Mathematical Reasoning

National Institute of Economic and Social Research
Lucy Stokes, Jake Anders, Richard Dorsett,
Anitha George, Nathan Hudson-Sharp and Heather Rolfe

Education Endowment Foundation

Evaluation Summary	
Age range	Year 2 (age 6-7)
Number of pupils	Around 7,200
Number of schools	160
Design	Cluster randomised trial
Primary Outcome	Progress Test in Maths (GL Assessment)

Intervention

The Mathematical Reasoning programme develops children's understanding of the logical principles underlying mathematics. Developed by Professor Terezinha Nunes and Professor Peter Bryant at the University of Oxford, the programme consists of 10 teaching units, delivered to pupils as part of their normal mathematics lessons, taking up approximately 12 lessons. Teachers receive training and are provided with lesson plans and materials to deliver the programme. Learning is also supported by online games, which can be used by pupils both at school and at home.

In this effectiveness trial, delivery will be via a scalable train the trainers model. Schools will be recruited and trained through the Maths Hubs, working with the National Centre for Excellence in the Teaching of Mathematics (NCETM). Within each participating hub, Work Group leaders will be identified and trained in delivering the intervention. This will include Work Group leaders using the programme materials in the schools in which they work. Work Group leaders will then train teachers within the participating schools that have been randomly allocated to receive the intervention. These teachers will receive one day of training; Work Group leaders will provide further support to the teachers through a school visit during the period in which they are delivering the programme. Teachers in the intervention group will also have access to online resources made available by the Oxford team.

Significance

Previous research by the developers of the Mathematical Reasoning programme (Nunes, Bryant, et al., 2007) has demonstrated the importance of logical reasoning for mathematical understanding. The Mathematical Reasoning programme focuses on quantitative reasoning, that is, the ability to understand the relationships between numbers and to use them to solve problems.

A previous EEF efficacy trial of the Mathematical Reasoning programme has shown the intervention to have a positive impact of 0.2 on pupils' mathematical attainment (Worth et al., 2015). Such efficacy trials are designed to test whether an intervention can work under ideal conditions, with considerable support from the original developer of the intervention. The effectiveness trial provides the opportunity to test whether the intervention can work at scale. As such it will be particularly important for the evaluation to explore those factors which have changed in order to implement the intervention on a larger scale, for example, whether the change in the delivery of training and method of support (greater distance from the original developers) has consequences for the effect of the programme.

Methods

Research questions

The primary research question this evaluation is designed to answer is:

- What is the effect of the Mathematical Reasoning programme on children's mathematical attainment (as measured by the Progress Test in Maths) at the end of Year 2?

The evaluation will also address the following questions:

- What is the effect of the Mathematical Reasoning programme on mathematical attainment at the end of Year 2, among pupils who are eligible for free school meals?

Design

This will be a cluster randomised controlled trial. Randomisation will take place at school level. Schools will be recruited to the trial through Maths Hubs, working with the NCETM. Eight hubs will be identified for the trial. Each hub will aim to recruit 20 schools, with schools randomly allocated to either the treatment arm (receiving the Mathematical Reasoning programme) or the control group. Thus there will be two arms of the trial. Schools in the control group will be expected to deliver 'business as usual' mathematics teaching and will be offered the opportunity to take part in the programme in the following school year (that is, for the pupils who will be in Year 2 in the following academic year).

Randomisation

Schools will be randomised within blocks defined on the basis of hubs, proportion of children eligible for FSM and prior attainment at KS1 (that is, school-level KS1 attainment in the academic year 2014/15). Within each of the 8 hubs, two FSM groups will be determined: 'high' and 'low' – with schools ranked within hubs by the proportion of pupils eligible for FSM, with thresholds for the 'high' and 'low' groups then chosen so that half of all schools fall into each group. Within each of these Hub/FSM groups, schools will then be allocated into two KS1 groups (again ranking schools on the basis of their KS1 attainment and allocating schools to 'high' and 'low' groups, so that half of schools fall into each category). With 8 hubs, two FSM groups and two KS1 groups, this will result in 32 blocks.

Randomisation of schools, to achieve a 50:50 allocation, will be performed as follows:

- Each school will be assigned a randomly generated number;
- Schools will be sorted by hub, block and random number
- The first school will be randomised to treatment or control;
- Each subsequent school will be assigned to have the opposite outcome of the previous school.

The randomisation process will be recorded in the syntax and log files used to carry out the randomisation.

Participants

Schools will be recruited through the Maths Hubs, working with the NCETM. Eight hubs will be identified to take part in the trial. All English state primary and infant schools within the recruited hubs will be eligible to participate in the trial. This includes schools with one class per year group as well as schools which have classes with mixed year groups. Each hub will be expected to recruit around 20 schools. Within each school, all Year 2 pupils and teachers will be eligible to participate.

Recruitment will be led by NCETM. Maths Hubs, working with the NCETM, will advertise the trial through various communication channels, such as newsletters, mailing lists, websites and twitter, using centrally produced communication materials. The Maths Hubs will select the 20 schools within each of their regions based on expressions of interest.

Schools wishing to participate in the trial will be asked to sign a Memorandum of Understanding, and must be willing to fully comply with the requirements of the trial, including supplying the necessary pupil data. Opt-out consent will be sought from parents of all eligible pupils for agreement for data sharing.

Outcome Measures

The primary outcome will be the GL Assessment Progress Test in Maths. Level 7 of the Progress Test in Maths will be used, as this is the recommended level for this age group. The paper version of the test will be used.

Tests will be administered by the National Centre for Social Research (NatCen). The test administrators will be blinded as to whether the school is within the treatment or control group.

Sample size calculations

The minimum detectable effect size (MDES) has been estimated based on the parameters for the efficacy trial, reported in Worth et al (2015). This reports an intra-cluster correlation, before controlling for covariates, of 0.12 and that 57 per cent of both school-level and individual-level variance is controlled for by covariates. With a proposed sample size of 160 schools, in 8 hubs, assuming 45 pupils per school, the MDES is estimated to be 0.11 (with 5 per cent significance level and 80 per cent power). The effect size in the efficacy trial was 0.2; there is likely to be some dilution as a result of the train the trainers model, but this sample size should therefore provide adequate power to detect the size of effect likely to arise.

Assuming that 15 per cent of the pupils in the sample are eligible for Free School Meals, and assuming that all other parameters remain the same, the MDES for this subgroup stands at 0.14.

Analysis plan

The primary outcome will be the overall standardised score on the Progress Test in Maths. This is the successor to the Progress in Maths test used in the efficacy trial (which will no longer be available by the time of testing for this trial).

The analysis will be carried out using linear regression, with robust standard errors to take account of the clustering at school level. The two types of school included in the trial are:
a) intervention schools
b) control schools.

The estimated impact will be based on the difference between these two groups, regardless of contamination of the control schools or drop out by intervention schools. This is in order to estimate the "intention to treat" (ITT) effect. In addition, the regression models used for the primary analysis will include controls for prior attainment (at school level).

Estimated impact in terms of pupil's mathematical achievement will be converted into a Hedges' g effect size (1981). This will use the estimated total pooled standard deviation of the treatment and control groups, rather than the within-school pooled standard deviation as this is a more conservative approach (although estimates will also be produced on this basis to check they do not differ much).

We will also conduct the analysis for the subgroup of pupils who have ever received free schools meals.

Implementation and process evaluation methods

The purpose of the process evaluation will be to establish fidelity and identify the factors influencing impact. The basic features of our design are:

- Attendance at the training for Work Group Leaders (WGLs) and at the teacher training days held in three hubs
- Analysis of resources and of data collected by the project team of relevance to the process evaluation
- Visits to 8 schools to interview teachers involved in the intervention and to observe sessions
- On-line survey of all schools

Evaluation of training: A key aim of the process evaluation will be to explore the consequences of the introduction of the "train the trainers" model. We will attend the training course held for WGLs as well as the teacher training days in three hubs to understand their initial experiences of the intervention, their expectations and any concerns. It is anticipated that the training days will include group discussion sessions and observations of these would be used to garner teachers' impressions of the training and materials. We will also review training content and resources to understand both

how WGLs are prepared to train other teachers and how all teachers are prepared to deliver the intervention. We anticipate that NCETM and the Maths Hubs will evaluate the training and provide us with extracted data. We will also review monitoring and support provided to the WGLs and by the WGLs to teachers. In addition, a small number of phone interviews will be conducted with WGLs to explore their experiences of training and supporting teachers in delivering the intervention.

Review of materials: We will review the training, project materials and resources developed by Oxford University and used by WGLs and Year 2 teachers. This will also be covered in interviews with teachers and in observation of sessions (see below). We will also review the content of online games, looking in particular at whether pupils across the range of maths and IT ability are able to use these unassisted. We would expect to work closely with NCETM to share qualitative and quantitative data relevant to the process evaluation.

Visits to schools: We will carry out research visits to 8 schools, with a full day of interviews and observations in each (timing visits to enable observation of lessons). We anticipate that these visits will take place towards the end of the intervention. Through visits to schools we will interview project leads and teachers on the intervention itself, and particularly their experiences of training, using the project resources and delivering the intervention. We will also carry out classroom observations.

Through interviews with teachers, we would wish to establish the schools' current approach to maths teaching, with the aim of ascertaining the extent to which the Mathematical Reasoning programme changed teachers' usual approaches to maths teaching. It will also be important to establish pupil attainment, schools' maths performance data, identified barriers to improvement and other strategies in place to improve maths performance. The focus of interviews will be on teachers' experiences of implementation, including views on the appropriateness and value of the resources, any adaptations they have needed to make and pupil response.

Through classroom observations we will assess pupil engagement and evidence relating to their understanding of the principles taught in the sessions. Where possible, we will observe pupils using on-line games.

Online survey of all intervention and control schools: The survey will enable us to gather qualitative data in a consistent way from all intervention schools on implementation and perceived outcomes and on factors that could affect fidelity. The survey will cover their experiences of using the lesson plans and guidance materials, including preparation and delivery time. We will also explore how the data collected on fidelity through the survey can be used in quantitative analysis to investigate any potential moderating of treatment effects.

The survey will also include teachers' and schools' usual approaches to teaching maths, in particular whether this includes:

- A focus on the principles underlying the Mathematical Reasoning programme
- Whole class teaching
- Use of online games, time allocated in school and efforts made to help pupils use these at home

Additionally, the survey will investigate teachers' knowledge development in terms of their understanding of mathematical reasoning as well as of how children learn mathematics. We will also ascertain changes in practice resulting from participating in the intervention.

We will also survey control schools about their usual approaches to teaching maths to Year 2. It will be important to keep the on-line survey short, requiring around 10-15 minutes to complete, and to focus on questions which contribute to broad understanding of implementation, feasibility and impact. The qualitative interviews will seek to understand the success or otherwise of the intervention in greater depth.

Interviews will be digitally recorded, with the agreement of teachers and transcribed. Lessons will be observed, using a pro-forma designed for the purpose. It will include assessment of pupil engagement as well as evidence of their understanding of the logical principles underlying maths. We will analyse the data using a social research 'framework' approach, drawing themes and messages from an analysis of interview transcripts, observations of training and of lessons and other materials collected by evaluation and project teams. The on-line survey will be delivered using Survey Monkey and analysed using NVivo software.

Costs

An estimate of the per-pupil cost of the intervention will be calculated by NIESR. This estimate will focus on cost from the perspective of a participating school and will be based on the marginal, financial costs of the intervention.

The cost estimates will in part make use of information from the project team (particularly with regard to the actual cost of delivering the intervention, e.g. the cost of providing the training), as well as that collected directly by NIESR. Information on costs, especially any hidden costs or resource implications, will be explored through the process evaluation as part of the interviews with teachers and school visits. The purpose of collecting such data in the process evaluation would be to identify areas of expenditure that the project entails and which would need to be considered for a wider roll out of the project. This process will also inform the development of appropriate questions on costs/resource use for inclusion in the survey. This will need to strike a balance between collecting sufficient cost information and not damaging response rates; it will also need to take account of whether a teacher is well placed to provide accurate information on particular types of costs.

Costs in terms of time will be reported separately from the financial costs – such as the amount of time for which schools need to arrange supply cover for teachers to attend training. Costs in terms of pre-requisites, for example, the availability of technology for whole class teaching and for enabling pupil access to online games, will also be considered.

An estimate of cost-per pupil per year will also be calculated based on a 3-year time period, as once trained, teachers would also be able to deliver the programme in subsequent years. Any costs associated purely with the trial will be excluded.

Ethics and registration

International Standard Randomised Controlled Trial Number: ISRCTN89670776

We take seriously the ethical issues raised in both the quantitative and qualitative elements of the research. NIESR adheres to the Ethics Guidelines of the Social Research Association (SRA). Members of the process evaluation team have Disclosure and Barring Service (DBS) clearance. Ethical review of the project is being undertaken by the University of Oxford. While NIESR has an ethics committee consisting of Trustees, we do not consider that this trial requires such additional clearance.

As the intervention will be delivered within school hours, consent from the school should be sufficient with regard to consent for the intervention; and as randomisation is taking place at the school level (rather than randomising individual pupils), the decision to enter into randomisation can also be made by the school. However, we will be collecting personal information on pupils as well as their performance in assessments. We will also be applying for, using and linking to data from the National Pupil Database. Whether participants are in the treatment or control groups will be identified from the school that they attend. Participants' confidentiality and anonymity will be safeguarded by the methods that we have in place. On this basis, we judge that it is appropriate for the trial to use an opt-out consent process, with participants' parents (or legal guardians) making an informed decision regarding whether they consent to data sharing based upon the information provided to them.

A parental information sheet provided with the opt-out form provides information on the aims of the research and the use of data in order that parents are able to make an informed decision about whether to withhold consent from data sharing. The form itself makes it clear and simple for parents to ensure their child(ren)'s data is not requested from the DfE and hence they will not be part of the trial.

Given the personal information collected in the course of this trial, data security is of utmost importance. As such, it will be transmitted and stored using the security principles underlined in the NIESR Data Security policy (attached in Appendix A) and the procedures set out in further detail for this specific project in a Data Sharing Agreement. This includes secure transfer of data and use of password-protection and encryption as appropriate during data storage.

Personnel

Project team (Oxford and NCETM):

- Terezinha Nunes (Oxford)
- Peter Bryant (Oxford)
- Rossana Barros Baertl (Oxford)
- Deborah Evans (Oxford)
- John Westwell (NCETM)
- Ione Crossley (NCETM)

Evaluation team (NIESR):

- Lucy Stokes (Principal Investigator)
- Anitha George (Qualitative Lead)
- Jake Anders

- Richard Dorsett
- Nathan Hudson-Sharp
- Heather Rolfe

Test administration (NatCen):

- Rakhee Patel

Design of the trial

- sample size calculations - NIESR
- refinement of randomisation approach - NIESR

Delivery of the intervention

- identification of hubs and recruitment of schools - NCETM
- delivery of training – NCETM/Oxford

Measurement of outcomes

- administration of Progress Test in Maths – NatCen
- NPD application– NIESR

Data collection

- collection of school and pupil data – NCETM/Maths Hubs/NIESR
- data for process evaluation – NCETM and NIESR
- data for cost evaluation – NCETM and NIESR

Impact analysis - NIESR
Qualitative analysis - NIESR

Risks

The data security policy of the National Institute of Economic and Social Research is included with this protocol.

Some of the key risks are listed below:

- School drop-out after randomisation reduces the integrity of the experimental design. To reduce the risk of drop-out, it will be important to ensure schools are well-informed about the programme and the trial from the start, so that they are clear as to what is expected of them before they commit to taking part. Schools will be asked to sign a memorandum of understanding as a signal of their commitment. It will also be important to maintain good communications with schools throughout the project in order to maximise retention. Drop-out of control schools is a particular risk; to help minimise this control schools will be offered the option to receive programme at later date.

- There may also be difficulties in recruiting schools to the trial. NIESR will work closely with NCETM and Oxford to convey the importance of the programme to schools and the value to them of them taking part.
- If individuals do not consent to data sharing this has the potential to reduce the sample size, and affect the internal and external validity of the trial. As consent is collected pre-randomisation, it should not affect the internal validity of the trial, as any withholding of consent should be just as prevalent in the treatment and control groups. In addition, as only opt-out consent is required, we judge that the risk of a large number of opt-outs is low.
- If pupils are not present on the day of testing this may also reduce the sample size by reducing the number of pupils for whom we are able to obtain a post-test; furthermore, it may introduce some bias if it is a non-random group of pupils who are absent. In those schools where a higher proportion of pupils are absent mop-up visits will be carried out to attempt to minimise this risk.
- There is a possibility that the delivery of the intervention will vary across schools. However, this reflects the reality of implementing such a programme; impact estimates therefore relate more to type of treatment likely to prevail in practice rather than that which might be observed under ideal conditions. Nevertheless understanding treatment variation is important and will be explored as part of the process evaluation.
- When randomising clusters rather than individuals the chances of a 'bad draw' increase because of the smaller number of units. We will use blocking to limit this problem

Timeline

Date	Activity
Oct-Dec 2015	Recruitment of hubs (NCETM) and further manualisation of the programme (Oxford)
Jan-Mar 2016	Recruitment of schools (Maths Hubs/NCETM) and training of WGLs (NCETM/Oxford)
Apr-Jun 2016	Collect consent and pupil data (NCETM/Maths Hubs/NIESR)
July 2016	Randomisation (NIESR)
Sept-Dec 2016	Training of teachers (NCETM/Maths Hubs)
Jan-Apr 2017	Delivery of programme (NCETM/Maths Hubs)
Jan 16-Jul 2017	Process evaluation ongoing throughout this period (NIESR)
Jun-Jul 2017	Post-tests administered (NatCen)
Sept-Dec 2017	Training of teachers from control schools (NCETM/Maths Hubs)
Jul-Dec 2017	Impact analysis ongoing throughout this period (NIESR)
January 2018	Evaluation report (NIESR)

References

Nunes, T., Bryant, P., Evans, D., Bell, D., Gardner, S., Gardner, A. and Carraher, J. (2007) 'The contribution of logical reasoning to the learning of mathematics in primary school', British Journal of Developmental Psychology, 25: 147–166. doi: 10.1348/026151006X153127

Worth, J., Sizmur, J., Ager, R. & Styles, B. (2015) "Improving numeracy and literacy" EEF evaluation report.

Appendix A: NIESR Data security policy

See attached document