondents having no recollection of receiving the communication, more frequent or multi-modal approaches may be necessary to achieve meaningful engagement and impact.

## Embedded Maths

### Impact evaluation findings

The intervention was originally designed for a larger-scale efficacy trial (RCT) but faced severe recruitment and retention challenges. Of the more than 50 providers that initially showed interest, only 21 signed up to deliver the intervention, and 19 of these subsequently withdrew, leaving only 2 providers participating (68 learners in total). This level of attrition made an RCT unfeasible.

### Implementation and process evaluation findings

The 2 providers that completed the intervention reported positive experiences, despite severe attrition that saw 19 of 21 recruited providers withdraw before or during delivery. The participating tutors valued the intervention's practical application to Health and Social Care roles and successfully delivered the content despite initial concerns about teaching maths. One tutor reported that the intervention "completely changed" their confidence in teaching maths, while the other adapted their practice to deliver maths content that related to learners' daily routines rather than abstract formulas. Learners engaged positively with the embedded content, with interviewees reporting improved confidence in workplace tasks such as medication calculations and patient monitoring.

The 2 participating tutors found the intervention materials and support effective. The comprehensive 10-hour pre-training programme and ongoing reflective practice sessions (approximately 2 hours per content area) provided structured guidance through a weekly Padlet[62] and PowerPoint presentations with detailed instructions. While one tutor initially viewed the intervention as "extra hassle," the training changed their perspective by demonstrating its value. The resources were described as allowing tutors to be "spoon fed to deliver", which eased delivery and minimised additional preparation. Learner engagement was strong, with participants particularly valuing the practical application to

[62] Padlets are visual collaboration tools for creative work and education

healthcare tasks, though some struggled to see the relevance of certain topics such as fractions and room decoration exercises.

Compliance with delivery requirements was generally successful among the 2 participating providers, though data sources showed discrepancies regarding training attendance, with management information and tutor interviews providing conflicting figures. Learner attendance at the Embedded Maths sessions was high, averaging 92% at one provider and 80% at the other. However, contamination was a significant issue – 95% of learners at one setting also enrolled in the Adapted Maths Approach Pilot RCT, while in the other setting, 15% of learners also enrolled in the Adapted Maths Mastery Pilot RCT and 23% in the Preparation for Maths GCSE Pilot RCT. This overlap made it impossible to attribute outcomes specifically to Embedded Maths.

The trial experienced substantial attrition, with only 2 of 21 contracted providers completing delivery. Reasons for withdrawal included insufficient resources (4 providers), challenges recruiting tutors (3 providers), challenges recruiting learners (3 providers), and lack of tutor confidence in delivering maths content. Maths anxiety among Health and Social Care tutors proved more challenging than anticipated, with some tutors remaining "completely bewildered" about delivering maths content despite attending 80% of the training.

## Contextualised Approach

### Impact evaluation findings

The Contextualised Approach intervention did not have a statistically significant effect on the primary outcome of FSQ Level 1 maths pass rates. Learners who received the intervention had a lower pass rate than the control group (45% vs 49%), though this difference was not statistically significant. This remained true after adjusting for learner characteristics including age, sex, ethnicity, employment status and deprivation.

However, these findings require cautious interpretation. The pilot trial was underpowered to detect small-to-moderate effects, with a final sample size lower than planned (204 learners in treatment and 515 learners in control[63]). The wide confidence intervals mean that it cannot definitively be concluded that the intervention had no effect; rather, any potential impact was too small to detect given the sample size.

The intervention also showed no statistically significant effect on the secondary outcome of maths confidence. Results improved slightly when analysis was restricted to compliant

[63] The gateway design in the FSQ Level 1 Maths trial allowed providers to participate in either the full 3-arm trial (Contextualised Approach, Adapted Mastery Approach, and control) or a 2-arm version excluding one intervention if they had prior engagement with it or strongly opposed it. As a result, control group learners came from both 3-arm and 2-arm trial sites, creating slight differences in control group composition between Contextualised Approach and Adapted Mastery Approach.

learners (those attending at least 80% of sessions) but remained below significance thresholds. Similarly, subgroup analyses revealed no evidence of differential impacts across demographic characteristics including sex, age, ethnicity, deprivation and health/disability status.

### Implementation and process evaluation findings

Providers, tutors and learners participating in the trial reported generally positive experiences. Learners expressed strong satisfaction with the programme, with 93% of treatment learners being satisfied with their course overall, comparable to the 92% satisfaction rate in the control group. Treatment learners particularly appreciated the real-world contexts, visual elements, and group work approaches, which increased their confidence and problem-solving capability.

Tutors demonstrated moderate to high compliance with training requirements, with 11 of the 14 tutors completing the full pre-training and attending at least 80% of the continuing professional development sessions.

The intervention was well-received by treatment group tutors despite initial concerns about workload. The training was praised as high-quality, well-paced and interactive, with tutors particularly valuing the structured lesson plans and clear, comprehensive teaching resources. Collaborative lesson study sessions provided valuable opportunities for peer discussion and professional development, helping tutors build confidence in delivering the learner-led facilitation approach.

Treatment group learners demonstrated reasonable engagement, with 59% attending at least 10 of the 12 sessions (the minimum dosage requirement). Attendance challenges stemmed from personal barriers such as work and family commitments rather than disengagement with the intervention itself.

The trial experienced relatively high attrition, with 42% of providers and 50% of tutors withdrawing before or during early delivery. Common reasons included late notification of training dates, lack of staff capacity, and insufficient learner recruitment. However, those who progressed to delivery showed good fidelity to the intervention design, implementing core principles whilst making minor adaptations for pacing and learner needs. The intervention proved broadly feasible, though challenges emerged, including the need for more complete FSQ curriculum coverage and better support for ESOL learners. Tutors planned to embed aspects of the approach into their long-term practice, suggesting potential for sustained impact beyond the trial period.

## Adapted Mastery Approach

### Impact evaluation findings

The Adapted Mastery Approach intervention did not have a statistically significant effect on the primary outcome of FSQ Level 1 maths pass rates. This remained true after adjusting for learner characteristics including age, sex, ethnicity, employment status and deprivation.

However, these findings require cautious interpretation. This pilot trial was underpowered to detect small-to-moderate effects, primarily due to lower than anticipated recruitment (470 learners in treatment and 444 in control[64]). The wide confidence intervals mean that it cannot definitively be concluded that the intervention had no effect; rather, any potential impact was too small to detect.

The intervention also showed no statistically significant effect on the secondary outcome of maths confidence. Results were mixed: the single confidence measure showed a negative but insignificant effect, whilst the composite confidence measure showed a nearly significant positive effect (p=0.066). When restricted to compliant learners (those attending at least 80% of sessions), results improved slightly but remained statistically insignificant. Subgroup analyses revealed no differential impacts across demographic characteristics including sex, ethnicity and health/disability status. Secondary outcome analysis was particularly limited by low survey completion rates, with only 13% of treatment learners and 31% of control learners completing the endline survey.

### Implementation and process evaluation findings

Providers, tutors and learners participating in the trial reported generally positive experiences. Tutors expressed strong engagement with the intervention and largely met delivery requirements, though some struggled with adapting materials for online delivery or different learner skill levels. Treatment group learners reported high satisfaction with their courses, frequently citing increased confidence, greater willingness to attempt maths problems, and improved ability to apply maths in everyday situations. Both treatment and control learners reported similarly high satisfaction rates with teaching quality and course content.

Treatment group tutors welcomed the intervention despite initial concerns about workload. The structured lesson plans and ready-made resources reduced preparation time and were praised for their clarity and quality. The 4 2.5-hour online training sessions

[64] The gateway design in the FSQ Level 1 Maths trial allowed providers to participate in either the full 3-arm trial (Contextualised Approach, Adapted Mastery Approach, and control) or a 2-arm version excluding one intervention if they had prior engagement with it or strongly opposed it. As a result, control group learners came from both 3-arm and 2-arm trial sites, creating slight differences in control group composition between Contextualised Approach and Adapted Mastery Approach.

were considered high quality and interactive, though some felt they focused too much on practical delivery rather than underlying mastery concepts. Weekly lesson study sessions provided valuable opportunities for peer discussion and professional development, but many found the time commitment challenging. Treatment learners demonstrated reasonable engagement, with around half meeting the minimum attendance requirement. Many particularly valued the visual elements, interactive sessions and confidence-building approaches.

Compliance with training requirements was high among tutors who delivered the intervention, with most attending all initial training sessions and the majority of lesson study sessions. However, the trial experienced substantial attrition. A total of 6 treatment providers (32% attrition rate) and 7 control providers (37% attrition rate) withdrew after randomisation, primarily citing lack of resources, staff turnover, or insufficient learner numbers. Some providers also considered the training requirements too burdensome, particularly for part-time staff.

The intervention demonstrated good fidelity overall, with most tutors delivering all 15 sessions as intended, though many made minor adaptations for learners' varying needs, delivery mode, and gaps in curriculum coverage. The approach proved broadly feasible, with many providers planning to continue or expand delivery of the approach, despite identified barriers including curriculum alignment (such as examinable curriculum missing from the intervention content), needing to adapt materials for different learner levels, engaging learners when delivering online, and the time commitment required for weekly lesson study sessions.

## Family Numeracy

### Impact evaluation findings

The Family Numeracy intervention had a statistically significant effect on the primary outcome of parent/carer progression to further maths courses. A total of 22 parents/carers in the treatment group (10%) progressed to a further maths course compared to none in the control group. After adjusting for learner characteristics and provider differences, the treatment effect remained statistically significant, with parents/carers in the treatment group being 17 percentage points more likely to progress than those in the control group. The effect was stronger among parents/carers who completed at least half the sessions and home learning tasks, though weaker for those experiencing ill health. As a pilot trial with a relatively small sample, these findings should be interpreted with caution (222 parents in the treatment group and 199 parents in the waitlisted control group).

The intervention showed mixed results for secondary outcomes. Parents'/carers' confidence in using maths in everyday life increased significantly, though this finding was

based on very limited data (only 25% survey response rate), making it difficult to generalise. The intervention did not significantly affect other secondary outcomes, including progression to non-maths courses, progression to qualification-bearing courses, or children's maths attainment. For example, 15% of the treatment group progressed to any further course compared to 11% of the control group, although this difference was not statistically significant. The study had limited power to detect small-to-moderate effects for these secondary outcomes, and low survey response rates undermined the reliability of confidence measure findings.

### Implementation and process evaluation findings

Providers, tutors and parents/carers participating in the trial reported generally positive experiences. Providers expressed enthusiasm about the importance of family learning and successfully met trial data requirements, though submissions were delayed. Tutors found the Family Numeracy course fitted well within school settings and appreciated the integrated learning approach bringing parents/carers and children together. Parents main motivation to join the course was to be able to better support their child's learning. Parents were overall very satisfied with the course, and they found the approach and format engaging.

The pre-delivery training was considered useful and comprehensive for tutors. Delivered through a 2.5-hour online session (with recordings available), it achieved high compliance with 92 out of 95 tutors attending. Tutors found the weekly lesson plans, PowerPoint slides and instructional videos valuable for maintaining consistent delivery. Parents/carers showed moderate engagement, with 34% completing the minimum dosage by attending at least half the sessions and completing half the home learning tasks. The main barriers to attendance were childcare commitments and other personal circumstances.

The trial experienced significant attrition during recruitment, with 51 providers and 264 schools withdrawing because they could not recruit enough parents/carers to make courses viable. However, once courses began, the intervention demonstrated high fidelity. All 6 sessions were delivered as intended in terms of timing and content, though tutors made appropriate adjustments for participants' needs and abilities. Schools found the intervention feasible to deliver, typically scheduling sessions around drop-off or pick-up times to maximise attendance.

The intervention appears ready for scaling to a full efficacy trial. Schools reported increased parental engagement both with their children's learning and the school more broadly, which motivated continuation of family learning approaches. Future delivery would benefit from simplified recruitment materials, confirmed course timings shared earlier and additional time for parent/carer recruitment.

## Preparation for Maths GCSE

### Impact evaluation findings

The Preparation for Maths GCSE intervention did not have a statistically significant effect on the primary outcome of GCSE Maths grades. This remained true after adjusting for learner characteristics including age, sex, employment status, deprivation and prior attainment.

However, these findings require cautious interpretation. The pilot trial was underpowered to detect small-to-moderate effects, primarily due to lower than anticipated recruitment (278 learners in treatment and 714 learners in the control group). The wide confidence intervals mean that it cannot definitively be concluded that the intervention had no effect; rather, any potential impact was too small to detect.

The intervention also showed no statistically significant effects across any of the 3 secondary outcomes. Learners who received the study skills lessons had a marginally higher completion rate than those in the control group, but this difference was not statistically significant. The Level 2 pass rate (achieving grade 4 or above) was actually lower for the treatment group, though this too lacked statistical significance. Maths confidence showed negligible differences between groups, though this finding was compromised by low survey completion – only 25% of learners completed the pre-survey and 42% post-survey. As with the primary outcome, the trial lacked sufficient power to detect realistic effect sizes for these secondary outcomes.

### Implementation and process evaluation findings

Providers, tutors and learners participating in the trial reported generally positive experiences. Providers expressed confidence in delivering the intervention and met trial data requirements, despite initial challenges faced by some in getting data into the correct format. Treatment group tutors found the course fitted well within their existing teaching commitments and curriculum delivery. Of those learners who were aware they were participating in a trial, 85% were satisfied with their experience.

Treatment group tutors particularly welcomed the intervention. The structured guidance and prepared materials required minimal additional preparation by tutors and helped formalise teaching of the topics covered. The 2-hour online webinar training was considered accessible and non-burdensome, fitting easily around existing teaching commitments. The growth mindset session proved particularly valuable, giving learners opportunities to reframe negative experiences with maths and build confidence. Treatment group learners demonstrated good engagement across all sessions, with 88-94% finding the content useful and relevant to their learning needs.

All participating providers successfully submitted their learner data, though most were delayed in doing this. The trial experienced moderate attrition, with 13 providers withdrawing – 7 before randomisation due to lack of interest or capacity, and 6 after randomisation, primarily citing insufficient tutor or learner numbers.

The intervention demonstrated high fidelity, with sessions delivered as intended in volume and timing. Tutors adapted content length and combined materials to fit the existing curriculum but maintained core delivery requirements. The approach proved feasible, with both tutors and learners reporting that the study skills sessions felt like a natural part of their GCSE Maths course rather than an additional burden.

# Implementation and Process Evaluation Matrix

## Figure 9: Implementation and process evaluation matrix

| | Trial Name | Compliance | | Fidelity – delivery as intended | Resource requirements | Scalability assessment |
|---|---|---|---|---|---|---|
| | | Tutor training attendance | Learner session attendance | | | |
|  | **Comms Progression** *(full-scale effectiveness RCT)* | N/A | N/A | Letters sent 25th-28th March 2024 as planned; both versions delivered as designed | Low – DfE has access to ILR data, so only costs would be letter design, printing and postage. | Highly scalable |
|  | **Family Numeracy** *(Small, pilot RCT)* | 97% attended the pre-delivery training | 34% of parents/carers completed at least 50% of sessions and at home tasks | All 6 sessions delivered as intended, with minor adaptations for individual learner needs | Low levels of training (2.5-hours, before delivery) and minimal additional preparation needed. Significant admin resource needed to set up sessions between primary schools and adult numeracy providers and to recruit parents | Feasible with improvements to recruitment of learners and linking schools and providers |
|  | **Adapted Mastery** *(Small, pilot RCT)* | 83% attended full pre-delivery training sessions; 69% attended 8+ CPD sessions | 50% learners attended at least 12 sessions (80% of intervention) | 95% of tutors delivered all 15 sessions with minor adaptations for learner needs | High levels of training and ongoing support (10 hours initial training plus 24 hours weekly lesson study sessions). | Feasible with minor adaptations for online delivery and reduced training burden |
|  | **Contextualised Approach** *(Small, pilot RCT)* | 79% attended full pre-delivery training; 83% attended 8+ of CPD sessions | 59% learners attended at least 10 sessions | 92% of tutors delivered all 12 sessions with minor adaptations for learner needs | High levels of training and ongoing support (10 hours initial training plus 19 hours weekly lesson study sessions). | Feasible with fuller curriculum coverage and reduced training burden |
|  | **Preparation for GCSE Maths** *(Small, pilot RCT)* | 100% attended the pre-delivery training | 67% average learner attendance per session | All 7 sessions delivered as intended with minor adaptions to fit the curriculum | Low levels of training (2 hours, before delivery) and minimal additional preparation needed. | Feasible, though product developers recommend longer initial training |
|  | **Embedded Maths** *(Implementation and Process Evaluation)* | 100% of those who delivered the intervention attended the pre-training and ongoing support sessions | Average 82% attendance across sessions; ranged from 69% to 94% per session | All sessions delivered as intended, with minor adaptations to support learners | Medium level of training (10 hours, before delivery) but this didn't work for some tutors, so may need to be higher. | More development work needed, including adaptation of the approach to training to address H&SC tutor maths anxiety and tutor burden. |

Source: DfE analysis of Adult Numeracy Trial reports

## Evidence and cost ratings

As well as the impact and IPE findings, each trial has been rated according to its cost and evidence. Cost ratings are qualitative judgements, based on the number of training hours needed and other costs to deliver the intervention e.g. printing and postage. Evidence ratings are judgements based on the Nesta Standards of Evidence[65], which consider the level of observed impact, the extent to which that impact can be attributed to the intervention being tested, the degree to which impact has been independently validated and whether the intervention could be replicated and scaled at reasonable cost. Ratings are summarised below in Table 17.

**Table 17: Summary of cost and evidence ratings**

| Trial name | Cost rating (on a scale of 0 to 3) | Evidence rating (on a scale of 0 to 3) |
|---|---|---|
| **Encouraging progression** *(full-scale effectiveness RCT)* | 1 | 3 |
| **Family numeracy** *(small, pilot RCT)* | 3 | 2 |
| **Adapted mastery approach** *(small, pilot RCT)* | 2 | 2 |
| **Contextualised approach** *(small, pilot RCT)* | 2 | 2 |
| **Preparation for maths GCSE** *(small, pilot RCT)* | 1 | 2 |
| **Embedded maths** *(implementation and process evaluation)* | 2 | 1 |

## Summary of trial findings

### Impact findings

Across the 6 adult numeracy interventions tested, only Family Numeracy achieved a statistically significant positive effect on its primary outcome, with parents/carers in the treatment group 17 percentage points more likely to progress to further maths courses compared to those in the control group. The other 4 interventions with impact

[65] Standards of Evidence | Nesta

assessments (Encouraging Progression, Contextualised Approach, Adapted Mastery Approach, and Preparation for Maths GCSE) showed no statistically significant effects on their primary outcomes. However, exploratory subgroup analysis for Encouraging Progression found a statistically significant positive impact for learners aged 50 or more.

Importantly, the remaining course-based trials were all underpowered meaning they could not detect small-to-moderate effect sizes due to small sample sizes and although Encouraging Progression was a full trial it was underpowered to detect the small effects likely from letters. Therefore, while no statistically significant impacts were found, it cannot be definitively concluded that the interventions had no effect. Any potential impacts may have been too small to detect with the available sample sizes.

## Implementation and process findings

The implementation and process evaluation findings across the 6 trials reveal important insights into the feasibility and acceptability of adult numeracy interventions. Learner satisfaction was consistently high across all interventions, with participants valuing both the content and delivery approaches. This positive reception extended to tutors, who generally appreciated the structured resources, lesson plans, and professional development opportunities provided. Many reported planning to continue using elements of the interventions beyond the trial period.

Training requirements significantly affected compliance rates and attrition. Interventions with less intensive tutor training requirements (such as a 2-2.5 hour online sessions) achieved near-universal tutor attendance, whilst those requiring 10+ hours of initial training plus weekly CPD sessions faced greater challenges with tutor engagement and retention. The burden of training requirements frequently led to withdrawals, particularly among part-time staff who struggled to accommodate extensive professional development alongside existing teaching commitments. However, tutors who completed more intensive training programmes often reported valuable professional growth, suggesting the need for balance: training requirements must fit within tutor workloads whilst still providing adequate support during delivery.

Provider attrition varied across the trials, ranging from minimal withdrawal after randomisation to severe drop-out rates. Common withdrawal reasons included insufficient resources, challenges recruiting tutors or learners, late notification of training dates, and lack of staff capacity. Interventions requiring fundamental changes to existing practice, such as asking vocational tutors to deliver mathematics content outside their expertise and additional recruitment of learners outside standard course recruitment faced particular challenges. Conversely, interventions that embedded within existing curriculum hours and required minimal adaptation to current teaching practices experienced low withdrawal rates during delivery.

Fidelity to intervention design was generally high among providers who completed delivery, with sessions typically delivered as intended in volume, timing, and core content. Tutors made appropriate adaptations for learner needs, including adjustments for different skill levels, online delivery requirements, and specific learning difficulties. These minor adaptations maintained the core principles of each intervention whilst ensuring accessibility for different learner populations. Session attendance varied considerably, influenced more by external factors such as work and family commitments than by engagement with the intervention content itself.

The evidence suggests a tension between intervention intensity and implementation feasibility. Interventions requiring minimal adaptation to existing practice achieved high compliance and low attrition but may have limited potential for transformative impact (such as Encouraging Progression and Preparation for Maths GCSE). Conversely, interventions demanding substantial changes to teaching practice and extensive professional development showed promise for innovation but faced significant implementation barriers that led to high levels of attrition (such as Embedded Maths, Contextualised Approach, Adapted Mastery Approach). Interventions that required recruitment of learners onto new courses not previously delivered by providers faced particular recruitment challenges (such as Family Numeracy). Future trials should seek an optimal balance – introducing interventions capable of detecting meaningful impact while remaining feasible for delivery within adult education settings.

# 10. Conclusions and recommendations


## Summary of chapter

This section presents conclusions and recommendations from the programme of Adult Numeracy Trials. It considers lessons learned from programme setup, recruitment, and data collection, providing practical recommendations for future trialling in adult education and beyond. Policy recommendations for commissioners and government are outlined, alongside conclusions about the programme's success in demonstrating trial feasibility and provider willingness to participate in future research and evaluation.


## Conclusions

### Programme set up

#### Partnership collaboration & communication

Strong collaboration between a large number of organisation including DfE, Evaluators, the MSS and product developers was critical to success. Formal partnership agreements established collaboration expectations from the outset, supported by regular bilateral meetings, all-partner workshops, and a culture of open discussion and collaborative problem-solving.

#### Evidence review prevented duplication

A systematic review ensured trials did not repeat existing research within adult education settings. It revealed limited robust evidence on what works in adult numeracy, with existing studies relying on opinions or completion rates rather than impact assessments and Randomised Control Trials (RCTs). This confirmed the evidence gap and provided the rationale for conducting trials in adult education. The systematic review also helped develop an understanding of topics and interventions of interest in the sector.

#### Clear governance structures

The programme governance structure incorporated a Delivery Board, Senior Responsible Officer (SRO) Steering Group, Technical Steering Group plus an independent Technical Advisor and Ethics Panel. This was more extensive than standard DfE research projects given the scale and complexity of the programme of work. This comprehensive structure provided appropriate oversight at each stage.

### Partnership with education developers

DfE partnered with specialist education developers to design interventions, embedding sector expertise within the interventions from the beginning.

### Comprehensive intervention selection process but barriers to quick decision-making

The Trial Feasibility Assessment (TFA) process effectively screened interventions using a consistent framework. The Evaluators developed a TFA template and populated it for each intervention, in collaboration with product developers, to inform a judgement on feasibility. However, the lack of trial-ready interventions meant that there were delays in decision making on which interventions to take forward. Additional work needed to identify trial-ready interventions reduced the recruitment window. Future programmes should set clear deadlines for deciding which trials to take forward.

### Centralised sampling and flexible design approaches enabled multiple simultaneous trials despite limited provider capacity in a small sector and lack of sector benchmarks

To manage multiple trials with limited provider capacity, the programme used a centralised sampling approach with provider-level cluster randomisation. This minimised both contamination between trials and burden on providers. Working with tight timelines and no existing sector benchmarks on recruitment, attrition, compliance or impact of interventions required continuous adaptation – protocols were adjusted and flexible design options maintained throughout. However, this programme has now created those missing benchmarks documented in this report and the individual trial level reports. Future researchers can use this data to design stronger, more effective trials in adult education.

## Recruitment and attrition

### Direct provider contact outperformed top-down recruitment

The initial 'top-down' strategy of engaging Mayoral Strategic Authorities (MSAs) and local authorities to recruit providers was less successful than planned, as the funding model was not attractive enough for them to take on the administrative burden of recruitment. Recruitment pivoted to direct provider engagement, with personalised emails and calls from the MSS proving most effective. While engagement with MSAs and local authorities did not generate formal sign-ups, they played a valuable role in raising awareness, with around a third of recruited providers having heard about the trials through their local authority.

### Lack of research and RCT experience meant some providers required intensive support

The adult education sector had limited previous experience or knowledge of research and more specifically randomised controlled trials, meaning additional time and resources were required to explain evaluation concepts like randomisation and control groups. The lack of experience made direct communication and support essential throughout recruitment.

### Financial incentives were critical for recruitment

The introduction of a £1,000 participation payment for providers improved recruitment rates. The MSS reported a notable increase in expressions of interest and sign-ups after this incentive was introduced. Before this, many providers could see the potential benefits, but participation was not financially viable.

### Limited sector and tutor capacity

The relatively small size of the adult education sector (around 2,000 providers compared to 17,000 primary schools) meant that launching multiple trials simultaneously created significant capacity constraints. Additionally, tutor shortages meant that some providers were unable to take on the extra work involved in participating in a trial, or in some cases were unable to find tutors to deliver the interventions (including to replace those who left). These tutor shortages led to withdrawal from some of the trials.

### Condensed recruitment timeframes

Adult education providers need as much time as possible to plan for the academic year ahead, ideally up to a year in advance. The compressed timeline for the trials meant that products and trial requirements were still being finalised after recruitment began, creating uncertainty that prevented some providers from signing up.

### Intensive training requirements proved unsustainable for some

Training demands proved particularly challenging for the FSQ Level 1 Maths trials. They required treatment tutors to attend an initial 10 hours of training plus weekly one hour lesson study sessions over intervention delivery. This was especially challenging for part-time tutors, which are common in adult education. In contrast, trials with less time-intensive tutor training requirements, such as Preparation for Maths GCSE, experienced lower attrition once delivery began, however this did mean less support readily available for tutors.

### Interest in taking part in future trials and research

Despite recruitment challenges and some withdrawals, providers and tutors expressed interest in taking part in future trials and research, provided these are cost effective and align with existing curricula.

## Data collection and linking

### Ipsos Data Portal with automated checks ensured quality

The online Ipsos Data Portal ensured high-quality data through automated quality assurance checks at the point of submission, allowing providers to correct issues before finalising. The portal also provided a secure data transfer system without requiring individual Data Processing Agreements with each provider. However, compressed timelines prevented proper pilot testing, leading to some errors in early data uploads.

### Use of 2 data portals caused confusion for providers

Providers were required to upload data to 2 different portals, one for the MSS and one for Ipsos. This was due to data sharing agreements and to prevent organisations receiving data they did not require. Future data collection should aim to create one portal designed to create 2 data outputs, so organisations only have access to the data they require.

### High levels of support required to assist providers with uploads

Although the Ipsos Data Portal ensured high quality data, many providers required substantial support to complete uploads due to lack of experience with Excel and formatting data. This included a dedicated inbox and phone support, information webinars demonstrating the upload process, and one-to-one calls.

### Mixed-mode surveys maximised response rates

A flexible approach to surveys for IPE data collection, starting online and then following up by telephone, boosted completion rates. Additional techniques included SMS reminders, language interpretation services and adjusting survey windows for different course schedules. Switching from open links at baseline to unique links at endline improved completion rates, although open links were initially necessary due to delays in data uploads.

### The Individualised Learner Record (ILR) was complex and required significant resource to manage

The ILR is an on-going collection of data about learners from training providers in the Further Education (FE) and Skills sector in England and was the source of primary and secondary outcome measures for the trials. Ipsos received monthly ILR data snapshots, which were valuable in enabling early quality assurance checks, practice-runs of data

linking and the development of code for use in the final impact analyses. DfE provided responsive support and clear documentation to facilitate accurate interpretation of the data. However, inconsistent provider data submissions undermined these benefits, requiring significant resources by DfE, the MSS and Evaluators to prompt course-based providers for complete and correctly code returns.

#### Data linking succeeded despite complexity

A hierarchical "waterfall" matching system[66] successfully linked learner data from multiple sources (management information, surveys, and the ILR), achieving high match rates through both exact and fuzzy matching techniques. This created a central data warehouse consolidating all quantitative learner information across the trials. However, extensive data cleaning was required to address issues including inconsistent column names, conflicting information between sources, and multiple ILR files. Creating clear audit trails and data source priorities proved essential for managing conflicts and maintaining transparency.

## Recommendations to inform future trialling

### Create conditions for collaborative and flexible working

Future programmes of trials should establish formal collaboration expectations from the outset, for example through requiring participating organisations to sign collaboration schedules as part of their contracts. They should create an environment where stakeholders feel comfortable challenging ideas and asking for help, whilst building flexibility into the design to allow adaptations as challenges emerge. These recommendations are particularly important for large scale complex research but could equally apply to smaller research and evaluation projects.

### Improve knowledge of RCTs within the adult education sector

Prior to this programme, providers had limited understanding of randomised controlled trials. Through participation, they gained practical RCT experience. Future programmes should build on this foundation by providing clear information upfront about trial requirements, offering ongoing training opportunities, and maintaining open dialogue with the sector about what enables successful trial participation.

[66] A waterfall system is a hierarchical matching process that attempts to match against criteria in priority order, cascading down to the next available option when higher-priority matches are unavailable or ineligible.

## Allow sufficient lead time for recruitment

Future trials in the adult education sector should ideally begin engagement with providers at least a year in advance of delivery, allowing sufficient time for market warming and curriculum planning and where possible staggering timings if more than one trial is commissioned. Timings should also consider recruitment for courses which may not follow the standard academic year (September to July). All trial materials, including intervention descriptions, data and evaluation requirements, funding models, should be finalised before recruitment begins to improve participation rates and reduce attrition. This includes sufficient lead time for trial design (and power assumptions), informed by recruitment and sample size estimates.

## Using a multi-channel recruitment approach

Future trials should adopt a multi-channel recruitment strategy to maximise reach and sign-ups. Whilst direct contact from recruitment teams proved most effective, this should be supported by awareness-raising through local authorities, sector bodies, and targeted government communications.

## Take measures to ensure data collected is high quality

Quality checks should be implemented wherever feasible to ensure the highest quality data. Ideally, checks should be conducted at the point of data uploads from participating organisations, reducing the need for evaluators to carry out retrospective checks. Trial timelines should include sufficient time to thoroughly test data collection systems before launch. Evaluators should be prepared to invest time supporting providers through the data submission process as, whilst resource-intensive, it helps ensure a complete and robust dataset for analysis.

## Recruit sufficient sample sizes to detect effects

Future trials should prioritise recruiting sufficient number of providers, for example by starting recruitment earlier, to enable fully powered trials. Consideration could also be given to concentrating resources on fewer well-powered trials that can generate definitive evidence, rather than dividing efforts across multiple studies. Where sector capacity presents genuine constraints, explore alternative trial designs that can work with available numbers, such as learner-level randomisation or stepped-wedge designs. As outlined earlier, sufficient sample sizes for fully powered trials require generous recruitment timelines that account for provider engagement, trial explanation, and relationship building.

### Trial interventions requirements should fit within existing structures in adult education and within tutors' workloads

Trials with interventions embedded within existing curriculum hours experienced lower attrition and better compliance, as seen with the Preparation for Maths GCSE intervention. Interventions should be designed to accommodate varied delivery patterns in adult education including evening classes, weekend provision, compressed courses, and rolling starts (not following the standard academic year). Training requirements should also be manageable for tutors within their existing workloads. Future trials should avoid interventions that require extensive additional hours or intensive training requirements that part-time tutors cannot accommodate.

### Balance between intervention intensity and implementation feasibility

Interventions requiring minimal adaptation to existing practice achieved high compliance and low attrition but may have limited potential for meaningful impact. Conversely, interventions demanding substantial changes to teaching practice and extensive professional development showed greater impact potential but faced significant implementation barriers, resulting in recruitment challenges and high attrition. Future trials should seek a middle ground – identifying interventions with the potential to generate measurable change while remaining feasible within adult education structures.

### Provide funding that makes taking part in trials cost effective

Trial designs should ensure funding covers participation costs including staff time, administration, and training requirements. Funding amounts should be confirmed and communicated early – the £1,000 participation payment for providers was critical for recruitment as was the £500 payment for schools. Future trials should also consider differential funding for providers with small class sizes where per learner funding models would not cover delivery costs. This should highlight non-financial incentives like access to Continuing Professional Development (CPD), which the sector particularly values given limited access to free training opportunities.

## Summary

This programme of Adult Numeracy Trials has demonstrated that conducting trials is both possible and welcomed within the adult education sector. Despite the challenges of working with a relatively small sector that had limited prior experience of RCTs. The programme successfully delivered 1 communication-based intervention as a full-scale effectiveness trial (Encouraging Progression) and recruited 96 providers, 72 schools and over 2,700 learners to the 4 smaller pilot RCTs (Family Numeracy, Adapted Mastery Approach, Contextualised Curriculum and Preparation for Maths GCSE) and 1

Implementation and Process Evaluation (Embedded Maths). The pilot RCTs have identified promising interventions and generated valuable methodological learning about conducting robust research in this sector, despite them being underpowered to definitively establish the impact of the interventions that were tested. The collaborative approach between DfE, Evaluators, the MSS and product developers proved essential to navigating the complexities of delivering research at this scale.

This programme has built capacity and awareness of trials within the adult education sector. Providers and tutors now better understand what participating in trials involves and expressed interest in future research opportunities (76% of providers involved in the trials agreed they would consider taking part in an RCT in the future). The methodological frameworks, data collection systems, and practical tools developed benefit DfE and other government departments commissioning large scale programmes. The lessons documented here provide a roadmap for future triallists, not just in adult education but across sectors, offering practical guidance on designing and delivering complex research that balance methodological rigour with delivery realities.

# Annex 1: Glossary of terms

**Adapted Mastery Approach for FSQ Level 1 in Maths:** An intervention designed to support teachers of FSQ Level 1 Maths to adopt a 'Mastery' approach. This involves deepening understanding and spending time building from context to abstract reasoning, instead of just 'covering the curriculum'. It includes 15 taught sessions.

**Adult Education Budget (AEB) – now ASF:** The AEB provides government funding for skills training for adults aged 19 and over in England. It covers basic English, maths and digital skills. On 1 August 2024, the government funded Adult Skills Fund replaced the Adult Education Budget.

**Adult education provider:** An organisation that offers learning, training, and courses to adults aged 19 and over. These can include a wide range of institutions, such as Further Education (FE) colleges, local authority services, independent training providers (ITPs), and voluntary or community organisations.

**Analysis sample:** The final group of participants whose data is used to generate the final statistical results. It is often smaller than the initial sample due to exclusions (e.g., attrition, missing data, loss to follow-up).

**Attrition rate:** The percentage of participants or records lost between two points in a study (e.g., from baseline to endline).

**Attrition:** When participants leave a study before it is finished. This reduces the final sample size for analysis. This can occur when providers or learners withdraw, fail to complete assessments at the end of an intervention or submit required data. Attrition by trial arm looks at the attrition in the treatment and control groups separately. It is used to check for differential attrition (unequal loss between arms) that could bias comparisons.

**Base size (base):** The number of respondents that a particular statistic (e.g., a percentage or mean) is calculated from. In surveys, this is the number of responses

**Baseline survey:** An initial data collection exercise conducted via survey and completed before an intervention begins, to establish a reference point.

**Baseline**: Data collected from participants at the start of a study, prior to any intervention. It serves as a benchmark for measuring change over time, checks that groups are balanced, and improves the precision of the final analysis.

**Bias**: A systematic error that can make study results misleading. Unlike random chance, this error does not disappear by simply increasing the number of participants in the study.

**Binary variable**: A type of variable that can only have two possible values, such as 'pass'/'fail' or 'yes'/'no', often coded as 1 or 0 in statistical analysis.

**Business as Usual (BAU):** The standard set of conditions or practices that participants experience if they are not assigned to receive a new intervention. BAU does not mean doing nothing; it means continuing with the existing approaches (e.g. the standard curriculum) rather than the new approach.

**Cluster Randomised Controlled Trial (RCT):** A research method where groups of people (or “clusters”), rather than individuals, are randomly assigned to different trial arms.

**Clustering**: How individuals are naturally grouped together, such as learners within the same class or provider. Because people within the same cluster often share similar characteristics or experiences, statistical analysis must account for this to avoid drawing incorrect conclusions.

**Coefficient (Standard error):** In a regression analysis, the coefficient is the number that represents the size and direction of the relationship between a predictor variable and the outcome. The standard error is a measure of the statistical accuracy of that coefficient; a smaller standard error means a more precise estimate.

**Cohen’s d:** A widely used statistical measure for calculating effect size, which shows the size of the difference between the averages (means) of two groups. A smaller number (e.g., around 0.2) indicates a small effect, meaning the difference between the two groups is minor. A larger number (e.g., 0.8 or higher) indicates a large effect, signifying a substantial and more meaningful difference between the groups.

**Cohen’s h:** This is a standardised measure of effect size used to compare two proportions (i.e., the difference between two percentages). It is useful for comparing results across different studies that might have different sample sizes or baseline rates. A small Cohen’s h (around 0.2) means a smaller difference between proportions. Even if the difference is statistically significant, it might not be very meaningful in a practical sense. A larger Cohen’s h (around 0.8) means a more substantial difference between proportions.

**Complete separation**: Complete separation is a situation where there are participants in the treatment group who have achieved an outcome but no one in the control group has achieved that same outcome.

**Compliance**: The extent to which providers adhered to the trial requirements outlined in the Trial Readiness Packs, including data submission requirements and adherence to eligibility criteria.

**Computer Assisted Telephone Interviewing (CATI):** A data collection method in which a researcher conducts a survey over the telephone while using a computer-based system to guide the interview and record responses.

**Confidence intervals:** Confidence intervals are used to express the certainty of an estimate. The interval is the range of values within which the 'true' value for the whole group is highly likely to lie. The smaller the range, the more certain the estimate. A 95% confidence level (the most common) means that if we repeated the study and analysis 100 times, 95 of the ranges calculated would include the true value of the population.

**Contamination:** A potential issue in a trial where the control group is unintentionally exposed to the intervention.

**Contextualised Curriculum for FSQ Level 1 in Maths:** An intervention designed to improve adults' attitudes and beliefs about the value of maths and attainment in maths by using a Realistic Mathematics Education (RME rme.org.uk) approach in classroom delivery. This means learners are engaged with context prior to learning a technique. It includes 12 taught sessions.

**Continuing Professional Development (CPD):** Ongoing learning activities for tutors/teachers to maintain and improve their professional knowledge and skills (for example, training, workshops, mentoring). Many adult education providers (e.g. FE colleges, training centres) require documented CPD for quality assurance and accreditation purposes.

**Control group:** A group that does not receive the intervention(s) being tested within an RCT design following randomisation. They are monitored alongside the group(s) receiving the intervention(s), and their results are compared to their treatment counterparts to understand what impact the intervention has had, compared to receiving no intervention. Any changes or effects detected within the control group over the course of the RCT can be interpreted as what would have happened normally.

**Cooperation rate:** A survey metric showing the proportion of contacted, eligible people who completed the survey. Typically calculated as completes divided by (completes + refusals + break-offs) among those reached and eligible.

**Correlation**: A correlation is a statistical measure that describes the strength and direction of an association between two or more variables. However, it is important to note that correlation does not imply causation, or how much such variables will change when a change is observed in the independent variable.

**Covariates:** Characteristics of participants (e.g., age, prior qualifications) that are measured at the beginning of a study and can be used in the statistical analysis to account for pre-existing differences between groups.

**Descriptive statistics:** Statistics used to summarise and describe the main features of a dataset, such as the mean, median, and standard deviation.

**Effect sizes:** A standard metric that quantifies the strength of a result. An effect size tells you not just if an intervention worked, but how much it worked, allowing for comparisons between different studies and contexts. Larger effect sizes indicate a stronger effect.

**Embedded Maths:** An intervention to improve numeracy skills in adults by embedding mathematical concepts into the curriculum of Health and Social Care Level 2 qualifications.

**Endline survey:** A survey completed at the end of an intervention period.

**Entry level:** Qualifications at entry level provide an introduction to education and can lead to certification of essential skills and knowledge for beginners.

**ESOL learners**: 'English for Speakers of Other Languages' learners; those taking part in a course for whom English is not their first language.

**Experimental**: An evaluation design where participants are deliberately assigned to groups, ideally through randomisation. By ensuring that there are no systematic differences between the treatment and control groups at the start (such as control groups and treatment groups formed by individuals with very different ages), this design provides the strongest possible evidence that the intervention caused any observed results.

**Exploratory analysis**: An early analysis of data to find initial patterns or interesting results. Findings from this stage are treated as suggestions that need to be tested properly in a future study, as they have a higher risk of being due to chance.

**Family Numeracy:** A programme hosted by primary schools for parents and their children to improve adult numeracy skills through family learning. Includes 12 hours of learning: 6 in-person group sessions and home learning activities.

**Feasibility:** An assessment of whether a proposed study or intervention can be practically implemented given the available resources, time and logistical constraints. It focuses on practical considerations such as recruitment volumes and data collection processes.

**Fidelity**: The extent to which an intervention is delivered as intended. It assesses whether what was implemented aligns with the original design, ensuring that the results reflect the true intervention rather than a diluted or altered version.

**Foundation Tier Maths GCSE**: Maths GCSE entry for grades 1-5.

**Functional Skills Qualification (FSQ) Level 1 in Maths:** A qualification that focuses on practical mathematical skills needed for everyday life and work, equivalent to GCSE grade 1-3. FSQ Level 1 Maths is usually targeted at individuals who require a more applied or vocational approach to learning maths and is suitable for learners aged 16+ years in further education, apprenticeships, or adult learning.

**Guided learning hours:** The amount of time a tutor is scheduled to be present to provide specific guidance to learners as part of a course of study.

**Health and Social Care L2**: Health and Social Care (H&SC) refers to a broad sector that includes both healthcare services (such as hospitals, clinics, and GP surgeries) and social care services (such as residential care, home support, and community-based assistance). The H&SC qualification at Level 2 (equivalent to GCSE grades 4–9) is a foundational qualification needed to access a variety of entry-level roles in both health and social care settings. It covers topics such as communication, safeguarding, personal care, and health and safety, and is suitable for those starting out in the sector or looking to formalise existing experience.

**Hedges' g**: Hedges' g is a standardised measure of effect size that expresses the difference between two means in terms of standard deviations. It includes a small correction for bias in small samples, making it slightly more accurate when the sample size is small. Around 0.2 is considered a smaller Hedge's g (a smaller difference between means), and 0.8 is considered a larger Hedge's g (a bigger difference between means).

**Impact evaluation:** A study designed to determine whether an intervention is the cause of an observed change. It works by comparing the outcomes of participants against a counterfactual – that is an estimate of what would have happened to those same participants had the intervention not taken place.

**Impacts**: The broader, long-term effects of an intervention on participants and their environment, such as improved employment prospects or sustained changes in teaching practices.

**Implementation and Process Evaluation (IPE):** A study designed to complement an impact evaluation by examining how the intervention was put into practice. While the impact evaluation measures outcomes, the IPE assesses factors like fidelity and participant engagement to explain why those results occurred.

**Imputation:** This is the statistical process of replacing missing data with substituted, plausible values.

**Individualised Learner Record (ILR):** The ILR is an on-going collection of data about learners from training providers in the Further Education (FE) and Skills sector in

England. It was used in the adult numeracy trials as a source of administrative data for outcomes such as grades, course completion and progression to further learning.

**Intention to treat (ITT):** This is a fundamental principle for analysing the results of a randomised controlled trial. It means that all participants are analysed in the group to which they were originally randomised, regardless of whether they actually received the intervention, completed it, or complied with it. This method preserves the benefits of randomisation and avoids bias.

**Interaction effects:** When the effect of an intervention differs depending on another factor rather than the intervention itself (for example, the impact varies by delivery mode, provider, or learner characteristics). This is also referred to as moderation.

**Intervention/ treatment arms**: In a trial, the "arms" are the different groups to which participants are assigned. An "intervention arm" or “treatment arm” is any group that receives a specific treatment or intervention being tested.

**Intervention:** In the context of a trial, an intervention refers to a specific programme, approach, or set of activities being tested for its effectiveness. It represents the treatment or change being implemented with participants, which is then compared against a control group or alternative approach to measure its impact on specified outcomes.

**Intraclass correlation coefficients (ICC):** A 0–1 measure of how similar outcomes are within the same cluster (e.g., provider or class) compared with across clusters. Also known as intra cluster correlation coefficients.

**Key stage:** A term used in the education system in England, Wales, and Northern Ireland to describe a specific stage of compulsory schooling, defined by a child's age. Key Stage 1 covers ages 5-7. Key Stage 2 covers ages 7-11. Key Stage 3 covers ages 11-14. Key Stage 4 covers ages 14-16 and concludes with national exams (GCSEs). Key Stage 5 covers ages 16-18.

**Kicktag**: Kicktag is the brand name of the data collection system used for these trials, also known as the Ipsos data collection portal.

**Learner progress**: A measure of the change in learner outcomes over the duration of a trial. It can involve tracking of development in areas such as academic performance, confidence, and study skills amongst both the treatment and control groups, enabling comparison of the relative progress of each.

**Level 1:** Qualifications at level 1 include or are equivalent to GCSE grades 3, 2, or 1 (previously D, E, F or G).

**Level 2:** Qualifications at level 2 include or are equivalent to GCSE grades 9, 8, 7, 6, 5, or 4 (previously grade A*, A, B, or C).

**Linear regression:** A statistical method used to estimate the relationship between an outcome variable (such as GCSE grade) and one or more predictor variables.

**Local area:** The collective term ‘Local areas’ covers the authorities that commission and coordinate Multiply programme delivery in their area. These are the Greater London Authority (GLA), Mayoral Strategic Authorities (MSAs) and upper tier and unitary local authorities outside of these areas.

**Longitudinal:** A type of research design that involves collecting data from the same subjects repeatedly over a period of time. This allows researchers to track changes, development, and long-term effects.

**Managed Service Supplier (MSS):** The MSS led on the recruitment and management of providers and schools who took part in the course-based adult maths trials. They were responsible for contract management, monitoring and reporting of delivery and issuing payments. The MSS also facilitated tutor training for providers assigned to treatment groups.

**Maths GCSE:** An academic qualification typically taken by students in the UK around the age of 15-16, but which can also be taken by people of all ages. It is graded 1-9 with anything above grade 4 considered a pass. The GCSE serves as an important benchmark for further academic and professional pursuits, often required for entry into further education, vocational qualifications, or employment.

**Mayoral Strategic Authorities (MSAs):** Regional governance bodies in England that replaces the previous Mayoral Combined Authority model. Led by a directly elected mayor, these authorities coordinate wide-scale economic growth, infrastructure, and public services across a defined geographical area without replacing local councils. There was a transition from Mayoral Combined Authorities to Mayoral Strategic Authorities (MSA) as part of the 2025’s Devolution Bill. This reflects a shift in UK devolution, moving from bespoke arrangements to a uniformed statutory framework.

**Meta-analyses**: Studies that systematically combine and statistically pool results from multiple studies on the same question to produce an overall estimate of impact.

**Minimal detectable effect sizes (MDES):** This is the smallest true effect (or impact) of an intervention that a study has a good chance (usually 80% probability) of detecting as statistically significant. It is calculated before a study is carried out (at the design stage) to determine whether the sample size is sufficient to find a meaningful result.

**Multimode:** This refers to surveys that can be completed through different channels or modes, e.g. online or by telephone.

**Multiple Imputation by Chained Equations (MICE):** A statistical method that handles missing data by creating several different plausible datasets. By analysing these combined datasets, researchers can account for the uncertainty of the missing values, leading to more accurate standard errors and conclusions.

**Observational studies:** Studies that observe what happens without assigning people to conditions (for example, tracking outcomes in naturally occurring groups). They are useful for describing patterns and associations but are weaker for evidencing causal claims.

**Outcomes:** The specific, measurable results of an intervention that are tracked to evaluate its effectiveness, such as learner pass rates, attendance, and changes in confidence.

**Outputs:** The immediate, tangible products or services delivered by an intervention as a direct result of its activities. They describe 'what was done' or 'what was produced' rather than the changes that resulted from it. For example, the number of learners who participated in a maths lesson.

**Pedagogy:** The theory and practice of teaching and learning (for example, the methods, principles, and strategies used by tutors).

**Percentage points (ppt):** A percentage point is the unit for the absolute arithmetic difference of two percentages. For example, moving from 10% to 12% is an increase of two percentage points.

**Pilot Randomised Controlled Trial (Pilot RCT):** A small-scale randomised study to assess the practical application of an intervention and the validity of the research methods. It focuses on gathering evidence regarding feasibility, implementation, and acceptability, allowing researchers to refine the programme design and data collection tools based on real-world feedback.

**Purposive sampling:** A sampling technique where researchers deliberately select participants based on specific characteristics relevant to the study's goals, rather than at random.

**p-value:** The p-value, or probability value, is the probability that a result occurred by chance. A small p-value (usually 0.05 or less) suggests the result is 'statistically significant', meaning it is unlikely to be a coincidence.

**Qualification-bearing Maths course**: This is a Maths course that, upon successful completion, results in a formal qualification.

**Qualitative data:** Refers to non-numerical data that is descriptive in nature, such as interview transcripts, observations, and case studies. It focuses on understanding experiences and perspectives.

**Quantitative data:** Refers to numerical data that can be measured and statistically analysed, such as test scores, pass rates, and survey ratings.

**Quasi-experimental:** An evaluation design that attempts to estimate impact without using random assignment. Instead, it relies on statistical techniques (such as matching or difference-in-differences) to construct a comparison group that resembles the treatment group as closely as possible.

**Randomisation:** The process of assigning participants to treatment or control groups using a random mechanism (such as a computer algorithm). This ensures that every participant has an equal probability of assignment, creating groups that are statistically equivalent at the start of the study.

**Randomised Controlled Trial (RCT):** An evaluation design where participants are randomly assigned to either a treatment group or a control group. This process ensures the groups are statistically equivalent at the start, meaning that any difference in final outcomes can be confidently attributed to the intervention rather than external factors.

**Readiness for trial:** An assessment of whether an intervention is sufficiently well-developed and stable enough to meet the requirements of a randomised controlled trial, and whether the proposed research methods are feasible.

**Recall bias:** A systematic error that occurs when people are asked to remember past events or experiences and their memories are incomplete or inaccurate.

**Regression model:** A statistical tool used to model and analyse the relationship between a dependent variable (the outcome) and one or more independent variables (the predictors).

**Robustness check**: An analysis to test whether the main results of a study hold up when the methods or assumptions are slightly changed.

**Robustness**: Whether or not the main results of a study hold up when the methods or assumptions are slightly changed.

**SEND learners:** Those taking part in a course who have Special Educational Needs and/or Disabilities.

**Sensitivity analysis**: A statistical method used to assess how the results of a study might change if key assumptions or population data were different.

**Skewed**: This describes a distribution of data that is not symmetrical. A distribution is "skewed" if the data points are not evenly distributed around the average (mean). For example, a sample may be skewed towards having lower levels of qualifications if the sample contains more data entries with lower levels than higher ones.

**Standard deviation:** This is a measure of the amount of variation or dispersion in a set of values. A low standard deviation means that all values tend to be close to the average (mean), while a high standard deviation means that the values are spread out over a wider range.

**Statistically powered:** A term describing a study that has a large enough sample size to have a high probability (typically 80% or more) of detecting a real effect if one exists.

**Statistically significant:** A result is "statistically significant" if it is unlikely to have occurred by random chance alone. Researchers typically set a threshold to make this determination (p-value). It suggests there is a real effect or relationship in the data.

**Technical Steering Group (TSG):** A Technical Steering Group was established by DfE to provide advice, guidance and oversight of key design elements of the trials. They also had ownership of technical risks and oversight of final outputs and ethical considerations.

**Theory of Change:** A model that explains how an intervention is expected to work. It maps the logical pathway from the inputs and activities to the intended short-term outcomes and long-term impacts.

**TIDieR framework:** TIDieR stands for the Template for Intervention Description and Replication. It is a 12-item checklist and guide designed to improve the completeness and quality of how research interventions are described in publications.

**Treatment group:** The group of participants randomly assigned to receive the specific programme or policy being tested. Their outcomes are compared against those of the control group to determine if the intervention caused a significant change.

**Trial participant:** An individual who meets the eligibility criteria, has provided informed consent, and has been formally enrolled or randomised into the study.

**ULN**: Unique Learner Number. A unique 10-digit number assigned to individuals over the age of 14 involved in education or training in the UK.

**Variance:** A specific statistical measurement that measures the spread of data points around their average value (the mean). A small variance means the data points are

clustered tightly around the average, while a large variance indicates they are more widely scattered.

# Annex 2: Overview of the research conducted by the consortium from May 2023 to March 2026

**Table A1: Overview of interventions**

| Name of intervention | Type of research/RCT | Organisation lead |
|---|---|---|
| Encouraging Progression | RCT, IPE | Learning and Work Institute |
| Preparation for Maths GCSE | Pilot RCT, IPE | Ipsos |
| Contextualised Approach in Functional Skills Level 1 | Pilot RCT, IPE | Institute for Employment Studies |
| Family Numeracy | Pilot RCT, IPE | Institute for Employment Studies |
| Embedded Maths | IPE | Institute for Employment Studies |
| Adapted Mastery Approach in Functional Skills Level 1 | Pilot RCT, IPE | The Policy Institute at King's College London |

Source: Ipsos

# Annex 3: Ethics panel terms of reference

The Terms of Reference outlines the ethical approach developed for the Multiply Education Research trials. This document was sent to the Ethics Panel convened for the programme and established the common ethical framework the Panel would apply when reviewing trial proposals.

## Introduction

Ipsos has been appointed by the Department for Education (DfE) to deliver Randomised Control Trials (RCTs) of adult numeracy interventions delivered through the Multiply programme. Ipsos is leading a consortium to deliver this work, which includes the Institute for Employment Studies, Learning and Work Institute, King's College London and RAND Europe. A bespoke Ethics Panel is being established to review and advise consortium members on ethical considerations associated with each trial. This paper sets out a Terms of Reference for the Multiply RCT Ethics Panel.

## Background to the Multiply RCT programme

The Multiply RCT programme aims to address a gap in the available evidence on what works in adult skills. There is an estimated £120m a year spent on adult numeracy in the UK, with limited robust evidence of effectiveness or impact. The aim of the RCTs is to add to the evidence of what works to inform future investment in the sector, ensuring it is being used in the most efficient and effective way.

The list of potential trial ideas has come from frontline practitioners, commissioners and sector experts. They cover the whole learner journey from initial engagement, the types of support that can help maintain motivation to stay on / complete courses, the types of teaching / peer support that works best and types of interventions that support progression to higher level courses. The trials are therefore varied in their design and will cover a range of different delivery methods, pedagogies, target groups, delivery settings, geographies, levels of ability and learning outcomes.

## Approach and rationale

As the Multiply RCTs will involve withholding or delaying access to numeracy interventions, careful consideration needs to be given to ensuring that they are conducted in an ethical manner to ensure the protection of participants and the validity of the findings. At the same time, it is important that the ethics review and approval process for the trials is proportionate to the risks involved for those participating in trials, considering the severity of personal or social harm that could transpire as a result.

The Multiply RCT Ethics Panel will be responsible for reviewing and advising on ethical issues associated with each of the Multiply RCT trials. The Ethics Panel will be made up of:

- Senior members of the Ipsos Ethics Group – who have expertise in the specific ethical issues associated with evaluations and RCTs, as well as quantitative and qualitative research with adults.
- A representative from each of the other consortium partners – each of the four partners will nominate one person to join the Ethics Panel who has expertise in the ethical issues associated with trials.
- A condition of membership to the Ethics Panel will be that members are not directly involved in any aspect of Multiply RCT delivery.

The rationale for establishing a bespoke Ethics Panel is to ensure:

- Consistency – having one dedicated group to review and advise on ethics across all trials will ensure that similar issues are addressed in a consistent manner.
- Efficiency – one induction to the panel will ensure that all members are fully informed of the context, background, aims and objectives of the RCT programme. As panel members build their knowledge and understanding of the trials, this should result in efficiencies in assessing the ethical considerations associated with each.
- Rigour – the specific ethical considerations associated with RCTs require a higher level of scrutiny than other forms of research and evaluation. Having a panel made up of representatives from five separate organisations with different areas of expertise will ensure a thorough and rigorous review of the ethics associated with all aspects of trial design and delivery.
- Independence – as noted, ethics panel members will not be directly involved in Multiply RCT design or delivery. Their role will be to advise on ethical issues associated with the Multiply RCTs, with the Technical Steering Group (TSG) providing final sign-off on all Trial Protocols.

## Process for ethics review and approval

The role of the Ethics Panel will be to review the completed ethics form and to provide comments, advice and guidance to consortium partners on any ethical issues or concerns associated with each. They will be asked to RAG rate each trial based on whether they consider it to be high, medium or low risk in terms of ethical considerations. Final sign-off of the Trial Protocol will be the responsibility of the Multiply RCT Technical

Steering Group (TSG), who will receive the report of the Ethics Panel and details of how consortium members have responded to this.

The section below detail the processes through which each trial would undergo ethics review and approval.

## Ethics review and approval process for Multiply RCTs

1. Completed Ethics Form is shared with Ethics Panel members for review
2. Panel meet to discuss and agree ethical issues associated with the trial, with the lead triallist attending part of the meeting to answer questions
3. Panel provide written advisory feedback to consortium partners and recommended RAG rating for the trial
4. Consortium partners address advisory feedback from Ethics Panel and update Trial Protocol accordingly
5. Report of the Ethics Panel and updated Trial Protocol shared with DfE for review
6. Report of Ethics Panel and updated Trial Protocol shared with TSG for review

Where trials are rated as medium or high risk by the Ethics Panel and the TSG agrees, additional ethical review will be sought from the DfE Director of Analysis, who will have the final say on whether the trial should proceed.

## Additional ethics process for trials rated as medium or high risk

7. Report of Ethics Panel and updated Trial Protocol shared with TSG for review
8. TSG agree with rating provided by Ethics Panel, or return with comments/questions if they disagree. Final decision on ethics made by TSG
9. For trials rated medium or high risk, additional ethical review and final decision made by senior member of DfE

Ethics Panel members will be expected to base their review and feedback on each trial on the six ethical principles set out in the Government Social Research (GSR) Ethics Guidance, which cover the following areas:

1. Clear and defined public benefit
2. Sound application, conduct and interpretation
3. Data protection regulations

4. Specific and informed consent
5. Enabling participation
6. Minimising personal and social harm.

In reviewing ethical considerations associated with individual trials, Ethics Panel members should also consider:

- Level of risk – feedback on ethical considerations associated with each trial should be proportionate to the potential risk associated with each, particularly the level of risk to participants.
- Mitigation measures – where the Ethics Panel identify potential ethical issues associated with individual trials, they should provide advice and guidance to consortium partners on potential mitigation measures.
- Good practice – in addition to the GSR principles set out above, Ethics Panel members should take account of wider good practice in informing their ethical review (such as the Economic and Social Research Council Framework for Research Ethics).
- Ongoing monitoring – consortium members will be responsible for feeding back to the Ethics Panel on any ethical issues arising during delivery of the RCTs and (where appropriate) seeking follow up advice on how to address these.

Breaches of good ethics practice will be treated as a serious matter. Where these occur, they will be subject to review by DfE.

## King's College London ethics clearance

King's College London are one of the RCT consortium partners. As a university, they have their own mandatory ethical clearance process that will need to be adhered to for those Multiply RCTs that they are leading. The College Research Ethics Committee (CREC) is responsible for providing ethical clearance for all KCL research projects that collect primary data, or for those that collect secondary data if data is identifiable. Only data that has been collected after ethical clearance has been granted can be used in KCL publications. Clearance must be granted before any data collection begins and retrospective clearance cannot be provided under any circumstances.

For those trials that KCL are leading, ethical issues and considerations will be reviewed by the Multiply RCT Ethics Panel first (which will include a representative from KCL who will be alert to the types of issues that are likely to be picked up by the CREC). Once feedback from the Panel has been addressed, it will then be submitted to the CREC for clearance.

## Expectations of Ethics panel members

Members of the Multiply RCT Ethics Panel will be expected to contribute half a day per trial to review ethical considerations and attend a meeting with to discuss these with other panel members. Most of the ethics reviews will take place during 2024. Panel members will be given an outline schedule of when ethics reviews are expected to take place for each trial to help with their planning. The aim will be to schedule Ethics Panel meetings at least one month in advance, with paperwork provided two weeks before this.

Panel members could be perceived as having a vested interest in trials proceeding given that their organisations stand to gain financially from this. This will be partly mitigated by DfE having the final say on which trials proceed. We will further mitigate this by ensuring that the lead organisation for each trial is not revealed to the Ethics Panel in the ethics form at the independent review stage. The trial lead will attend the meeting after the panel have had initial discussions – at this stage it may be clear which organisation is leading.

Following feedback from the first Ethics Panel, it was also agreed that Panel members will avoid discussing the specifics of the trials, forms and what was discussed in the panel within their own organisations. The Chair of the Panel will not be subject to the same restrictions given the need to communicate about the trial process to Ipsos project team members.

The Chair of the Panel will have an important role in reaching consensus across members. This role will be taken by an Ipsos Research Director (senior member of the Ipsos Ethics Committee).

There is a risk that feedback from the Ethics Panel conflicts with the view of the TSG, including on whether trials should proceed. The TSG will have full sight of all feedback from the Panel. In cases where there is a conflict, the trial will be escalated to DfE for ethics review who will have the final say on whether trials go ahead.

## Expectations of lead trialists

The lead trialist will be responsible for drafting the Ethics Form based on the available information about the trial. The form will be provided approximately 2 weeks in advance of the Panel meeting.

The lead trialist will be invited to attend part of the Ethics Panel meeting to answer questions or clarify the approach proposed. This should not compromise the anonymity of the review since it takes place following the independent review of the form and once views have been gathered from the panellists.

## Sample agenda for Multiply RCTs ethics panel meeting

**Table A2: Sample agenda for Multiply RCTs ethics panel meeting**

| Time | Item | Actions |
|---|---|---|
| **30m** | Panellists discuss the Ethics form | Chair to invite all members of Panel to share views and key questions in case not answered in the completed forms shared prior to the meeting. |
| **20m** | Questions posed to lead triallist | Lead trialist to be invited to attend the meeting to respond to questions. Triallist to leave meeting once questions are answered. |
| **10m** | Panel agreement on proposed RAG rating for trial.<br>Next steps | Update Ethics Form with RAG rating<br>Confirm availability of panellists for review of the report / completed form and any other feedback. |

## Indicative task plan and timeframe for Multiply Ethics Panel review per trial

**Table A3: Indicative task plan and timeframe per trial**

| Task | Timeframe | Responsible |
|---|---|---|
| **Schedule Ethics Panel with members and Chair** | Approx. 1 month before Panel due to meet | Ipsos |
| **Send calendar hold to lead trialist (NB not the same invitation that will be sent to Panel to avoid risk of exposing lead trialist to Panel members in advance)** | Approx. 1 month before Panel due to meet | Ipsos |
| **Send ethics form to evaluator/lead trialist to complete** | Approx. 3-4 weeks before Ethics Panel | Ipsos |

| Task | Timeframe | Responsible |
|---|---|---|
| **Complete and return Ethics form to Ipsos** | 2 weeks before Ethics Panel | Lead trialist / evaluator |
| **Send ethics form to Panel for independent review** | 2 weeks before the Ethics Panel | Ipsos |
| **Collect written forms/feedback if necessary** | Up until date of meeting | Ipsos/Multiply Ethics Panel Chair |
| **Ethics Panel takes place.**<br><br>**Send contact details of lead trialist to Chair on the day to invite into the call at the appropriate moment** | Date as scheduled – approx. 1 hour per trial (to be determined by Chair) | Ipsos/Multiply Ethics Panel Chair |
| **Complete Ethics form including any amendments from the panelists** | Up to 2 weeks after the meeting | Multiply Ethics Panel Chair |
| **Send completed Ethics form to Ipsos delivery team for review/actions** | 2 weeks after the meeting | Multiply Ethics Panel Chair |
| **Send completed Ethics form to lead trialist for review / feedback** | 2 weeks after the meeting | Ipsos |
| **[If LOW RISK trial] Any actions to be updated in Trial Protocol and sent back to Ipsos and DfE along with completed ethics form.** | 3 weeks after the meeting | Lead trialist / Ipsos |
| **[If MEDIUM OR HIGH RISK trial] Follow additional approval steps as outlined in Process above** | 3 weeks after the meeting | DfE |

# Annex 4: Ethics review form

## Instructions

### Introduction

This ethics form should be completed by the lead triallist for each Multiply RCT. It will then be submitted to the Multiply RCT Ethics Panel for review and feedback.

Process for completing and submitting the form

- Please complete each of the sections of the form drawing on information from the Trial Protocol and other relevant documentation (e.g. product description, Theory of Change, etc).
- For any questions where you answer yes, please provide details of what you plan to do and how you aim to mitigate any ethical risks.
- Please avoid acronyms when completing the form as the Ethics Panel members may not be familiar with them.
- Once completed, please email the form to the Multiply RCT Programme Manager and Programme Director.
- Be prepared to attend the Ethics Panel meeting for 20-30 minutes to answer ethics questions about the trial design and approach.
- Please do not include the name of your organisation or any personally identifiable information within this form to ensure independent review from the Ethics Panel.

**To be completed by the trial lead**

**RCT name:**

**Date of form submission:**

**Brief description of RCT:**

**Please reference:**

**The overall design of the RCT and rationale for decision-making on, for example, if a waitlist approach was considered:**

**Evidence base (strong/weak):**

**'Business as usual' for the intervention being trialled:**

**Research methods and approach:**

**Sampling:**

**To be completed by the Ethics Panel following review of the form**

**Recommended ethics RAG rating for this trial (delete as appropriate):**

HIGH RISK

MEDIUM RISK

LOW RISK

**If identified as Medium or High risk, brief summary of ethical concerns**:

HIGH RISK: Do not proceed until appropriate mitigation strategies have been put in place to address ethical concerns.

MEDIUM RISK: Proceed with caution: ethical risks have been identified and while mitigation strategies are in place there are (limited) concerns raised with the approach.

LOW RISK: Proceed with caution: no ethical risks are identified, or risks have been identified and a convincing mitigation strategy is in place.

## 1. Does the trial involve vulnerable people or those who may be unable to give fully informed consent?

Please outline any considerations of vulnerabilities participants may be experiencing and any implications for the trial. This could include groups that may experience vulnerabilities (in particular those outlined below), vulnerabilities linked to power or discrimination, as well as other complex and potentially fluid vulnerabilities, such as financial vulnerability, health outcomes, life events etc.? Examples could include:

- Children
- People with learning disabilities or cognitive impairment
- People with mental or physical health issues
- People with a substance addiction including drugs or alcohol
- Homeless people

- Old / very elderly or frail elderly people
- Prisoners
- People who speak limited English
- People with accessibility needs e.g. Deaf people who communicate with British Sign Language

**Please detail:**

Steps taken by project team

Ethics review feedback

## 2. Will the trial involve delaying or withholding access to support for people who could potentially benefit?

Please describe the process for randomising participants to treatment and control arms of the trial, including how many treatment arms there are, and likelihood of participants being assigned to control. If access to treatment will be provided later (e.g. through a waitlist design), please provide details. If those in the control group will not receive access to treatment either now or in the future, please detail any potential harms to them from this being withheld.

Yes/No

If yes, please detail

Steps taken by project team

Ethics review feedback

## 3. Is there a clear and defined public benefit?

**Is the trial designed in the public interest either directly or indirectly?**

**Have any risks to the public been determined and possible mitigations been established?**

**Does the potential benefit of the trial outweigh any potential risks?**

Yes/No

If no, please detail

Steps taken by project team

Ethics review feedback

## 4. Have participants been given the opportunity to opt-out of the trial?

In most cases, participants should be given the opportunity to opt-out of the trial if they do not wish to take part. If you think opt-out is inappropriate for your trial (for example, if participants are being identified and recruited via administrative datasets), please provide details.

Yes/No

If no, please detail

Steps taken by project team

Ethics review feedback

## 5. Will the research involve discussion of sensitive issues?

Could the questions asked of RCT participants cause anxiety/stress beyond the risks of everyday life? Is there the potential for participants to become distressed? Examples could include:

- Topics that are highly personal in nature (e.g. sexual orientation or behaviour)
- Experience of traumatic events (e.g. domestic violence, human trafficking)
- Mental or physical health difficulties
- Criminal behaviour e.g. drink driving
- Addictions (e.g. gambling, drug or alcohol)
- Debts, personal finances, issues with benefits

Yes/No

If yes, please detail

Steps taken by project team

Ethics review feedback

## 6. Is there a heightened risk of disclosure of harm?

Is there a heightened risk of disclosure of harm that could require you to break confidentiality? This could happen on projects that ask vulnerable audiences about sensitive issues (e.g. children's experience of online harm), and/or where clients require potential safeguarding issues to be identified and escalated to a third party.

Yes/No

If yes, please detail

Steps taken by project team

Ethics review feedback

## 7. Is there a raised risk for researchers / interviewers?

**This may be due to:**

- the topics or nature of the questions (e.g. asking people about traumatic experiences)
- information you have about participants (e.g. ex-offenders)
- the areas where research is taking place (e.g. high crime areas, homes of individuals known to have a chaotic lifestyle)

Yes/No

If yes, please detail

Steps taken by project team

Ethics review feedback

## 8. Will the trial require the co-operation of gatekeeper(s)?

For example, to provide access to the target population, introduce the trial to potential participants, recruit participants to the trial, deal with consent and / or collect data. Have the ethical implications of the requirements on gatekeepers, such as providers, been given due consideration e.g. the administrative burden on providers and teaching staff to collect data and who may need to explain why learners cannot opt-out of the intervention even if they can opt-out of the evaluation; or maths assessment being added to a non-maths course?

**Examples of this include:**

- Schools
- Teachers
- Adult Education Providers
- Parents of children
- Employers
- Healthcare workers (e.g. a doctor or a nurse)
- Social workers
- Prisons

Yes/No

If yes, please detail

Steps taken by project team

Ethics review feedback

## 9. Sharing data, data linkage or identifying participants

- Does the trial involve sharing identifiable or potentially identifiable data outside of the project team (intention at this point, not future possibilities)?
- Does the study involve data linkage with administrative sources?
- Is the information difficult to anonymise, so that there is the risk of identification at the reporting stage?

Yes/No

If yes, please detail

Steps taken by project team

Ethics review feedback

## 10. Covert observations

**Will it be necessary for participants to take part in the study without their knowledge and consent at the time?**

Yes/No

If yes, please detail

Steps taken by project team

Ethics review feedback

## 11. Does the trial involve the use of social media or internet forums (e.g. social media monitoring, listening etc.)?

- Does your methodology collect identifiable information about users without their consent? (i.e. through a third party tool where data is collected on the basis of the terms and conditions)
- Is the social media content covered by the research of a particularly personal or sensitive nature? (e.g. sexual behaviours or research with an interest in young people)
- Do you intend to re-publish individual social media posts to either the client or publicly?

Yes/No

If yes, please detail

Steps taken by project team

Ethics review feedback

## 12. Diversity and inclusion

- Are there any particular sensitivities or issues regarding diversity and inclusion that the trial needs to take account of? If yes, please detail. You may want to consider the following points in your response:
- The representation of a diverse range of characteristics within your sample
- How you will ensure appropriate wording for any questions relating to characteristics such as gender, ethnicity, sexual orientation and disability
- Steps to maximise accessibility for people with different characteristics or needs.

Yes/No

If yes, please detail

Steps taken by project team

Ethics review feedback

## 13. Are there any other issues that may raise ethical concerns with this trial?

**Other issues that may raise ethical concerns include:**

- Particularly burdensome interviews or prolonged involvement in the trial
- Levels or types of incentive payments (e.g. large incentives that could impact benefit payments; payments to particular groups e.g. offenders, drug users)
- Use of online communities where there is potential for participants to interact without being moderated by a researcher
- Collection of very intrusive data
- Any other elements that you are concerned about from an ethical perspective.

Yes/No

If yes, please detail

Steps taken by project team

Ethics review feedback

## Actions taken (box to be completed by the project team AFTER the ethics review process)

**Please use this box to record the action(s) you have taken or the changes you have made as a result of the ethics review process, if any**:

# Annex 5: RACI framework

## RACI framework

Table A4 below sets out the key responsibilities by organisation for each stage through the “RACI” Framework.

Key: RACI stands for:

- **(R) R**esponsible: The organisation which is going to implement the task. The “owner” of the task at that particular time.
- **(A) A**ccountable: The organisation which will sign off on the work and judge its completion and how it meets quality standards.
- **(C) C**onsulted: These are the knowledge sources whose collaboration is necessary to reach the task and project goals. Typically, these are Subject Matter Experts. Their input is needed before actions are taken or decisions are made.
- **(I) I**nformed: Those who must be kept informed of the work but not necessarily consulted. They are given all the latest task updates on a regular basis. They are informed after actions are taken or decisions are made. Typically, a one-way communication.
- N/A = not applicable.

**Table A4: Key responsibilities by organisation**

| Stage number | Description of stage | MCA/LA | Local Providers | MSS (Etio) | Evaluator (Ipsos and partners) | Product developer | DfE (as client) |
|---|---|---|---|---|---|---|---|
| 1 | Develop intervention | C | C | I | C | R | C / A |
| 2 | Develop Protocol | N/A | N/A | C | R | C | A |
| 3 | Ethics approval | N/A | N/A | N/A | R / A | N/A | I / A |
| 4 | IDEA workshop(s) | N/A | N/A | R | C / A | C | I |

| Stage number | Description of stage | MCA/LA | Local Providers | MSS (Etio) | Evaluator (Ipsos and partners) | Product developer | DfE (as client) |
|---|---|---|---|---|---|---|---|
| 5 | Develop provider specification | N/A | N/A | R | C / A | C / A | C / A |
| 6 | Recruit institutions | C | C | R / A | C / I | N/A | A |
| 7 | Assign to treatment/control | I | I | R / I | R / I | N/A | A |
| 8 | Provide training | C / I | C / I | A | I | R | I |
| 9 | Implement the intervention | R | R | R / A | I | C | C / A |
| 10 | Contract management of providers, monitoring and reporting | C | C | R / A | I | N/A | A |
| 11 | Exit of MSS and final transfer of data | N/A | I | R | C | I | C / A |
| 12 | Measurement methods | I | I | C | R | N/A | I / A |
| 13 | Analysis, synthesis and reporting | C | C | C / I | R | N/A | A |
| 14 | Exit of Evaluation Supplier | N/A | N/A | N/A | R | N/A | C / A |

Source: Ipsos

# Annex 6: Summary of trial readiness packs (TRPs)

This Annex provides an overview of the content included in the Trial Readiness Packs (TRPs) for providers, schools and learners participating in the course-based Adult Numeracy Trials. It details the information provided to different audiences (providers, schools, tutors, teachers and learners) and their respective responsibilities in relation to the trial, including any differences treatment and control group participants.

TRPs were produced for the following trials: Preparation for GCSE Maths, Adapted Mastery Approach, Contextualised Approach, Family Numeracy and Embedded Maths.

---

## 1. Provider TRPs

The provider TRPs were written for adult education organisations participating in the trials. They did 5 things:

1. **Set the scene**
    - Explained that the organisation had agreed to participate in a research trial on adult numeracy.
    - Informed them that allocation to treatment or control (where relevant) had been done randomly.
    - Outlined, in a paragraph or two, what the trial was trying to find out.
2. **Spelled out what the organisation had to do**
    - Listed the main tasks and when they needed to happen (typically supported by a one-page checklist).
    - Made clear that learners could still attend their usual course whether or not they took part in the research.
3. **Described data and evaluation requirements**
    - Collect and share a standard set of learner details: learner reference numbers (if available), name, date of birth, postcode and contact details (email addresses and phone numbers).
    - Give each learner the opportunity to opt-out of their data being shared for the purposes of the research and / or to be contacted for surveys or interviews.

- Ensure tutors gave learners access to short online surveys at start of the course and arranged use of devices where needed.
- Take part in occasional interviews or feedback requests from the evaluation team.
- Noted that a sample of learners might be invited to take part in longer interviews, and would be offered a voucher as a thank you for their time and contribution.

4. **Explained how to handle existing administrative data requirements**
   Where relevant, the packs asked providers to:

   - Use a specific code or flag when submitting learner data to the Individualised Learner Record (ILR) so that trial participants could be identified for analysis.
   - Submit ILR returns in a timely manner for trial participants (for example, monthly rather than quarterly or waiting until the end of the academic year).
   - Understand that trial funding was provided via a grant arrangement, not via ILR returns.

5. **Covered technical and financial processes**

   - Set out how to use the Ipsos Data Collection Portal to upload learner data using supplied templates.
   - Described basic data protection points: DfE acted as Data Controller, evaluation partners processed data on a Public Task basis, and only necessary data were shared.
   - Outlined how the grant or supplementary funding worked, including:
     - Any flat participation payment.
     - Any administration premium linked to learner numbers.
     - For treatment providers, any payment linked to additional guided learning hours.
     - Gave step-by-step instructions for submitting invoices and who to contact with queries.

## Difference between treatment and control (providers)

In control group TRPs, providers were usually told to:

- Continue delivering their courses as they normally would (business-as-usual).

- Submit learner data via the Ipsos Data Portal, support administration of learner surveys, adjust ILR coding to flag trial participants, and be available for interviews.

In treatment group TRPs, providers were additionally asked to:

- Deliver a specific short course or set of sessions in place or alongside normal delivery.
- Ensure tutors were trained on the new content to be delivered.
- Record, for each session delivered: attendance, name of tutor and whether they had completed the required training, and whether the session was delivered in full or in part.

---

## 2. School TRPs

Where a trial took place in or through schools, the school-level TRPs typically focussed on:

1. **Context and rationale in school terms**
    - Short explanation of the trial, who it was for (e.g. parents/carers of children in Year 2) and what it was aiming to achieve.
    - Noted that schools were allocated to “sooner” or “later” delivery by random selection, with both groups making valuable contributions to the research.
2. **What the school was responsible for**
    - Recruiting eligible parents/carers using information sheets, leaflets and forms provided.
    - Collecting enrolment forms and sharing key details from these securely via the Ipsos Data Portal.
    - Working with a matched external adult learning provider to schedule sessions and make any room or timetable adjustments.
    - Arranging for children to join part of the joint sessions where that was part of the model.
    - Providing outcome data (for example, teacher-assessed maths grades) at agreed points.
3. **Data and information flows**

- Listed exactly what information about parents/carers and children needed to be entered into the secure Data Portal.
- Emphasised that personal data should never be emailed and instead only uploaded via the portal.
- Explained the legal basis for data collection (consent / opt-in) and data sharing (Public Task) and DfE's oversight role.

4. **Evaluation activities involving families and staff**

    - Described plans for short online surveys to be administered to parents/carers before and after the programme.
    - Noted that a sub-sample of parents/carers and staff might be invited to take part in longer interviews, with parents / carers offered a voucher as a thank you for their time and contribution.
    - Made clear that research activities were voluntary and separate from access to any parenting or learning support.

5. **Incentives for schools (if applicable)**

    - Explained whether there was a fixed incentive for recruiting a minimum number of eligible parents/carers.
    - Set out what counted as "eligible", how many parents needed to be recruited, and by when.
    - Described the invoicing process, payment timetable and contact details for the team managing these payments.

6. **Timetables and support**

    - Listed deadlines for uploading enrolment data and any cut-offs for recruitment.
    - Confirmed when schools would be told whether they were in the treatment or control group.
    - Gave brief details of any scheduled "keeping in touch" calls or support meetings and how to book them.

The main differences between the control and treatment group versions of school TRPs were around dates (with control delivery happening later), recruitment targets and invoicing windows.

---

## 3. Tutor TRPs

Tutor TRPs were developed for adult learning staff responsible for delivering specific interventions being tested. They tended to be shorter and more practical, typically covering:

1. **Why the tutor was involved**
    - Included a short statement that their organisation was taking part in an Adult Numeracy Trial.
    - Summarised the new / additional sessions they were being asked to deliver and how these sat alongside normal teaching.
2. **Training and materials**
    - Confirmed that tutors were required to attend a training session (often online) before teaching the new material.
    - Explained how to access recordings of the compulsory training if they could not attend at the scheduled time, and how any replacement tutors should be brought up to speed.
    - Explained where to download lesson plans and any “background and rationale” notes that detailed the thinking behind the design of the intervention.
3. **What tutors needed to do in class**
    - Delivering the extra sessions according to the timing and structure set out (for example, a set number of sessions before a given holiday, with remaining sessions by another point in the year).
    - Supporting learners to complete short online surveys at or near the start of the course.
    - Ensuring every learner had seen the relevant information sheet and knew about their participating in the trial and how to opt-out of research activities or their data being shared.
4. **Recording and reporting**
    - Keeping attendance records for each intervention session.
    - Noting where sessions were shortened, or skipped, and why.
    - Passing these records to the person in the organisation responsible for uploading data to the Ipsos Data Portal.
5. **Working with the evaluation team**

- Making tutors aware that some learners might be invited to one-to-one interviews and that they would be offered a voucher for their participation.
- Making tutors aware that they themselves might be invited to interviews or to contribute to online diaries or brief end-of-course surveys; participation was voluntary.

6. **Frequently asked questions**

    - How to use the QR code in class to enable learners to complete the baseline survey.
    - What to do if a learner did not have a device to complete the survey or struggled with digital tools.
    - How to explain particular survey questions (e.g. about qualifications or work status).
    - What support was available for learners who needed translation or additional help.

---

## 4. Teacher and school staff TRPs

Where school-based teachers or senior staff had a specific role in the trials, this was typically covered within the school-level materials or associated briefings. The content aimed to:

1. **Clarify their part in the process**

    - Helping promote the opportunity to parents (e.g. through letters, meetings or informal conversations).
    - Allowing children to attend sessions or activities.
    - Providing teacher-assessed child attainment grades or other internal data at agreed times.

2. **Flag any evaluation activities**
   Noting that teachers might be contacted for a short discussion about how the programme had worked in practice, with reassurance that this was optional and could be scheduled around other commitments.

---

## 5. Learner and parent/carer TRPs

Learner and parent/carer facing documents sat alongside the organisational TRPs. They were shorter, more straightforward, and focused on what participation meant for individuals. They generally included:

1. **Explanation of the research**
    - What was being studied and why.
    - Who was running the research (DfE and research organisations).
    - Why the individual had been invited (e.g. “you have enrolled on this course” or “you are a parent of a child at a participating school”).
2. **Information about choice**
    - A direct statement that taking part in the research was optional.
    - Reassurance that not taking part did not affect access to the course or support.
    - A straightforward way to opt out (or opt in for the Family Numeracy trial) and who to contact to do that.
3. **What participation usually involved**
    - Sharing a limited set of personal details so that records could be linked.
    - Completing one or 2 short online surveys, normally at the beginning and end of a course or programme.
    - In some cases, the possibility of being asked to take part in a longer interview, often with a voucher as a thank you for their contribution.
4. **Practical survey information**
    - When the survey should be done (ideally in a specific session, or by a certain date).
    - How to access it using a link or QR code, with simple instructions.
    - What devices could be used and what to do if a learner did not have access to a suitable device.
    - Whether the survey was only available in English and what support could be offered if that was an issue.
5. **How information would be used and protected**
   This was usually set out in two layers:

- A short “questions and answers” section covering:
    - What data were collected.
    - Who saw the answers.
    - That results were reported in summary form, without naming individuals.
- A more formal Privacy Notice that:
    - Named DfE as Data Controller and explained the role of research contractors.
    - Stated the legal basis for data being shared (Public Task).
    - Described how long identifiable data were kept and when they would be removed.
    - Listed data protection rights (such as access, correction, deletion, restriction) and how to exercise them.
    - Gave contact details for DfE, the research contractor and the Information Commissioner’s Office.

Across different trials and groups, the structure of these learner/parent materials was very similar. The main points of variation tended to be:

- Whether the default was inclusion unless the person opted out, or explicit opt-in.
- Whether a longer interview (with a voucher) was mentioned.
- The specific dates and deadlines used.

---

Taken together, the Trial Readiness Packs and associated information sheets formed a consistent framework for explaining each trial to key audiences. They clearly set out what each group was expected to do, and how data and participation in the evaluation would be managed across both treatment and control groups.

# Annex 7: Scoring approach

**Evaluation criteria used by DfE to score RCT ideas**

**1– 2 Limited evidence / Below threshold**

1: No evidence that the proposal meets the relevant criteria

2: Some evidence that the proposal meets the relevant criteria

**3 – Acceptable evidence / Meets threshold**

3. Good evidence that the proposal meets the relevant criteria (acceptable at this high-level stage)

**4-5 – Strong evidence / Above threshold**

4. Strong evidence that the proposal meets the relevant criteria

5. The proposal exceeds the necessary criteria

**Section 1: the need for a trial**

- The results will add to our understanding of what works in adult numeracy
- The reasons for the trial are explained
- The results will impact on real world practice
- The hypotheses to be tested and/or the study objectives are specified and described clearly

**Section 2: the proposed trial**

- The study design is appropriate to answer the research questions posed
- The outcomes, and their measures are clearly described and appropriate to answer the research question posed
- The trial population has been defined adequately in relation to the target population so that the results will have meaning
- The control group is appropriate

**Section 3: trial management and delivery**

- This trial will be developed and delivered in the time available

- The proposed budget to deliver the trial appears sufficient
- This idea could be easily implemented in order to test the outcomes
- Recruitment for the trial will be straightforward

**Section 4: Value for Money additional considerations**

- The RCT is the most efficient way available to source answers to the research question
- The suggested cost of the trial represents a reasonable spend to identify an answer to the research question
- Answers to the research question will not duplicate another trial being considered
- The trial is focussed on an area of significant spend

**Section 5: Risks**

For this section, the following scoring scale was used:

1 - serious ethical or risk concerns

2 - some ethical or risk concerns, but should be able to mitigate

3 - no ethical or risk concerns

A score of 2 or 3 is a pass and means that the trial can continue in the process at this stage. Scores of 1 would need to be discussed at moderation before the trial can continue in the process.

# Annex 8: Overview of product ideas put forward for Trial Feasibility Assessments

**Table A5: Overview of trials put forward for trial feasibility assessments**

| Number | Intervention name | Intervention description | Place in the learner journey | Outcome |
|---|---|---|---|---|
| 1 | Comms 1: language in digital advertising | Testing the effectiveness of different types of Multiply marketing materials, which differ in terminology to describe the maths programme. | Initial engagement | Passed stage 2 assessment but did not proceed through TFA gateway |
| 2 | FE Workforce | Testing the effectiveness of signposting teachers and tutors - who deliver maths courses to adult learners without formal teaching qualifications - to teacher training. | Ways of teaching (pedagogy) | Passed stage 2 assessment but did not proceed through TFA gateway |
| 3 | Numeracy Champions (employers) | Testing whether training 'numeracy champions' located within workplaces signpost staff to numeracy programmes. | Engagement, addressing barriers | Passed stage 2 assessment but did not proceed through TFA gateway |

| Number | Intervention name | Intervention description | Place in the learner journey | Outcome |
|---|---|---|---|---|
| 4 | Numeracy Champions (local areas) | Testing whether training 'numeracy champions' located within local areas signpost people with relevant needs to numeracy programmes. | Engagement, addressing barriers and supporting progression | Passed stage 2 assessment but did not proceed through TFA gateway |
| 5 | Homework at GCSE / FSQ and L1/L2 | Testing whether provision of homework leads to greater outcome achievement (GCSE/FSQ qualification). | Ways of teaching (pedagogy) | Passed stage 2 assessment but did not proceed through TFA gateway |
| 6 | Tutoring for GCSE/FSQ L2 | Testing the impact of small group tuition at the adult learner level on attainment in GCSE, FSQ or equivalent course attendance, completion, and learner confidence in maths (using pre-/post- survey evidence). | Delivery channels / places of learning | Passed stage 2 assessment but did not proceed through TFA gateway |
| 7 | Bite-sized learning | Testing the efficacy of delivering short numeracy sessions of around 20 minutes to employees in their workplace. | Identification of issues, addressing barriers, and delivery channels. | Passed stage 2 assessment but did not proceed through TFA gateway |

| Number | Intervention name | Intervention description | Place in the learner journey | Outcome |
|---|---|---|---|---|
| 8 | Confidence workshops | Investigating the effects of confidence-building workshops on learner's course attendance, completion rates, and their maths-related subjective stress and anxiety levels. | Addressing barriers and supporting | Did not continue after TFA |
| 9 | Personalised Assessment | Testing the impact of a personalised maths needs assessment for jobseekers on the uptake rate of maths courses provided under Multiply. The intervention would provide jobseekers who do not have GCSE level maths with a personalised needs assessment related to their maths ability. | Identification of barriers, engagement, addressing barriers, ways of teaching and supporting. | Did not continue after TFA |
| 10 | Count Me In | Testing to what extent the CMI programme in prisons helps to boost the numerical literacy of prisoners. | Addressing barriers, delivery channels / places of learning. | Did not continue after TFA |

| Number | Intervention name | Intervention description | Place in the learner journey | Outcome |
|---|---|---|---|---|
| 11 | Contextualised Curriculum | Testing to what extent an RME approach with adults and in adult FSQ level 1 classes leads to increased understanding of maths. | Addressing barriers | Proceed to trial (‘Contextualised Approach) |
| 12 | Maths by Stealth | The intervention consisted of Health and Social Care tutors engaging with a comprehensive package of pre-training to explain the principles of Embedded Maths. | Addressing barriers, delivery channels and ways of teaching. | Proceed to trial (‘Embedded Maths’) |
| 13 | Comms 2: motivational messages | Testing the effectiveness of 2 different letters (one urgent, one supportive) in encouraging adults who had completed a Multiply course to progress to a maths course that could result in a qualification. | Supporting progression | Proceed to trial (‘Encouraging Progression’) |
| 14 | Parental engagement / Family Learning | Testing to what extent learning sessions designed for adults and their children bolster both numeracy and family learning. | Addressing barriers, places of learning, ways of teaching, supporting and progression. | Proceed to trial (‘Family Numeracy’) |

| Number | Intervention name | Intervention description | Place in the learner journey | Outcome |
|---|---|---|---|---|
| 15 | Maths Mastery: general curriculum | The Adapted Mastery Approach (AMA) for 16-19-year-old learners in further education settings testing interactive teaching and group work, facilitated by adult education tutors, with the intervention focusing on 5 key principles for mastery of maths.[67] | Ways of teaching (pedagogy) | Proceed to trial ('Adapted Mastery Approach') |
| 16 | Preparation for Maths GCSE | Testing the impact of delivering preparation courses for maths GCSE/FSQ of around 26 hours to adults currently completing Level 1 maths courses. The intervention aims to build confidence in students to take up higher level maths GCSE courses. | Addressing barriers to learning, supporting and progression | Proceed to trial ('Preparation for Maths GCSE') |

Source: DfE and Ipsos

[67] The principles were 1) Teaching that allows learners to develop an understanding of mathematical structure, 2) Valuing and building on learners' prior learning, 3) Prioritising curriculum coherence and connections, 4) Developing both fluency and understanding of key ideas, and 5) Developing a culture in which everyone believes everyone can succeed.

# Annex 9: Template for assessing the feasibility of trial ideas

## Feasibility of trials template

### Purpose

This document provides a template and guidance to be followed when conducting feasibility assessments of Multiply trial ideas. Please work through each of the relevant sections and complete them as far as you can. There will be some elements that cannot be fully completed during the scoping phase of the study as we won't have full information on all the trial ideas being considered. Please note where this is the case.

### Summary

*This section should be completed last.*

**Table A6: Feasibility criteria**

| Criteria | RAG rating |
|---|---|
| Operational feasibility<br><br>(is the intervention well defined and is there a delivery organisation to do it?) | |
| Technical feasibility<br><br>(is an RCT feasible for this intervention, or is there a viable robust alternative?) | |
| Scheduling feasibility<br><br>(Can it be done by March 2025?) | |
| Economic feasibility<br><br>(Could it be done within the budget envelope?) | |
| Recommendation | |

**RED:** Relevant risks need to be mitigated to be able to proceed to commission an evaluation – if identified risks do not have a mitigation strategy by the first 6 months of the evaluation – do not proceed.

**AMBER:** Proceed with caution, some risks will need to be mitigated as the evaluation is scoped further, If mitigation strategy for identified risks not in place within the first 6 months of the evaluation – do not proceed.

**GREEN:** Proceed

## Detailed assessment

**Table A7: Detailed assessment criteria**

| Intervention characteristic | Description |
|---|---|
| **Intervention name:** | What is the intervention called? |
| **Which part of the learning cycle does this seek to address?** | **Which of these does it seek to address?**<br>Engaging Learners<br>Delivery Channels<br>Teaching Practice<br>Additional Support |
| **Type of intervention** | **Which of the following types of intervention apply?**<br>Product based<br>Activity based<br>Course based<br>DfE developed |
| **Bolt-on or innovation?** | Are we going to do this as part of an existing practice (bolt-on) or is this an innovation (new service) that would need to be introduced? |

| Intervention characteristic | Description |
|---|---|
| **Intervention Description** | Here provide a brief description of the intervention: origins, intended outcomes, and delivery model. Include why you think this intervention might produce an effect on outcomes, including any evidence you have.<br><br>You may want to use some of these prompts to guide you (but you don't need to answer all of them):<br><br>How are participants identified and referred to the service?<br><br>What are the key components/activities of the service?<br><br>How frequently are these activities carried out? (e.g. 4, weekly training sessions are delivered)<br><br>Where is the programme or service delivered? (e.g. within adult education institutions)<br><br>How is the service delivered? (e.g. group or individual setting, face-to-face, telephone service, online)<br><br>Who delivers the service (e.g. Adult Education Providers, community organisations, Local Authorities, DfE)?<br><br>In some cases, you might not have a well-defined intervention but a practice that needs to be manualised. If that is the case, this additional time needs to be reflected in the scheduling. |
| **Current situation** | • Where has it currently been tried?<br>• What is the state of the evidence base (e.g. are there any studies or reviews on this intervention)? What did those find? |

## Operational feasibility

**Table A8: Operational feasibility criteria**

| Intervention characteristic | Description |
|---|---|
| **Targeting of the intervention / referral mechanisms** | • Who are the potential participants (could it be applicable for multiple groups)?<br>• What are the eligibility criteria (if known)<br>• Are participants already identified or do they need recruiting?<br>• How can participants be contacted?<br>• Who is the service provider?<br>• What is the setting where services are provided?<br>• Is there a very clear referral mechanism into the study? Can we identify participants and where they are? Do we have an indication of volume? (Tip: you might want to consider the number of people/learners who are eligible for the trial in a given quarter or last year who are consistent with the eligibility criteria)<br><br>**RED:** Substantial challenges to identify participants, with very small samples that are likely to take part.<br>**AMBER:** Eligibility criteria needs to be refined, or there are some challenges to identify who the participants are. OR, caseload is likely to be small.<br>**GREEN:** Eligibility criteria is clear and participants can be readily identified. There is a substantial caseload to choose from. |

| Intervention characteristic | Description |
|---|---|
| **Delivery of the intervention** | **Who will deliver the intervention?**<br>• Are there organisations in England willing and able to deliver the intervention? These could be organisations already delivering, organisations with similar skills and/or organisations identified through the call for practice?<br>• Is there a managed service supplier in place?<br>• Who will be the delivery partners on the ground (e.g., Colleges)<br>**What is BAU?**<br>• Describe what 'business as usual' is.<br>• Is there a clear group receiving similar support from other agencies?<br>• How different is this intervention from BAU (i.e. 'differentiation')? Is this intervention done on top of BAU (i.e. 'additionality')?<br>• Is information needed from the DfE Sector Advisory Group about BAU?<br><br>**RED:** Intervention that needs better definition and clear boundaries to be set before it could be evaluated. Or, intervention with very limited 'differentiation' and 'additionality' to BAU.<br>**AMBER:** Well-defined intervention or intervention with small adaptations, and requires identifying a delivery organisation<br>**GREEN:** Well-defined intervention with clear organisations that could deliver it. |

## Technical feasibility

**Table A9: Technical feasibility criteria**

| Technical element | Description |
|---|---|
| **Methods** | • If this is an RCT, tell us what type;<br>  • At what level do we recommend that randomisation occurs?<br>  • How will randomisation operate in reality and how would it fit into the trial set up process?<br>  • Is it a parallel design? Cross-over? Stepped-wedge?<br>  − Is it a single or multi-site trial?<br>  • Is it a complex or rapid trial?<br>  • Is non-inferiority or superiority trial? Do you know what the control group will be?<br>  • Is this phase I (pilot), phase II (efficacy - ideal conditions) or phase III (effectiveness - real conditions)?<br>Studies are expected to be 2-arm trials, unless there is a strong argument for multiple arms (e.g. small variations or nudges).<br>We anticipate that RCTs are the default for all of these studies. However, if a QED approach would be preferable and possible, please describe and specify the approach you would propose. This may include:<br>• Difference-in-Differences approaches (when intervention starts in multiple areas but not others)<br>• Statistical matching (when a comparison group of similar units can be created using existing data)<br>• Synthetic controls (when a single place starts to do something, and there are no other changes happening in that area)<br>• Avoid proposing theory-based evaluations (these would be out of scope)<br>If more than one option is possible, please outline them and explain your preferred option.<br>**RED:** An RCT is not possible and QED methods have very substantial risks; or, only methods that are not based on counterfactual are possible (e.g. process tracing, before-and-after comparisons)<br>**AMBER:** An RCT is possible, but there are substantial risks that need to be mitigated for it to be feasible; or a QED is feasible, with some limited risks.<br>**GREEN:** An RCT is deemed as feasible or only with limited risks. |

| | |
|---|---|
| **Sample Size** | • Considering costs and/or previous studies, what would you expect the Minimum Detectable Effect of Relevance to be? (small=0.2; medium=0.35; large=0.5). Please specify the target MDES. This should be informed by your assessment of the intensity of the intervention, the costs of the intervention, and the outcome measure being considered.<br><br>Please provide a worked example of what this means in the context of the trial being scoped and its specific outcomes. For example, if the Cohen's d is used as a measure of effect size:<br>• Primary outcome: Achievement of a GCSE level 2 math qualification (0/1), assuming a binomial distribution.<br>• Proportion of individuals in the treatment group who achieve the qualification = 0.3, with a standard deviation of 0.46.<br>• Proportion of individuals in the control group who achieve the qualification = 0.2, with a standard deviation of 0.4.<br>• The pooled standard deviation is 0.43 in this example.<br>• Cohen's Delta would be: (0.3-0.2)/0.43= 0.23. This means that the difference between the proportion of treated and control individuals in the trial achieving a GSCE level 2 math qualification is around 1/5 the variability in the sample, i.e., a difference of 13 percentage points on the original scale with the assumed standard deviation.<br>• Using the standard assumptions for the design you have proposed, how many participants would you include (in treatment and control)? To do the calculations, please use this spreadsheet and this guidance.<br>• What is the expected effective sample after attrition and non-consent?<br>• What is the number that needs to be randomised and What is the number of referrals needed to reach that figure?<br>• If you would like to propose any changes to the assumptions, please do so and specify.<br><br>You could propose 2 options: the minimum viable size and a well-powered trial.<br>**RED:** The sample size would require operating in a much larger scale or reaching a very large number of Local Authorities or partners, increasing the complexity of the project.<br>**AMBER:** The sample size could be challenging in the time period needed based on the information on referral rates, so a large number of Local Authorities or partners would be required.<br>**GREEN:** The sample size is attainable in a reasonable time period (see scheduling) based on the information available about referral rates. |

| Technical element | Description |
|---|---|
| **Outcome measure** | • Specify here the primary and secondary outcome measures for this trial.<br>• In general, primary outcomes are expected to be related to adult numeracy. This could include enrolment in a course; number (or proportion) of sessions attended, whether or not a participant passes, or their grade (if appropriate) but there are some circumstances when this would not be appropriate. For example, these might be passed further down stream, such as labour market outcomes, or in the case of marketing / outreach trials, more proximately, such as clicks through. In those cases please explain your choice of primary outcome.<br>• Specify when each outcome might need to be measured (e.g. short term or long term). This could be informed by previous studies and/or by the intervention's theory of change.<br>• Specify how data could be collected (admin data or survey-based)<br>**RED:** Outcomes data would require data matching between multiple government departments or agencies, which could prove challenging.<br>**AMBER:** Outcomes are well defined but would require extensive and costly primary data collection, or rely on administrative data that might be difficult to access, clean and present (e.g. data from LAs)<br>**GREEN:** Outcomes are well defined and could be collected through readily available administrative data from DfE |

## Scheduling

**Table A10: Scheduling feasibility criteria**

| Scheduling element | Description |
|---|---|
| **Timeline** | Please provide rough estimates for each phase:<br><br>**Programme design (optional):**<br><br>• Is this a well-defined intervention or is it necessary to develop it? If so, allow for at least 3 months.<br><br>**Recruitment of delivery organisation and participants:**<br><br>• Are sites and delivery partners lined up or do they need to be recruited?<br>• Will participants be recruited all at once or on a rolling basis?<br>• When would the last person be recruited?<br>• Implementation of the intervention<br>• How long is the intervention?<br>• When would the last person complete the intervention?<br><br>**Measurement**<br><br>• How long do you need to wait after the intervention to collect outcome data?<br>• If outcome data is from admin sources, consider how long it might take to be reported and accessed for the last person?<br>• When could we have results? (Usually allow for 3 months for reporting after data collection ends)<br><br>**RED:** Relevant outcomes for the sample needed are likely to be available after March 2026.<br>**AMBER:** Short-term outcomes / interim report are expected to be available before March 2026, even if other outcomes might be needed after that point<br>**GREEN:** All relevant outcomes are expected to be measured before March 2026 |

## Economic feasibility

**Table A11: Economic feasibility criteria**

| Economic element | Element description |
|---|---|
| **Estimated costs** | **Provide ballpark estimates of costs for:**<br><br>Delivery. Consider all costs of a treatment, not just those incurred by the client. Please include the total cost and the cost per client. If no information is available, say so, but you might want to give a ballpark figure - even if it is only indicative at this point.<br><br>Evaluation. Between 200 and 400k, depending on the methods and complexity of the project.<br><br>Total. Please include the total for the project.<br>**RED**: Total budget over 1.5M, reducing the number of projects that could be conducted.<br>**AMBER**: Total budget between 750k and 1.5M<br>**GREEN**: Total budget under 750k |

# Annex 10: Overview of the interventions taken forward to trial

**Table A12: Overview of interventions taken forward to trial**

| Intervention name | Intervention description | Proceeded to trial? | Rationale / existing evidence | Link to published protocol |
|---|---|---|---|---|
| Preparation for Maths GCSE | This involved delivery of 7 lessons of one hour each to adult learners on GCSE maths courses. The lessons took place within existing GCSE curriculum hours and covered theories of resilience and growth mindset, study skills, revision and exam skills; designed to be delivered face to face but could be adapted for online or hybrid courses. Mathematics in Education and Industry (MEI) was sub-contracted by National Numeracy[68] to develop the intervention. | Yes | A review of existing evidence conducted by MEI did not find any examples of study skills courses specifically designed for GCSE adult maths learners. General study courses were identified, but the content was less relevant to maths. There was some evidence that teaching cognitive and meta-cognitive skills can help school aged learners, including to address learning difficulties in maths[69]. The Pilot Trial tested whether adult learners who were explicitly taught study and revision skills relevant to maths and were prepared for the additional work that comes with a maths GCSE were more likely to have better outcomes than those who were not. | OSF Registries \| 'Preparation for Maths GCSE' study skills sessions embedded into Maths GCSE - Pilot RCT |

[68] An independent charity aimed at improving numeracy skills in the UK commissioned by DfE to develop trial courses

[69] e.g. Behzadi et al., 2014; Schneider & Artelt, 2010; Nolting, 1994

| Intervention name | Intervention description | Proceeded to trial? | Rationale / existing evidence | Link to published protocol |
|---|---|---|---|---|
| Encouraging Progression | The trial tested the effectiveness of 2 different letters in encouraging adults who had completed a Multiply course that did not lead to a qualification, to progress to a Maths course that could result in a qualification. The 2 letters tested were sent from senior officials at DfE. One included a quote from a peer saying how Multiply had made them feel more confident about undertaking further learning in Maths. The letter encouraged them to sign up for a Maths course expected to result in a qualification (a 'peer support' letter). The second letter highlighting 'urgency' in signing up for a Maths course likely to result in a qualification. This highlighted the potential for improved numeracy skills to enhance career prospects. (an 'urgent' letter). | Yes | DfE and local authorities send out letters to learners to advertise courses to try and encourage lifelong learning. This was the first systematic attempt to assess the effectiveness of such communications on learners' subsequent enrolment behaviour. | OSF \| Encouraging Progression of Multiply learners to qualification-bearing courses using targeted letters - a 3-arm randomised control trial |

| Intervention name | Intervention description | Proceeded to trial? | Rationale / existing evidence | Link to published protocol |
|---|---|---|---|---|
| Embedded Maths in Health and Social Care | The intervention consisted of Health and Social Care tutors engaging with a comprehensive package of pre-training to explain the principles of Embedded Maths. Tutors were supported by ongoing continuing professional development (CPD) in the form of reflective practice sessions led by expert practitioners. The 18 guided learning hours (GLH) for students covered approximately 60% of the FSQ Level 1 maths curriculum. The intervention covered: interpreting data and averages; understanding units, precision and context; measuring; using scales; budgeting; cost and profit; understanding time; recording and presenting data; understanding temperature; and charts and graphs, as these relate to working in H&SC. Due to the elevated levels of maths anxiety faced | Yes[70] | The intervention was adaptable to courses of different lengths, such as 15 or 30 weeks, and delivery modes including face-to-face or online. The intervention designers intended that it could be adapted to fit within most adult Level 2 Health and Social Care courses. Moreover, a unique feature was that the sessions met both the needs of FSQ Level 1 Maths and the maths component of Health and Social Care qualifications.<br><br>Whilst the intervention was considered feasible for trailing, challenges with recruitment of adult education providers and high levels of attrition (described in more detail in Section 6:Attrition), meant that the evaluation was taken forward as a small-scale, non-randomised descriptive, implementation and process study with 2 | Not published |

[70] Started as a trial but later pivoted to an Implementation and Process Evaluation (IPE) due to high levels of provider attrition – see Section 6: Attrition for further details.

| Intervention name | Intervention description | Proceeded to trial? | Rationale / existing evidence | Link to published protocol |
|---|---|---|---|---|
| | by adults, the Embedded Maths curriculum was intended to work by providing a supportive and engaging environment to overcome barriers to learning. The product was developed by the Education and Training Foundation, who also provided the tutor training. | | providers (rather than the larger scale efficacy trial originally developed for). | |

| Intervention name | Intervention description | Proceeded to trial? | Rationale / existing evidence | Link to published protocol |
|---|---|---|---|---|
| Contextualised Approach for Functional Skills Qualification (FSQ) Maths Level 1 | The intervention was designed to be delivered within adult Functional Skills Level 1 maths classes by functional skills tutors. While classes could be mixed ability, only learners working towards FSQ Level 1 were eligible for the trial. Tutors were expected to deliver 12 weekly 1.5-hour contextualised sessions to learners aligned with the existing lesson study schedule. The intervention involved testing a Realistic Mathematics Education (RME) approach in adult Functional Skills maths Level 1 classes. RME is a method which builds maths knowledge from contexts that are meaningful to learners. The product was developed by the Education and Training Foundation and the RME team from Manchester Metropolitan University. | Yes | Prior studies have used the RME approach within primary schools and up to Key Stage 4, demonstrating positive qualitative and mixed quantitative outcomes. Until this pilot trial, the RME approach had not been trialled in England with adult learners undertaking Functional Skills Level 1 maths. Furthermore, the intervention was developed with flexibility to accommodate diverse FE settings. The product developer provided information on the degree of tailoring tutors could apply, such as varying contexts to better match learners' experience/recognition, but not the maths content. The intervention was deemed proceedable for a 3-armed, pilot RCT (with the third arm including Adapted Maths Mastery – see below). | OSF \| Multiply: An Adapted Mastery Approach and Contextualised Approach for FSQ Level 1 maths |

| Intervention name | Intervention description | Proceeded to trial? | Rationale / existing evidence | Link to published protocol |
|---|---|---|---|---|
| Adapted Maths Mastery approach FSQ Maths Level 1 | The Adapted Mastery Approach (AMA) intervention was an adaption of an existing intervention for delivery to adult learners. The Maths Mastery (MM) approach was originally developed in the UK for teaching in primary and secondary schools. The DfE-funded Centres for Excellence in Maths programme (CfEM) updated the approach for GCSE resit and FSQ Levels 1 and 2 teaching for 16-19-year-old learners in further education settings. Delivery incorporated interactive teaching and group work, facilitated by adult education tutors, with the intervention focusing on 5 key principles for mastery of maths.[71] | Yes | Previous experimental and quasi-experimental trials of mastery interventions have found promising results. Assessment of evidence from quasi-experimental trials, meta-analyses and observational studies pointed to practices that showed promise, in relation to potential impact of mastery pedagogy on pupil attainment (Boylan et al., 2018). The efficacy of the mastery approach has also been evaluated in a randomised controlled setting for older participants, with more promising results. (Wake et al., 2023). The intervention was deemed proceedable for a 3-armed, pilot RCT (with the third arm including Contextualised Curriculum – see above). | OSF \| Multiply: An Adapted Mastery Approach and Contextualised Approach for FSQ Level 1 maths |

[71] The principles were 1) Teaching that allows learners to develop an understanding of mathematical structure, 2) Valuing and building on learners' prior learning, 3) Prioritising curriculum coherence and connections, 4) Developing both fluency and understanding of key ideas, and 5) Developing a culture in which everyone believes everyone can succeed.

| Intervention name | Intervention description | Proceeded to trial? | Rationale / existing evidence | Link to published protocol |
|---|---|---|---|---|
| Family Numeracy | The intervention featured a programme of structured, in-person group learning sessions designed for adults and their children to bolster both numeracy and family learning. The programme focussed on family learning skills and practise with a focus on addressing 'maths anxiety' and confidence to support whole family maths learning, leading to child progression in formal education, adult progression onto further numeracy interventions or programmes of study, and adult opportunities to expand or progress workplace numeracy use. The intervention targeted adults and children: adults as parents/carers, as numeracy learners and as employees/workers; and children as family members and learners. The intended outputs and outcomes from taking part were split into beneficial effects for both parents/carers and children exposed to the intervention. | Yes | Many Family Learning studies focus on literacy rather than numeracy. This trial was an opportunity to test how Family Learning impacted parent maths confidence, as well as parents'/carers' progression in maths and children's maths attainment. A pilot RCT was conducted to examine the feasibility of delivering Family Numeracy within a school setting, delivered by adult learning providers. | OSF \| The impact of whole family maths courses on parents' progression to further maths courses – a 2-armed cluster randomised controlled trial |

Source: Ipsos

# Annex 11: Analytical framework

**Table A13: Analytical framework and approach agreed**

| Analytical consideration | Approach agreed |
|---|---|
| Descriptive statistics (baseline stats per group) | A subset of variables will have to be reported using the same coding (e.g. sex, age, ethnicity, disability status, FSM status, residence in area on top quintile of deprivation index). Additional variables are optional (e.g. prior level of attainment, number of learners per class, etc.). |
| ILR data, key basic variables included for analysis (incl. baseline characteristics and outcomes) | i) The primary outcome will depend on each trial; ii) confidence in maths will share the same scale across trials; iii) baseline variable is not included in most trials according to protocol. |
| ILR data vars coding (categories) for impact estimation (e.g. age, gender, ethnicity, FSM status, outcome scale/categories) | Regression models will include the same coding for shared variables across trials. |
| Outcome descriptive analysis (including endline and baseline statistics per group, if available) | Outcome analysis pre and/or post will be optional. |
| Outcome data coding and standardisation procedure | Estimation will be implemented using non-standardised coefficients. Results will be reported using Hedge's G (doing calculation manually ex-post). |
| Balance assessment (formal comparison between treatment vs control group at baseline) | All trials reporting unconditional t-test (p-value reported) to assess balance across variables. EP trial reporting this in appendix. |

| Analytical consideration | Approach agreed |
|---|---|
| Model specifications (e.g. empty/full model) | Model specifications will depend on each trial. Main results will be reported for a "full" model as described in the protocol, which incorporates all possible confounders as covariates to improve statistical power and precision of estimates. When the same variables are used these will be coded in the same way across trials. Sensitivity analyses for different specifications could be included in appendix. These analyses are not mandatory unless included in the protocol. |
| Data clustering/standard error estimation | Clustering will be taken into account using STATA's analytical formula (Rogers, 1998), regardless of number of clusters. Bootstrapping will not be implemented. |
| Missing data management | Decisions will be made once trial leads review data completeness at endline. Flexibility will be the main criteria. As general guidelines: i) Outcomes: No imputation of endline or baseline for primary outcome will be performed. Initially, no imputation of endline for secondary outcomes either. However, for secondary outcomes, the option of imputation remains open for Multiple Imputation or mean imputation, depending on response rates and missingness. ii) "Secondary" vars (covariates, non-outcome data): We will implement multiple imputation including chained equations (MICE) for all variables with missingness above the 5% threshold. If missingness is too high, we will assess mean imputation. |
| Common effect size metric used | Hedge's G will be used for all trials. |
| Primary Analysis | Primary analysis will use regression framework. Covariates included are specified in each protocol. No multilevel (random effects) modelling is considered. |

| Analytical consideration | Approach agreed |
|---|---|
| Secondary Analysis | Secondary analysis will use a regression framework. Covariates included are specified in each protocol. No multilevel (random effects) modelling is considered. |
| Subgroup impact analysis/heterogeneous impacts | The use of interaction terms for subgroup analysis is preferred unless otherwise specified in the protocol. If subgroup regression analysis using separate subsamples is implemented, an analysis that includes results for similar models with interaction terms will be included in appendix. These subgroup analyses are likely to be highly underpowered in most trials. |
| Multiple hypothesis testing | Does not apply. |
| Compliance Analysis | 2SLS estimation will be implemented with definition of compliance according to set criteria per protocol. |
| Robustness checks included | Analysis based on compliance levels (CACE), with cut off point (minimum dosage) depending on each trial. |
| Effect size interpretation | Trialists will try to reach similar language at reporting stage, in particular for final summary. |
| Measures reported for binary outcomes (odds vs marginal effects) | Average Marginal Effects will be used across all trials. |
| Effect size reporting (e.g. confidence intervals, p-values) | p-values and SEs to be reported for all trials. 95% Confidence Intervals could be included for programme level reporting. |
| Key ex-post metrics to be reported: Ex-post power, Attrition, Balance. | Attrition and balance will be reported. This reporting is mandatory for primary outcomes. Reporting for secondary outcomes optional and based on individual trial findings. Ex-post power reporting optional and in appendix. |

| Analytical consideration | Approach agreed |
|---|---|
| Attrition diagram | Attrition diagrams will be included only when provider-level dropouts occurs, or if learners formally opt-out of the trial. |

Source: Ipsos

# Annex 12: Survey response rates

This Annex presents the call outcomes for all attempted telephone surveys[72], using the definitions in Table A14.

[72] The total number of outcomes exceeds the total number of learners who completed interviews because multiple attempts were made per learner.

**Table A14: Call outcomes and their definitions**

| Call outcome | Definition |
|---|---|
| Abandoned | Respondent started the interview but declined to continue at some point |
| Answer machine | Call was answered by answering machine |
| Appointment made but survey not completed | Respondent did not start survey but made an appointment to be called back to complete it in the future, however they did not complete it when called back |
| Busy | Respondent's phone was busy, such as in the process of another phone call when they were dialled |
| Request to call back another time | Respondent requested to be called back another time but did not express an intention to complete the survey when called back |
| Completed | Completed the survey |
| No answer | The phone was not picked up |
| Disconnected | The phone number was invalid |
| Language barrier | The respondent could not complete the survey due to a language barrier |
| Not available | The designated respondent was not available within the timescale of the project, even if they were willing to take part |
| Refused | Respondent answered the call but did not consent to start the survey |
| Wrong number | The call was answered but the person was not the intended respondent on the sample |
| Other | Any other call outcome, such as being a business number or the respondent made a ‘do not contact’ request for all Ipsos calls in the future |

**Additional definitions:**

**Response rates** were calculated based on the proportion of learners whose details were uploaded to the Ipsos Data Portal who completed the survey. This includes learners who were eventually excluded from the final analysis. It also includes sample cases deemed to be ‘unusable’ due to being incorrect phone numbers for the intended participant, deadlines, or invalid phone numbers.

**CATI response rates** are based on all records with phone numbers. Online response rates are based on all records with email addresses.

**Cooperation rates** are calculated from the number of completes divided by the sum of respondents who answered the call and were the intended respondent. This includes completed telephone interviews, refused telephone interviews, and interviews that couldn’t be completed due to a language barrier or other communication difficulty. Refused interviews are where the participant responded to the phone call but refused to continue with the interview.

## Baseline survey

For the online baseline surveys, it was not possible to develop an accurate response rate because it cannot be determined how many learners received the link. The mode of distribution relied on providers distributing this, in the form of a QR code embedded within an information sheet, to learners.

**Table A15: Family Numeracy baseline cooperation rate**

| Description | Rate |
|---|---|
| CATI cooperation rate | 102/158 (65%) |

**Table A16: Family Numeracy baseline call outcomes**

| Call outcome | Frequency | Percentage |
|---|---|---|
| Abandoned | 10 | 2% |
| Answer machine | 152 | 36% |
| Appointment made but survey not completed | 1 | 0% |
| Busy | 13 | 3% |
| Request to call back another time | 4 | 1% |
| Completed | 102 | 24% |
| No answer | 19 | 4% |
| Disconnected | 1 | 0% |
| Language barrier | 20 | 5% |
| Not available | 1 | 0% |
| Refused | 20 | 5% |
| Wrong number | 74 | 17% |
| Other | 6 | 1% |
| **Total** | **423** | **100%** |

**Table A17: Preparation for Maths GCSE baseline cooperation rate**

| Description | Rate |
|---|---|
| CATI cooperation rate | 42/75 (56%) |

**Table A18: Preparation for Maths GCSE baseline call outcomes**

| Call outcome | Frequency | Percentage |
|---|---|---|
| Abandoned | 5 | 2% |
| Answer machine | 109 | 42% |
| Appointment made but survey not completed | 0 | 0% |
| Busy | 9 | 4% |
| Request to call back another time | 2 | 1% |
| Completed | 42 | 16% |
| No answer | 19 | 7% |
| Disconnected | 4 | 2% |
| Language barrier | 5 | 2% |
| Not available | 0 | 0% |
| Refused | 21 | 8% |
| Wrong or unallocated phone number | 34 | 13% |
| Other | 7 | 3% |
| **Total** | **257** | **100%** |

**Table A19: Maths FSQ baseline cooperation rate**

| Description | Rate |
|---|---|
| CATI cooperation rate | 61/88 (69%) |

**Table A20: Maths FSQ baseline call outcomes**

| Call outcome | Frequency | Percentage |
|---|---|---|
| Abandoned | 9 | 4% |
| Answer machine | 73 | 32% |
| Appointment made but survey not completed | 2 | 1% |
| Busy | 5 | 2% |
| Request to call back another time | 4 | 2% |
| Completed | 61 | 26% |
| No answer | 14 | 6% |
| Disconnected | 2 | 1% |
| Language barrier | 4 | 2% |
| Not available | 0 | 0% |
| Refused | 8 | 3% |
| Wrong or unallocated phone number | 37 | 16% |
| Other | 12 | 5% |
| **Total** | **231** | **100%** |

## Endline survey

For the online Maths FSQ endline survey, it is not possible to develop an accurate response rate because it cannot be determined how many learners received the link. The mode of distribution relied on providers distributing the link, in the form of a QR code embedded within an information sheet, to learners.

**Table A21: Family Numeracy endline response rate and cooperation rate**

| Description | Rate |
|---|---|
| CATI cooperation rate | 97/179 (54%) |
| CATI response rate | 97/531 (18%) |
| Online response rate | 48/570 (8%) |

**Table A22: Family Numeracy endline call outcomes**

| Call outcome | Frequencies | Percentage |
|---|---|---|
| Abandoned | 23 | 4% |
| Answer machine | 181 | 34% |
| Appointment made but survey not completed | 12 | 2% |
| Busy | 13 | 2% |
| Request to call back another time | 6 | 1% |
| Completed | 97 | 18% |
| Language barrier | 33 | 6% |
| No answer | 26 | 5% |
| Not available | 2 | 0% |
| Refused or opt-out | 39 | 7% |
| Wrong or unallocated phone number | 76 | 14% |
| Other | 43 | 8% |
| **Total** | **531** | **100%** |

**Table A23: Preparation for Maths GCSE endline response rate and cooperation rate**

| Description | Rate |
|---|---|
| CATI cooperation rate | 207/256 (81%) |
| CATI response rate | 207/842 (25%) |
| Online response rate | 228/967 (24%) |

**Table A24: Preparation for Maths GCSE endline call outcomes**

| Call outcome | Frequencies | Percentage |
|---|---|---|
| Abandoned | 11 | 2% |
| Answer machine | 182 | 29% |
| Appointment made but survey not completed | 1 | 0% |
| Busy | 18 | 3% |
| Request to call back another time | 5 | 1% |
| Completed | 207 | 33% |
| Language barrier or other communication difficulty | 6 | 1% |
| No answer | 34 | 5% |
| Not available | 4 | 1% |
| Refused | 21 | 3% |
| Wrong or unallocated number | 126 | 20% |
| Other | 4 | 1% |
| **Total** | **619** | **100%** |

**Table A25: Maths FSQ endline response rate and cooperation rate**

| Description | Rate |
|---|---|
| CATI cooperation rate | 77/100 (77%) |
| CATI response rate | 77/369 (21%) |

**Table A26: Maths FSQ endline call outcomes**

| Call outcome | Frequencies | Percentage |
|---|---|---|
| Abandoned | 4 | 2% |
| Answer machine | 68 | 28% |
| Busy | 3 | 1% |
| Request to call back another time | 1 | 0% |
| Completed | 77 | 31% |
| Language barrier | 10 | 4% |
| No answer | 11 | 4% |
| Not available | 7 | 3% |
| Refused | 8 | 3% |
| Do not contact request or other | 10 | 4% |
| Wrong or unallocated number | 48 | 19% |
| **Total** | **247** | **100%** |

Source: Ipsos

# Annex 13: Data flows and matched data information

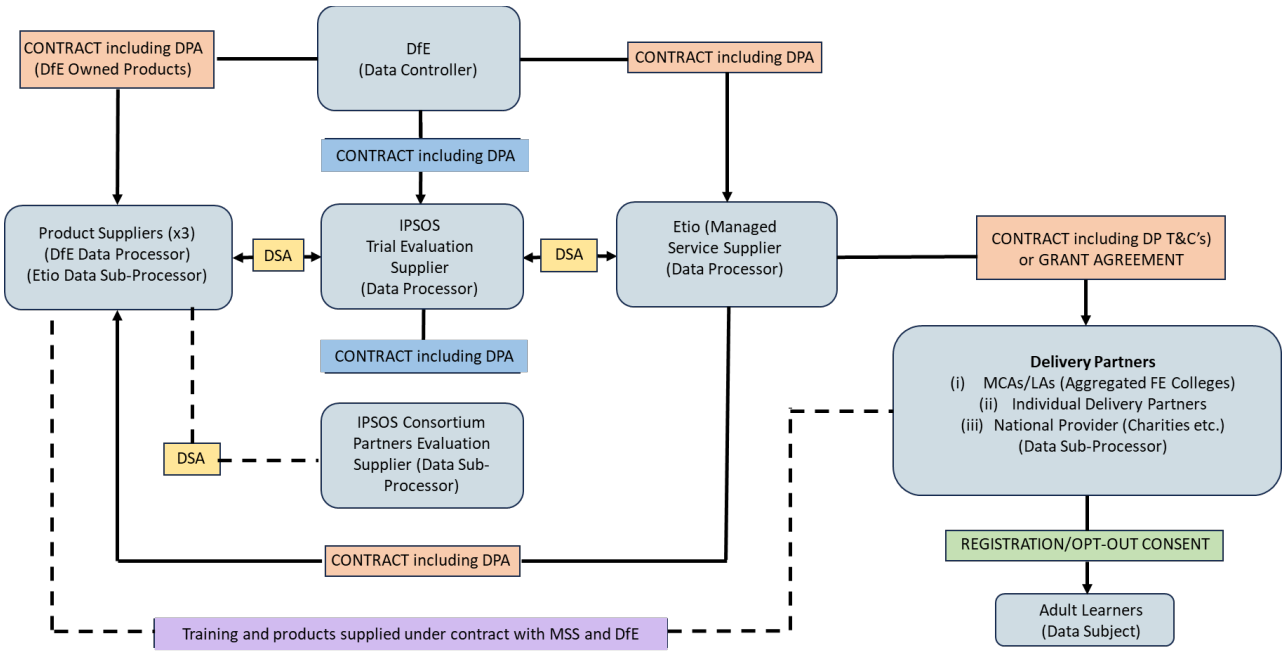
**Table A27: Total number of matches between Kicktag, ILR, survey and child attainment datasets**

| Dataset | Key 1 | Key 2 | Key 3 | Key 4 | Key 5 | Number of matches |
|---|---|---|---|---|---|---|
| ILR | Unique learner number | UKPRN | Date of birth | Post code | z | 947 |
| ILR | Unique learner number | Full name | Date of birth | Post code | z | 499 |
| ILR | Unique learner number | UKPRN | Full name | Date of birth | z | 52 |
| ILR | Unique learner number | UKPRN | Full name | z | z | 11 |
| ILR | Unique learner number | Full name | Date of birth | z | z | 46 |
| ILR | UKPRN | Learner reference number | Full name | Date of birth | Post code | 78 |
| ILR | UKPRN | Learner reference number | Full name | Date of birth | z | 5 |
| ILR | UKPRN | Learner reference number | Full name | z | z | 1 |
| ILR | UKPRN | Learner reference number | z | z | z | 64 |
| ILR | Full name | Date of birth | Post code | Email | z | 174 |
| ILR | Full name | Date of birth | Post code | z | z | 249 |
| Survey | Unique learner ID | z | z | z | z | 1,532 |
| Survey | Full name | Date of birth | Post code | z | z | 127 |

| Dataset | Key 1 | Key 2 | Key 3 | Key 4 | Key 5 | Number of matches |
|---|---|---|---|---|---|---|
| Survey | Full name | Date of birth | Post code | z | z | 498 |
| Survey | Full name | Post code | Email | z | z | 4 |
| Child attainment | School name | Child first name | Child surname | z | z | 425 |

Source: Ipsos, z = not applicable

## Figure A1: Data flow for management information from providers




Source: DfE

**Figure A2: Process flow for the Ipsos Data Portal**

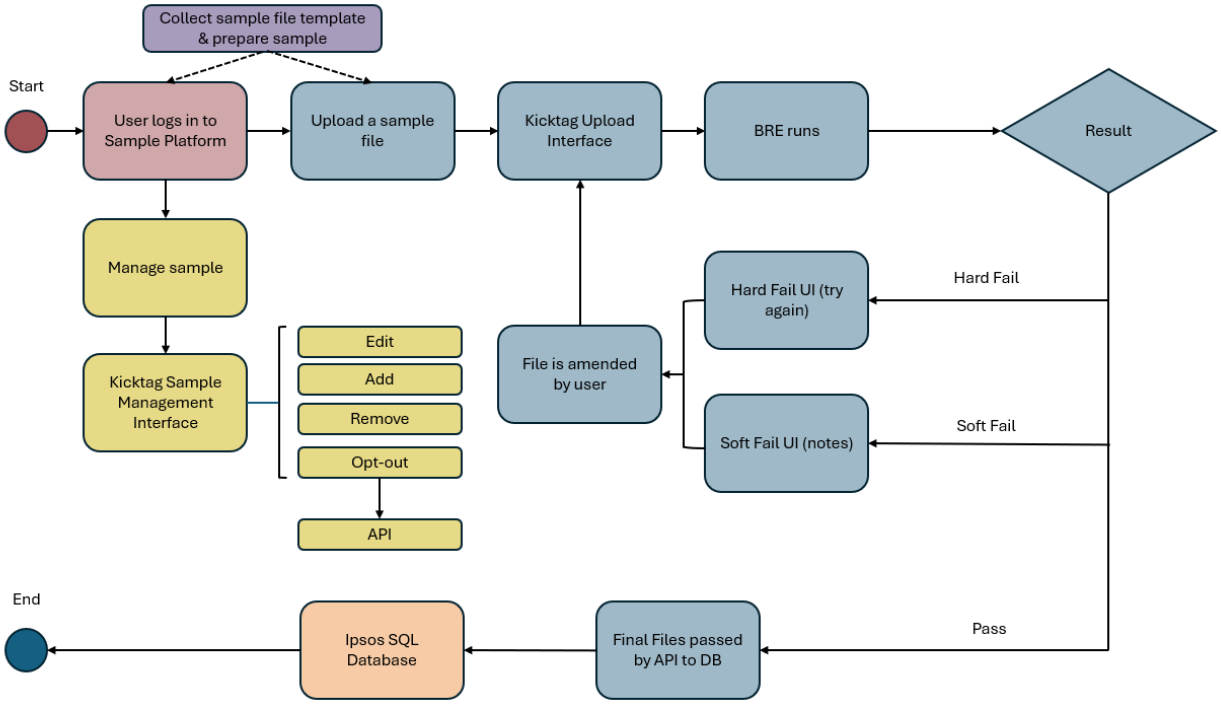



**Table A28: Kicktag match results to ILR**

| Trial | Matched | Unmatched | Total |
|---|---|---|---|
| Contextualised curriculum | 202 | 2 | **204** |
| Embedded maths | 67 | 1 | **68** |
| Family numeracy | 231 | 342 | **573** |
| FSQ Level 1 | 503 | 12 | **515** |
| Maths mastery | 489 | 5 | **494** |
| Preparation for maths GCSE | 1,024 | 34 | **1,058** |
| **Total** | **2,516** | **396** | **2,912** |

Source: Ipsos

**Table A29: Kicktag match results to survey (including fuzzy matching)**

| Trial | Matched | Unmatched | Total |
|---|---|---|---|
| Contextualised curriculum | 153 | 51 | **204** |
| Embedded maths | 42 | 26 | **68** |
| Family numeracy | 376 | 197 | **573** |
| FSQ Level 1 | 304 | 211 | **515** |
| Maths mastery | 175 | 319 | **494** |
| Preparation for maths GCSE | 690 | 368 | **1,058** |
| **Total** | **1,740** | **1,172** | **2,912** |

Source: Ipsos

**Table A30: Kicktag match results to child attainment data**

| Trial | Matched | Unmatched | Total |
|---|---|---|---|
| Contextualised curriculum | - | 204 | **204** |
| Embedded maths | - | 68 | **68** |
| Family numeracy | 425 | 148 | **573** |
| FSQ Level 1 | - | 515 | **515** |
| Maths mastery | - | 494 | **494** |
| Preparation for maths GCSE | - | 1,058 | **1,058** |
| **Total** | **425** | **2,487** | **2,912** |

Source: Ipsos

# Annex 14: Adult Numeracy RCT Costs

The total delivery, research and evaluation costs for the programme of Adult Numeracy Randomised Controlled Trials (RCTs) were £5,702,220[73].

The RCT costs included:

- the product development of the interventions to be tested (£330,290)
- the funding allocation to support delivery (£693,320)[74]. This included support for adult education providers and schools taking part in each RCT and to product developers to support testing and tutor training to deliver the trials.
- the managed service supplier (£1,293,710) and

the evaluation of the trials (£3,384,900)

[73] Figures rounded to the nearest £10
[74] Please see individual trial reports for individual trial grant funding spend.

Department for Education


**Reference**: RR1628/1
**ISBN:** 978-1-83870-791-0

For any enquiries regarding this publication, contact www.gov.uk/contact-dfe.

This document is available for download at www.gov.uk/government/publications.