Education Endowment Foundation

Effectiveness trial of Mathematical Reasoning

Further appendices

September 2026

Lillian Flemons, Chris Morton, Eleanor Bradley, Rachel Clutterbuck, and Helen Poet

NFER
National Foundation for Educational Research

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

We do this by:

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

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

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

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

0207 802 1653

info@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Table of contents

About the evaluator

The project was independently evaluated by a team from the National Foundation for Educational Research. The trial director and principal investigator for this study was Helen Poet, Research Director. Lillian Flemons, Research Manager, led the evaluation team and the implementation and process evaluation (IPE). Andrew Smith, Senior Evaluation Analyst, and Chris Morton, Senior Statistician, led the impact evaluation. They were supported by Eleanor Bradley and Gustavo Lopes as IPE researchers, and Rachel Clutterbuck and Juan Manuel Pozo Segura as statisticians. Kathryn Hurd led the research operations with Elaine McCann and Katharine Stoodley as operations researchers.

Contact details:

Helen Poet
National Foundation for Educational Research,
The Mere,
Upton Park,
Slough, Berkshire
SL1 2DQ,
United Kingdom

Email: h.poet@nfer.ac.uk

Acknowledgements

We are grateful to all of the primary schools that participated in this evaluation. We are especially thankful to the school staff members who completed the surveys, administered the baseline assessments, and co-ordinated the endpoint assessments with the National Foundation for Educational Research test administrators, and the staff and pupils who took part in the case study visits and interviews.

We would like to thank the delivery team, including Terezinha Nunes, Gabriel Stylianides, Louise Matthews, and Rossana Barros Baertl, for supporting the evaluation.

We would also like to thank the team at the Education Endowment Foundation, including Daniela Alvarado, Ben Sillitoe, and Amy Clark.

Office for National Statistics (ONS) Secure Research Service publication status and statistical results

Appendix E: Confirmatory Factor and Associated Analyses of the PTM7

Data Preparation and Method of Analysis

Confirmatory Factor Analyses (CFA) were performed on the PTM7 items to test the suitability of a four-subscale structure to the measure, as proposed by GL Assessment. The structure proposed by GL Assessment comprised the following four subscales: fluency in facts and procedures (5 items), fluency in conceptual understanding (9 items), problem-solving (4 items), and mathematical reasoning (16 items). Data of 879 pupils, who were missing baseline and/or endpoint PTM7 overall scores, were excluded from the analyses. This, therefore, gave a final sample size of N = 5,373 pupils for the confirmatory factor and associated analyses. Missing data on PTM7 items were imputed with zeros – this was deemed the most appropriate method given this is how missing response data is handled when calculating pupil scores on the PTM7, and, as there was no time limit for the assessment, it can be assumed that where there are missing item responses, the pupil was unable to answer the item.

Prior to CFA, correlations between items were evaluated to understand the strength and directionality of the relationships between the items – correlations that are close to zero and negative correlations may indicate weak or inverse relationships between the items. Cronbach's alpha, McDonald's omega, and composite reliability statistics were also calculated to estimate the reliability of the scale overall and the individual subscales. Alpha-if-item-removed statistics were calculated for subscales with lower reliability estimates to assess whether any single item may be lowering the reliability of the scale.

The PTM7 items have categorical indicators, with most items being binary alongside a small number of three- and four-category items (27 binary items, five items with three categories, two items with four categories). As a formal check to test for multivariate non-normality, Mardia's test (Mardia, 1970) was applied to each of the four subscales and all the items together. The results of the Mardia's tests provided clear evidence against multivariate normality, with each subscale and all items together deviating from normality on the skewness and/or kurtosis components of the Mardia's test. Therefore, in line with best practice for situations of multivariate non-normality, and where items have categorical indicators, a diagonally weighted least squares estimator (WLSMV) was used in the CFAs (Li, 2016).

The aim of the CFAs was to determine whether a four-factor structure of the PTM7, as proposed by GL Assessment, is a better fit to the data than a one-factor structure. To test this statistically, a four-factor CFA and a one-factor CFA were both performed on the data. The fit statistics of both CFA models were evaluated using the commonly referenced criteria by Hu & Bentler (1999); Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) close to 0.95, Root Mean Square Error of Approximation (RMSEA) close to 0.06 and Standardised Root Mean Square Residual (SRMR) close to 0.08. Standardised factor loadings were also evaluated alongside fit indices to understand individual item fit; factor loadings greater than 0.3 were considered sufficient (Hair *et al.*, 1998). Owing to the ordinal estimation method used in this analysis, fit indices were interpreted jointly rather than by single cutoffs (e.g., Marsh, Hau and Wen, 2004). Due to the sensitivity of the chi-squared statistic to large samples(Barrett, 2007), fit indices and factor loadings were used to determine model fit. The two CFA models were also compared statistically using a likelihood ratio test.

Analyses

Correlations

Correlations between all PTM7 items are presented in Appendix Table 5. There were no negative correlations between items, and correlations ranged from $r = 0.064$ to $r = 0.518$. Although several of the inter-item correlations were close to zero, the positive directionality of the correlations supported the suitability of the data for factor analysis. It was expected that there would be stronger correlations between items within the same subscale, supporting the proposed four-subscale structure; however, this was not observed.

Appendix E Table 1: Correlation coefficients between PTM7 items

	Q1a	Q2a	Q3a	Q4a	Q4b	Q1b	Q1c	Q2b	Q2d	Q3b	Q6a	Q8a	Q9a	Q9b	Q1d	Q3d
Q1a																
Q2a	0.28															
Q3a	0.18	0.20														
Q4a	0.15	0.17	0.14													
Q4b	0.15	0.15	0.14	0.17												
Q1b	0.37	0.26	0.22	0.15	0.18											
Q1c	0.32	0.27	0.21	0.16	0.16	0.37										
Q2b	0.26	0.31	0.16	0.13	0.13	0.26	0.27									
Q2d	0.28	0.37	0.19	0.17	0.16	0.32	0.30	0.36								
Q3b	0.17	0.17	0.39	0.13	0.10	0.23	0.19	0.14	0.17							
Q6a	0.24	0.24	0.17	0.16	0.16	0.26	0.23	0.22	0.25	0.16						
Q8a	0.25	0.29	0.17	0.14	0.10	0.25	0.25	0.26	0.29	0.14	0.22					
Q9a	0.28	0.30	0.21	0.16	0.16	0.30	0.26	0.27	0.31	0.21	0.27	0.27				
Q9b	0.24	0.25	0.18	0.15	0.15	0.28	0.25	0.22	0.26	0.20	0.24	0.23	0.35			
Q1d	0.17	0.15	0.11	0.06	0.08	0.19	0.16	0.16	0.18	0.15	0.16	0.13	0.25	0.18		
Q3d	0.21	0.19	0.32	0.15	0.15	0.25	0.24	0.19	0.24	0.26	0.17	0.18	0.24	0.23	0.12	
Q5c	0.15	0.12	0.11	0.09	0.10	0.15	0.13	0.12	0.14	0.13	0.13	0.15	0.18	0.15	0.16	0.09
Q10c	0.15	0.14	0.10	0.08	0.10	0.16	0.13	0.13	0.17	0.12	0.15	0.12	0.22	0.18	0.24	0.10
Q2c	0.18	0.19	0.15	0.09	0.12	0.19	0.20	0.27	0.25	0.13	0.18	0.17	0.24	0.16	0.20	0.14
Q2e	0.22	0.23	0.17	0.10	0.15	0.24	0.22	0.29	0.38	0.14	0.19	0.19	0.27	0.20	0.22	0.17
Q3c	0.26	0.30	0.33	0.16	0.16	0.30	0.24	0.25	0.31	0.31	0.28	0.25	0.31	0.29	0.20	0.33
Q3e	0.20	0.15	0.18	0.20	0.13	0.20	0.21	0.12	0.14	0.22	0.15	0.14	0.21	0.21	0.10	0.17
Q4c	0.22	0.19	0.18	0.15	0.18	0.21	0.23	0.23	0.22	0.18	0.22	0.16	0.28	0.21	0.18	0.21
Q5a	0.32	0.38	0.21	0.20	0.16	0.36	0.29	0.29	0.36	0.23	0.31	0.34	0.35	0.34	0.19	0.26
Q5b	0.24	0.27	0.19	0.14	0.11	0.27	0.24	0.26	0.32	0.21	0.24	0.25	0.32	0.25	0.25	0.20
Q6b	0.30	0.35	0.21	0.19	0.17	0.32	0.32	0.31	0.39	0.21	0.32	0.34	0.37	0.32	0.20	0.27
Q7a	0.25	0.37	0.15	0.13	0.11	0.24	0.23	0.27	0.32	0.14	0.24	0.35	0.28	0.25	0.12	0.17

Q7b	0.25	0.39	0.13	0.14	0.12	0.25	0.23	0.26	0.29	0.12	0.25	0.34	0.28	0.27	0.11	0.18
Q7c	0.33	0.42	0.20	0.21	0.16	0.33	0.32	0.34	0.41	0.21	0.32	0.38	0.39	0.35	0.20	0.25
Q7d	0.21	0.19	0.13	0.08	0.10	0.19	0.17	0.20	0.25	0.13	0.19	0.19	0.27	0.21	0.26	0.14
Q8b	0.16	0.13	0.10	0.06	0.10	0.14	0.13	0.16	0.17	0.11	0.14	0.11	0.22	0.15	0.26	0.14
Q8c	0.22	0.28	0.11	0.13	0.12	0.21	0.21	0.21	0.24	0.10	0.19	0.33	0.26	0.22	0.12	0.15
Q10a	0.19	0.17	0.15	0.09	0.09	0.19	0.19	0.16	0.21	0.14	0.18	0.17	0.24	0.20	0.16	0.14
Q10b	0.21	0.19	0.16	0.12	0.11	0.18	0.18	0.18	0.21	0.17	0.19	0.18	0.24	0.21	0.18	0.17

Appendix E Table 1 continued

	Q5c	Q10c	Q2c	Q2e	Q3c	Q3e	Q4c	Q5a	Q5b	Q6b	Q7a	Q7b	Q7c	Q7d	Q8b	Q8c	Q10a
Q5c																	
Q10c	0.13																
Q2c	0.12	0.19															
Q2e	0.17	0.19	0.31														
Q3c	0.13	0.20	0.21	0.26													
Q3e	0.23	0.09	0.09	0.11	0.19												
Q4c	0.13	0.16	0.23	0.23	0.24	0.15											
Q5a	0.17	0.19	0.20	0.23	0.44	0.23	0.24										
Q5b	0.22	0.24	0.26	0.28	0.36	0.15	0.24	0.44									
Q6b	0.17	0.18	0.24	0.28	0.37	0.21	0.27	0.47	0.37								
Q7a	0.12	0.12	0.15	0.20	0.27	0.14	0.18	0.39	0.25	0.35							
Q7b	0.12	0.13	0.14	0.19	0.27	0.14	0.16	0.39	0.24	0.36	0.51						
Q7c	0.19	0.20	0.24	0.27	0.41	0.20	0.23	0.52	0.41	0.49	0.47	0.48					
Q7d	0.16	0.24	0.24	0.27	0.23	0.08	0.17	0.22	0.32	0.26	0.19	0.19	0.30				
Q8b	0.16	0.19	0.24	0.22	0.19	0.09	0.17	0.16	0.23	0.20	0.10	0.11	0.20	0.26			
Q8c	0.12	0.13	0.15	0.19	0.24	0.14	0.17	0.29	0.22	0.28	0.29	0.31	0.34	0.15	0.12		
Q10a	0.13	0.29	0.18	0.19	0.23	0.15	0.17	0.22	0.23	0.22	0.18	0.17	0.25	0.21	0.15	0.18	
Q10b	0.15	0.23	0.16	0.18	0.23	0.16	0.16	0.23	0.23	0.26	0.20	0.20	0.27	0.20	0.14	0.18	0.45

Numbers shown are Pearson's correlation coefficients (*r*).

Items within the same subscale have cells of the same colour. Correlations between two items from the same subscale have cells shaded with a lighter version of the corresponding colour.

Subscales: *fluency in facts and procedures* (Q1a, Q2a, Q3a, Q4a, Q4b), *fluency in conceptual understanding* (Q1b, Q1c, Q2b, Q2d, Q3b, Q6a, Q8a, Q9a, Q9b), *problem-solving* (Q1d, Q3d, Q5c, Q10c), and *mathematical reasoning* (Q2c, Q2e, Q3c, Q3e, Q4c, Q5a, Q5b, Q6b, Q7a, Q7b, Q7c, Q7d, Q8b, Q8c, Q10a, 10b).

Confirmatory Factor Analyses (CFA)

Four-factor CFA model. Initially, a four-factor CFA model was run on the PTM7 items, testing the four-subscale structure proposed by GL Assessment (*fluency in facts and procedures*, *fluency in conceptual understanding*, *problem-solving*, and *mathematical reasoning*). The fit indices of the four-factor CFA model showed that a four-factor structure was a good fit to the data ($\chi^2(521) = 6646.232$ (scaled), $p < .001$; CFI = 0.937, TLI = 0.933, RMSEA = 0.047 (90% CI [0.046, 0.048]), SRMR = 0.059); CFI and TLI were close to .95, RMSEA was below 0.06, and SRMR was below 0.08, suggesting acceptable model fit (Hu & Bentler, 1999; Marsh, Hau, & Wen, 2004). As would be expected in large samples, the chi-squared statistic was significant. Supporting the acceptable-fit of the model as indicated by the fit indices, factor loadings for each item were sufficient on each factor (fluency in facts and procedures; 0.365 – 0.764, fluency in conceptual understanding; 0.484 – 0.733, problem-solving; 0.376 – 0.642, mathematical reasoning; 0.427 – 0.846); factor loadings are reported in full in Appendix Table 6. However, latent correlations between the four factors were high, ranging between 0.942 and 1.023 (see Appendix Table 7), suggesting high overlap and potentially poor discrimination between the subscales. A one-factor CFA model was performed on the data, to understand whether a one-factor model was a better fit to the data than a four-factor model.

Appendix E Table 2: Standardised factor loadings from the four-factor CFA model of the PTM7

Factor	Item	Loading
Fluency in facts and procedures	Q1a	0.581
	Q2a	0.764
	Q3a	0.550
	Q4a	0.384
	Q4b	0.365
Fluency in conceptual understanding	Q1b	0.674
	Q1c	0.648
	Q2b	0.621
	Q2d	0.733
	Q3b	0.484
	Q6a	0.522
	Q8a	0.650
	Q9a	0.646
	Q9b	0.627

Problem-solving	Q1d	0.642
	Q3d	0.515
	Q5c	0.376
	Q10c	0.552
Mathematical reasoning	Q2c	0.557
	Q2e	0.598
	Q3c	0.729
	Q3e	0.427
	Q4c	0.474
	Q5a	0.846
	Q5b	0.731
	Q6b	0.796
	Q7a	0.768
	Q7b	0.795
	Q7c	0.798
	Q7d	0.575
	Q8b	0.629
	Q8c	0.569
	Q10a	0.527
	Q10b	0.547

Note. Standardised loadings are presented, all $p < .001$ where the null hypothesis is $\lambda = 0$. WLSMV estimator was applied.

Appendix E Table 3: Latent correlations between the four PTM7 factors

	Fluency in facts and procedures	*Fluency in conceptual understanding*	*Problem-solving*	*Mathematical reasoning*
Fluency in facts and procedures	--	1.022	1.023	0.942
Fluency in conceptual understanding	1.022	--	1.013	0.957
Problem-solving	1.023	1.013	--	1.005
Mathematical reasoning	0.942	0.957	1.005	--

Note. Inter-factor correlations, all $p < .001$.

One-factor CFA model. The fit indices of the one-factor CFA on the PTM7 items showed that a one-factor structure is also a good fit to the data ($\chi^2(527) = 6790.435$ (scaled), $p < .001$; CFI = 0.936, TLI = 0.932, RMSEA = 0.047 (90% CI [0.046, 0.047]), SRMR = 0.059), with fit indices being similar to the four-factor CFA model; CFI and TLI were close to .95, RMSEA was below 0.06, and SRMR was below 0.08. The chi-squared statistic was again significant, but this is to be expected with large samples. Factor loadings were sufficient for all items, ranging between 0.360 and 0.839 (see Table 4).

Statistical comparison of the four-factor and one-factor CFA models was performed on the data, using a likelihood ratio test for comparing nested CFA models (Satorra, 2000). The analysis revealed that the four-factor CFA model of the PTM7 items was a significantly better fit to the data than the one-factor CFA model ($\chi^2_{\text{difference}}(6) = 157.51$, $p < .001$). These results provide support for the use of a four-subscale structure of the PTM7.

Appendix E Table 4: Standardised factor loadings from the one-factor CFA model of the PTM7

Item	Loading
Q1a	0.572
Q2a	0.752
Q3a	0.541
Q4a	0.378
Q4b	0.360
Q1b	0.665
Q1c	0.640
Q2b	0.614
Q2d	0.725
Q3b	0.478
Q6a	0.516
Q8a	0.643
Q9a	0.639

Q9b	0.620
Q1d	0.653
Q3d	0.524
Q5c	0.383
Q10c	0.562
Q2c	0.552
Q2e	0.592
Q3c	0.723
Q3e	0.424
Q4c	0.470
Q5a	0.839
Q5b	0.724
Q6b	0.789
Q7a	0.761
Q7b	0.788
Q7c	0.791
Q7d	0.569
Q8b	0.623
Q8c	0.564
Q10a	0.521
Q10b	0.541

Note. Standardised loadings are presented; all $p < .001$ where the null hypothesis is $\lambda = 0$. WLSMV estimator was applied.

Reliability

Reliability estimates were calculated for the four subscales and the overall PTM7. The widely used Cronbach's alpha statistic is reported alongside McDonald's omega and composite reliability estimates. Reliability estimates were interpreted using the classic guidelines outlined by Nunnally (1978), suggesting an acceptable lower bound of $\alpha = 0.70$. The reliability estimates of the overall scale were excellent, in line with previous research (GL Assessment, 2015) . Reliability estimates of the four subscales proposed by GL Assessment were mixed, with *fluency in conceptual understanding* and *mathematical reasoning* showing acceptable-to-good reliability coefficients, but *fluency in facts and procedures* and *problem-solving* showing poor reliability estimates (see Table E). Average variance extracted (AVE) estimates, which indicate the amount of variance captured by the construct relative to the amount of variance due to measurement error, are below the suggested lower bound threshold of 0.5 on all subscales and for the overall scale, potentially indicating weak construct validity (Fornell & Larcker, 1981).

Appendix E Table 5: Reliability estimates and average variance extracted (AVE) of the PTM7

	Cronbach's α	Omega	Composite reliability	AVE
Fluency in facts and procedures	0.51	0.54	0.67	0.30
Fluency in conceptual understanding	0.75	0.77	0.85	0.39
Problem-solving	0.39	0.42	0.62	0.31
Mathematical reasoning	0.84	0.86	0.92	0.44
Overall scale (all PTM7 items)	0.90	0.91	0.95	0.38

Note. AVE = average variance extracted. Standardised Cronbach's alpha and McDonald's omega total coefficients are reported. Composite reliability and AVE were calculated using Fornell & Larker's (1981) method.

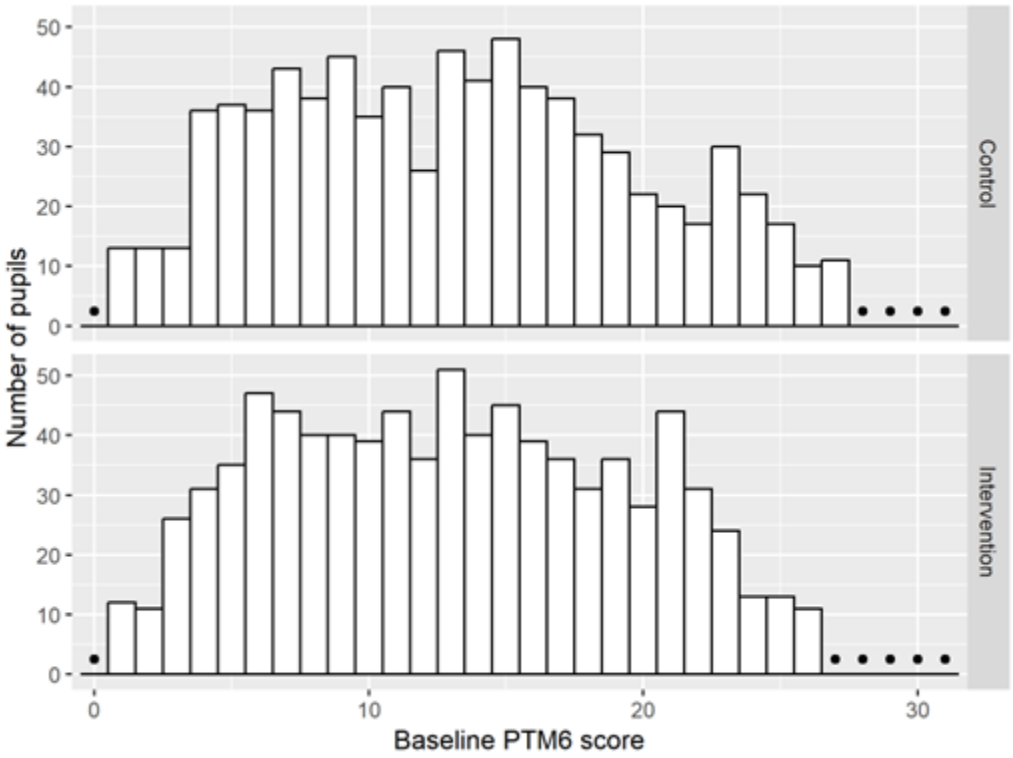
Alpha-if-item-removed statistics were calculated for the two subscales which had lower reliability estimates (*fluency in facts and procedures*, and *problem-solving*) to understand whether any single item could be lowering the reliability of these subscales. However, the alpha-if-item-removed statistics showed that removal of any item on these subscales would not improve the overall reliability of the subscale; fluency in facts and procedures alpha-if-item-removed ranged from α = 0.44 (Q1a) to α = 0.48 (Q4b), problem-solving alpha-if-item-removed ranged from α = 0.26 (Q1d) to α = 0.39 (Q3d). Therefore, no items were removed from the subscales.

Conclusion

Overall, the confirmatory factor analyses support the use of the four-subscale structure of the PTM7, as proposed by GL Assessment. The results of the factor analyses suggested that both the four- and the one-factor structure of the PTM7 are a good fit to the data, but the follow-up likelihood ratio test indicated that the four-factor structure is a better fit to the data than the one-factor structure. Correlational and reliability analyses, however, raise concerns about the four-factor structure. Correlations between items were generally low and we did not observe clustering of stronger correlations within the four subscales, as might be expected. In addition, reliability estimates were poor for two out of the four subscales, and AVE estimates were low for all subscales and the overall scale. While it is possible that the low reliability estimates may be due to there being fewer items on those poorly performing subscales, the vast number of low correlation coefficients within subscales is arguably concerning. Future research on the psychometric properties of the PTM7 should consider using exploratory methods to better understand the optimal structure of the PTM7.

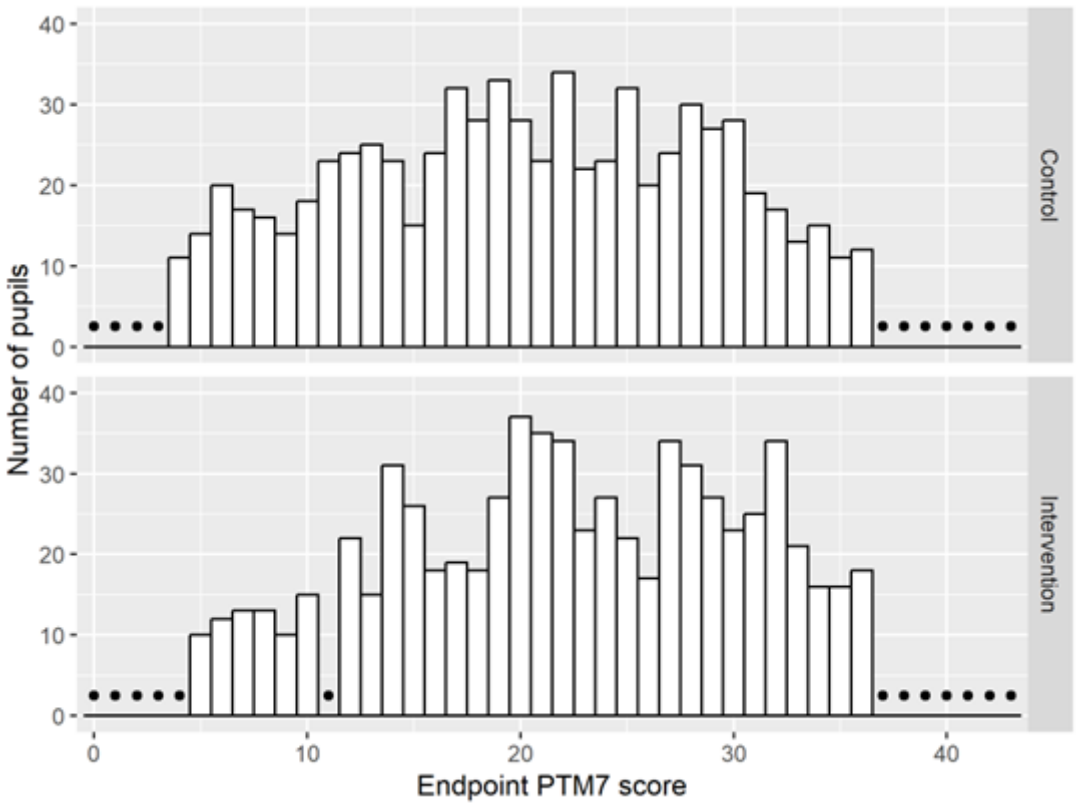
Appendix F: Distribution of PTM6 and PTM7 scores

Appendix F Figure 1: Distribution of PTM6 scores among FSM-eligible pupils in the control and intervention groups



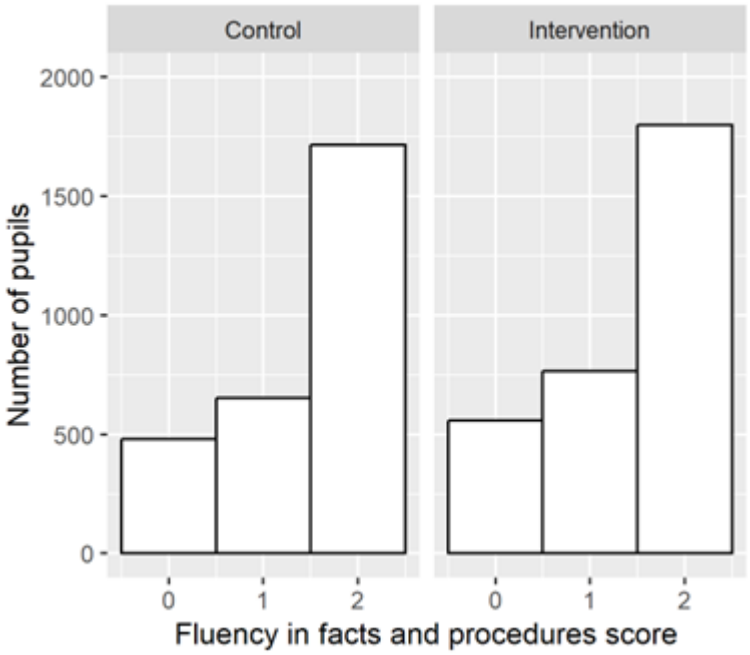
Black dots on the histogram represent pupil counts below ten. These have been suppressed to ensure individual pupils cannot be identified from the figures in this report.

Appendix F Figure 2: Distribution of PTM7 scores among FSM-eligible pupils in the control and intervention groups

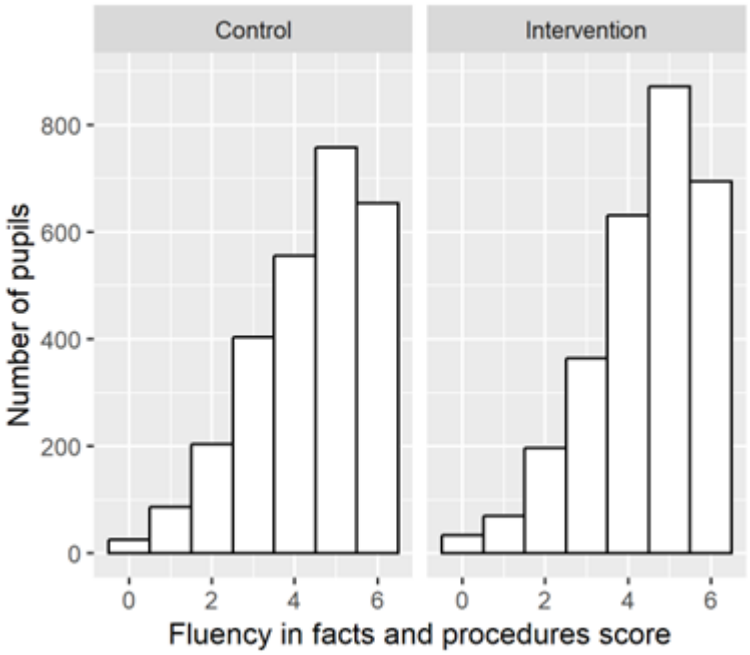


Black dots on the histogram represent pupil counts below ten. These have been suppressed to ensure individual pupils cannot be identified from the figures in this report.

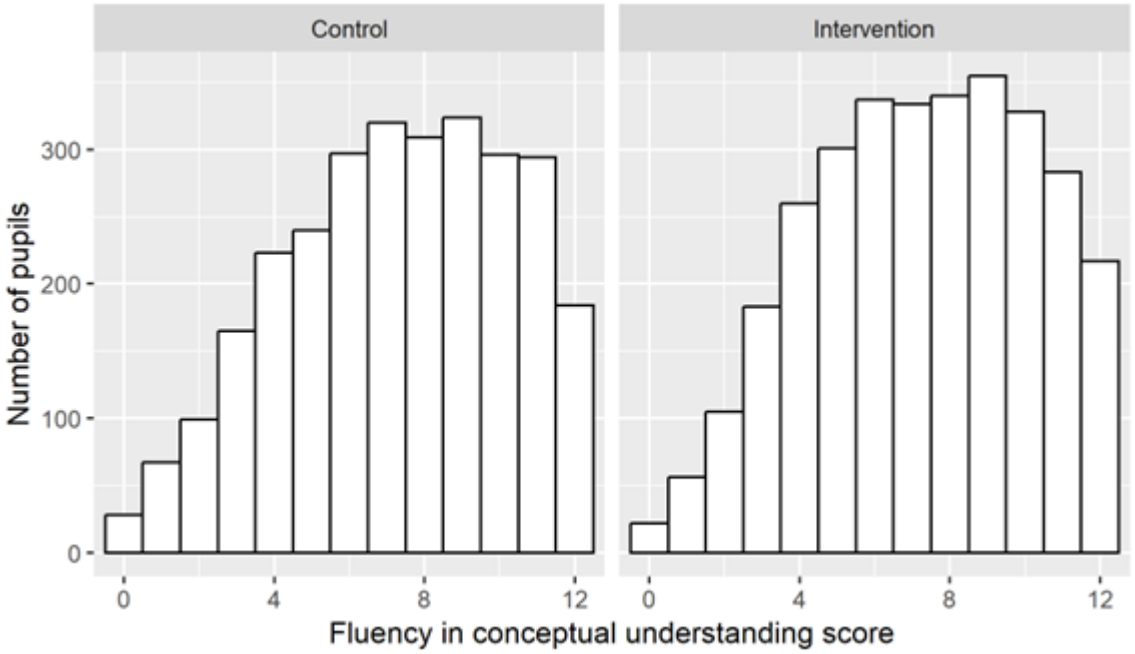
Appendix F Figure 3: Distribution of scores from the 'Fluency in facts and procedures' subscale of the PTM6, in the control and intervention groups



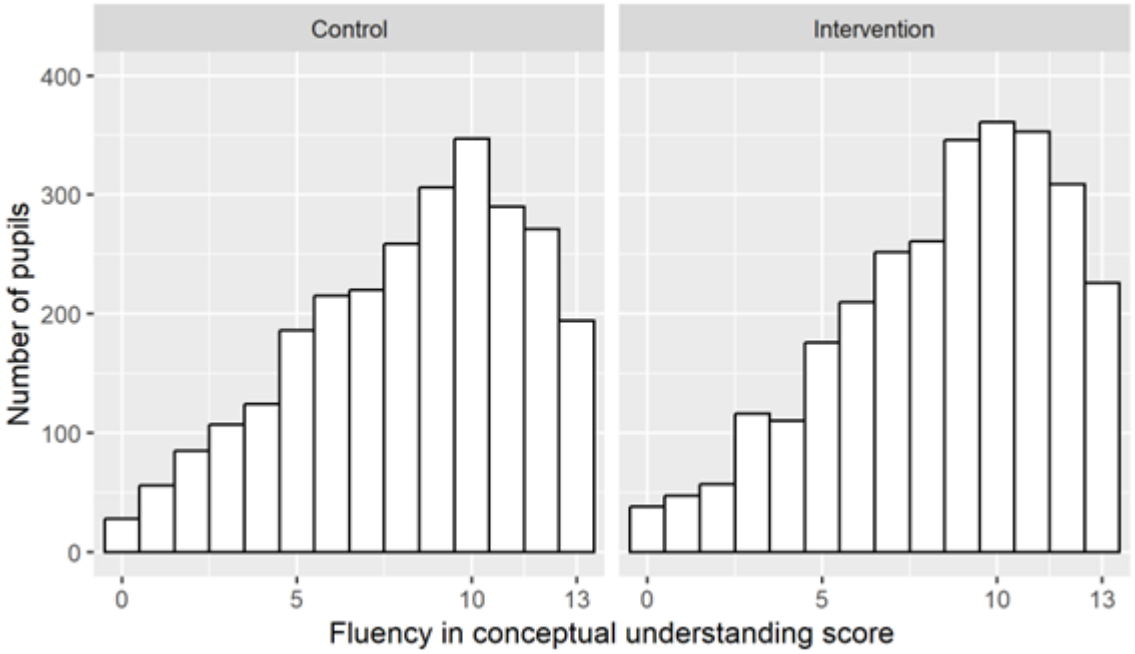
Appendix F Figure 4: Distribution of scores from the 'Fluency in facts and procedures' subscale of the PTM7, in the control and intervention groups



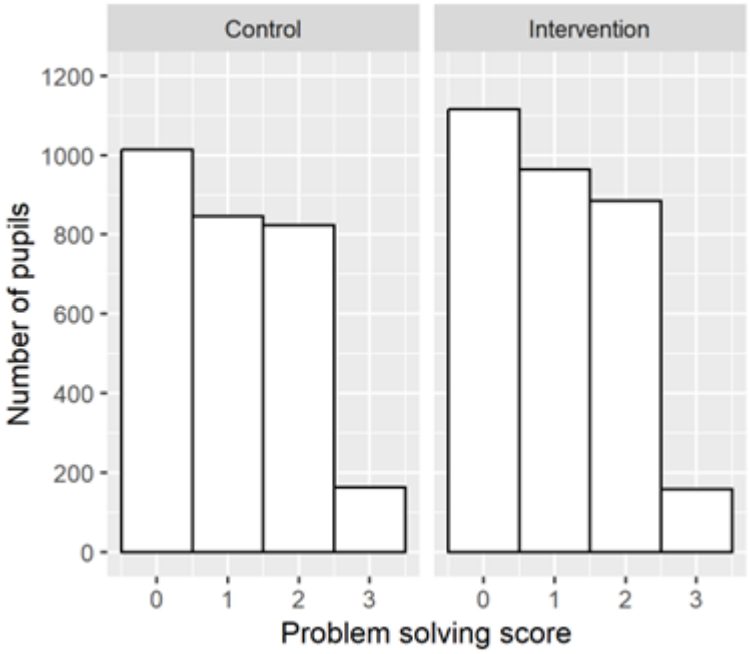
Appendix F Figure 5: Distribution of scores from the 'Fluency in conceptual understanding' subscale of the PTM6, in the control and intervention groups



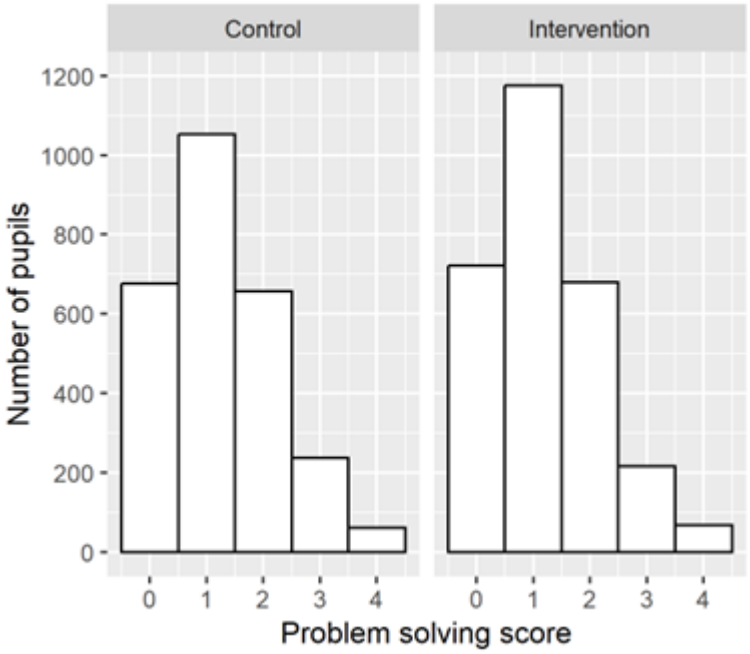
Appendix F Figure 6: Distribution of scores from the 'Fluency in conceptual understanding' subscale of the PTM7, in the control and intervention groups



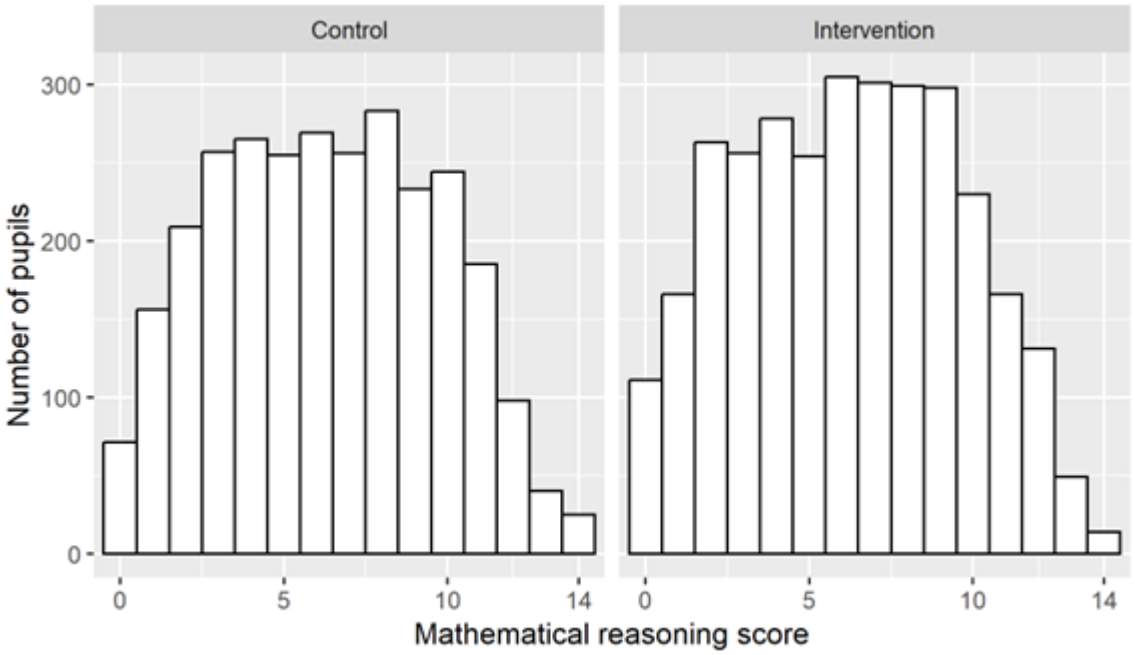
Appendix F Figure 7: Distribution of scores from the 'Problem solving' subscale of the PTM6, in the control and intervention groups



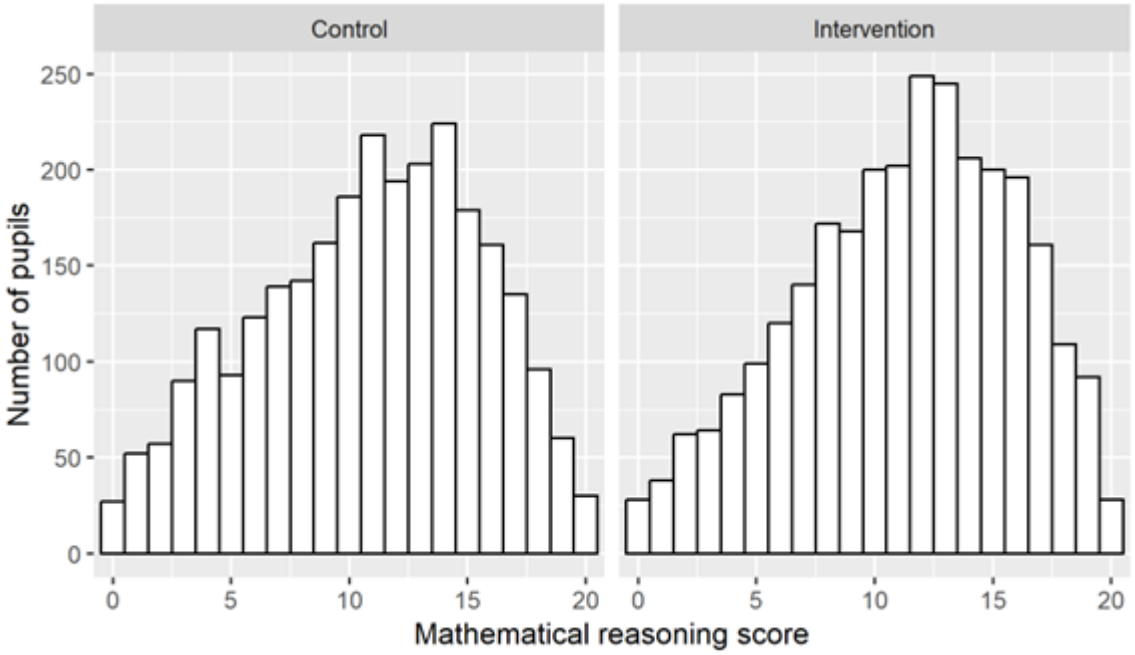
Appendix F Figure 8: Distribution of scores from the 'Problem solving' subscale of the PTM7, in the control and intervention groups



Appendix F Figure 9: Distribution of scores from the 'Mathematical reasoning' subscale of the PTM6, in the control and intervention groups



Appendix F Figure 10: Distribution of scores from the 'Mathematical reasoning' subscale of the PTM7, in the control and intervention groups



Appendix G: Missing data analysis

Appendix G Table 1: output from two multilevel logistic regression models. In the first model, the outcome is whether baseline PTM6 scores are missing (1) or not (0); in the second model, the outcome is whether endpoint PTM7 scores are missing (1) or not (0).

Outcome	Predictor	Estimate (SE)	Odds ratio (95% CI)	p-value
Missingness of PTM6 score	Intervention group	-0.327 (0.210)	0.72 (0.48, 1.09)	0.120
	PTM7 raw score	-0.027 (0.010)	0.97 (0.95, 0.99)	0.006
	Eligible for FSM	-0.080 (0.195)	0.92 (0.63, 1.35)	0.680
	Has SEN	0.494 (0.198)	1.64 (1.11, 2.42)	0.013
	% of half days absent in autumn and spring terms	0.131 (0.013)	1.14 (1.11, 1.17)	<0.001
	Average pupils per teacher at school	0.026 (0.032)	1.03 (0.96, 1.09)	0.410
	% of pupils eligible for FSM at school	0.003 (0.007)	1.00 (0.99, 1.02)	0.690
Missingness of PTM7 score	Intervention group	0.178 (0.125)	1.2 (0.94, 1.53)	0.153
	PTM6 raw score	-0.027 (0.007)	0.97 (0.96, 0.99)	<0.001
	Eligible for FSM	0.186 (0.109)	1.20 (0.97, 1.49)	0.089
	Has SEN	0.545 (0.114)	1.72 (1.38, 2.15)	0.000
	% of half days absent in autumn and spring terms	0.110 (0.008)	1.12 (1.10, 1.13)	<0.001
	Average pupils per teacher at school	-0.002 (0.018)	1.00 (0.96, 1.03)	0.907
	% of pupils eligible for FSM at school	0.003 (0.004)	1.00 (0.99, 1.01)	0.401

Appendix H: Computer games sensitivity analysis

Appendix H Table 1: sensitivity check in which the computer game models reported in Table X are altered to also include the N=11 pupils with over 150 games played

Group	Model	Predictor	Estimate (95% CI)	p-value
Intervention pupils	11	Number of games played	0.013 (0.004, 0.022)	0.003
	12	Number of subject topics	0.430 (0.227, 0.632)	<0.001
	13	Number of games played	0.029 (-0.003, 0.062)	0.080
		Number of subject topics	0.455 (0.213, 0.695)	<0.001
		Number of games played × Number of subject topics	-0.004 (-0.010, 0.002)	0.208

Appendix H Table 2: sensitivity check in which the computer game models reported in Table X are altered to exclude pupils with an amber quality match, leaving only those with a green match

Group	Model	Predictor	Estimate (95% CI)	p-value
Intervention pupils	14	Number of games played	0.056 (0.027, 0.084)	<0.001
	15	Number of subject topics	0.578 (0.237, 0.918)	<0.001
	16	Number of games played	0.114 (0.040, 0.188)	0.003
		Number of subject topics	0.788 (0.242, 1.334)	0.005
		Number of games played × Number of subject topics	-0.017 (-0.033, -0.001)	0.039

Appendix I: Training and preparation for Teacher Leaders

The training of TLs blends in-person and online activities. All TLs receive three days of in-person training with the Oxford University team. The first in-person training session provides an introduction to the programme and the online training course and encourages reflection and discussion around their own delivery of the programme sessions for practice (see below). It also includes discussion of themes that TLs will need to address when running webinars, such as why teach mathematical reasoning, and watching videos in order to identify observed "keys to success". These videos are part of the CPD for teachers in Module 5. The second training session prepares TLs for delivering a launch event and the three webinars, as well as how to support teachers and TAs with the training process. The third session focuses on feedback and reflection. New TLs also attend webinars delivered by the Oxford University team. These are different from the webinars the TLs themselves deliver to the teachers and TAs as part of the programme and are designed to support the Teacher Leaders with their own teaching of the programme.

All TLs are required to complete an online course tailored to their role. The course introduces TLs to the concept of mathematical reasoning, the expectations of the TL role, how to support teachers during their training (including through promoting fidelity) and how to deliver each of the three webinars. Participants are asked to respond to regular interactive activities to encourage them to reflect on what they have learnt and how they will apply this in practice. The course comprises training videos, research briefs, an online forum, interactive elements, and drafts of emails and other documents that TLs will be required to send to schools. TLs receive additional training and support material and templates for the programme as part of the in-person training session, as well as a slide deck for each of the webinars they deliver to the teachers and TAs as part of the programme. They are also provided with responses to FAQs from schools.

All TLs are expected to complete the online training course for schools and practise delivering eight sessions of the programme in non-trial schools. They will also shadow delivery of the school support webinars (see below) by an experienced TL before delivering their own.[1]

All training and preparation activities except the third day of in-person training and the webinar shadowing must be completed prior to in-school implementation of the programme.

Appendix J: Memorandum of Understanding

Effectiveness Trial of the Mathematical Reasoning Programme 2024-2025

Memorandum of Understanding

This document sets out the roles and responsibilities relating to the effectiveness trial of the Mathematical Reasoning (MR) programme for participating schools, the University of Oxford and the National Foundation for Educational Research (NFER).

Please read all the information below, along with the School Information Letter and privacy notices before signing this Memorandum of Understanding (MoU) and completing the information on page 4. Please return the completed MoU to Louise Matthews (**reasoningfirst@education.ox.ac.uk**) at your earliest convenience before 01.07.24

Responsibilities

The following outlines our expectations of schools and staff taking part in this evaluation. Please read the following statements and sign below to confirm you have read the document and agree to these terms. Please also sign and keep a copy for your reference. Please return the signed MoU as soon as possible to confirm your involvement in the trial.

The school will:

- Share the contact details for the Year 2 teacher for each participating class so that NFER and the University of Oxford can liaise with them as needed.
- Share the names, date of birth, UPN, FSM status and gender for all children in the participating class(es) with NFER via NFER's secure portal when requested.
- Be a point of contact for parents/carers, including providing them with full information about the project. Distribute an information sheet and a withdrawal form (provided by NFER) to all parents.
- Administer the Progress Test in Maths 6 (PTM6) to all children in participating classes in September 2024 (using booklets supplied by NFER) and return them to NFER using NFER's courier.
- Facilitate administration of the Progress Test in Maths 7 (PTM7) to all students by NFER Test Administrators in June 2025.
- Inform any NFER staff members visiting the school of the school's safeguarding processes and provide details of which school staff member to contact should a disclosure be made to them, or a safeguarding concern arise, during test administration or research visits.
- Ensure that NFER researchers and test administrators are never left unsupervised with children. A member of school staff should always be present to supervise activities.
- Support and encourage staff to complete staff surveys in September 2024 and June 2025 when requested by NFER to support understanding of the processes and implementation of the programme.

If allocated to the **Comparison Group**, the school will:

- Continue with usual practice over the trial period.
- Facilitate the data collection by NFER as outlined earlier.

If allocated to the **Intervention Group**, the school will:

- Release the teacher and TA for each participating class for 3 days in total to complete the core modules of the online training course (2 days) and to attend each of the three training webinars (total of 1 day) delivered over the course of the intervention period.
- Share contact details for the TA in each participating class so that NFER can contact them for data collection purposes.
- Deliver all units of the MR programme across a 12 to 15-week period between December 2024 and March 2025 in accordance with the programme training, guidance, and handbook.
- Deliver the whole-class component of each session using the PowerPoint provided and provide computers or tablets for children to use during the sessions when working on the computer games.
- Complete the session delivery and attendance log for each of the programme sessions and share this with NFER at the end of the 15-week delivery period.
- If invited by the NFER, make every reasonable effort to accept the invitation to participate in additional data collection activities for the purpose of case studies (interviews, session delivery observations, pupil focus groups).

Oxford University will:

- Provide a Teacher Leader who will support schools during the online teacher professional development course and programme delivery in Year 2, including via three webinars.
- Provide Intervention Group schools with:
  - Access to the online teacher professional development resources.
  - The MR programme, including printed Teacher Handbook and child workbooks, PowerPoints, access to online games for pupils to use and other printed resources required to deliver the programme.
- Provide Comparison Group schools with:
  - A £500 thank-you payment for schools allocated to the comparison group upon completion of the evaluation requirements.
- Share a brief report with schools at the end of the project.

NFER will:

- Provide a key contact who will be available to support schools throughout the evaluation period.
- Provide a parent information letter and withdrawal form to all participating schools.
- Provide schools with a secure portal to share pupil data and ensure that all data is shared in a secure manner.
- Provide PTM6 test booklets for schools to administer prior to the intervention, and a courier to collect the completed booklets from the school.
- Carry out the randomisation of settings to either the intervention or comparison group and notify schools and University of Oxford of group allocations.
- Ensure all NFER staff with access to evaluation data have DBS clearance.
- Liaise with the school to ensure all NFER staff visiting the school are familiar with standard safeguarding procedures, have received appropriate training on safeguarding protocols and have had an enhanced DBS check.

- Submit an anonymised safeguarding incident report to EEF covering any safeguarding concerns that emerge over the course of the evaluation.
- Carry out additional data collection activities (staff interviews, session delivery observations, pupil focus groups) with a subsample of eight schools.
- Send NFER Test Administrators who are trained and have Disclosure and Barring Service (DBS) clearance to all schools in June 2025 to complete endpoint assessment of all children in participating classes using PTM7.
- Provide all schools with assessment feedback reports at the end of the evaluation period.
- Store, share and analyse all evaluation data using secure systems.
- Analyse evaluation data, prepare and disseminate research findings.
- Provide schools with links to the published report.

Confirmation

I confirm that I have read and understood the information provided about the project, and I have passed a copy of the School Information Letter to my designated named key contact. I have had the opportunity to ask questions and have had these answered satisfactorily.

I understand that my school's participation is voluntary and that I am free to withdraw my school at any time. I will let the University of Oxford and NFER know if I choose to withdraw my school from the research.

I understand that, if my school withdraws from the evaluation, NFER will use the evaluation data my school has already provided, unless a parent has requested that their child's data not be used before the data has been anonymised.

I agree that my school will deliver the intervention as prescribed if allocated to the Intervention Group.

I know whom I can contact if I have any concerns or complaints about the project.

My school **will** take part in this research and agrees to the conditions stated in the Memorandum of Understanding (MoU).

We commit to the implementation (if allocated to the intervention group) and evaluation of the MR programme as set out in this document.

Headteacher's signature: _____

_____

Date: _____

Please sign and submit this form at your earliest convenience to , **reasoningfirst@education.ox.ac.uk** before 01.07.24

Headteacher's name	
Headteacher's telephone number	
Headteacher's email	
School name	
School address and post code	

Please provide details of key contact person for this research whom NFER and University of Oxford will primarily liaise with.

Key contact's name	
Key contact's phone number	
Key contact's email	
Key contact job role	

Year 2 Teacher's name	
Year 2 Teacher's email	

School Bursar's name	
School Bursar's email	

Appendix K: School Information Sheet

Effectiveness Trial of the Mathematical Reasoning Programme 2024-2025

School Information Sheet

About the Mathematical Reasoning Programme

Mathematical Reasoning (MR) is a 12 to 15-week programme developed by the University of Oxford for pupils in Year 2. The programme aims to improve mathematical attainment by developing pupils' understanding of the logical principles underlying mathematics, primarily:

- Quantitative Reasoning - the ability to reason about quantities and relations between quantities with or without numbers; and
- Arithmetic - the ability to reason about relations between numbers using the four operations, with a specific focus on additive composition and the inverse relation between addition and subtraction.

What does the Mathematical Reasoning programme involve?

The MR programme consists of twelve teaching units, designed to last approximately 12 to 15 weeks, with children receiving approximately one hour of MR content per week as part of their normal mathematics lessons. The programme is fully resourced and includes online games for pupils to use.

The teacher and TA for each participating Year 2 class are supported to deliver the programme through an online training course and support provided by the University of Oxford. The training course includes:

i. Six core modules designed to:

- Develop teachers' and TAs' understanding of why mathematical reasoning is important and what quantitative reasoning, arithmetic and number sense involve.
- Provide practical guidance on programme delivery.
- Support teachers and TAs to maintain gains and to expand the teaching of mathematical reasoning after the programme ends.

ii. Three *optional* modules for how mathematical reasoning can be taught to other age groups. These modules cover teaching fractions, using diagrams to reason about problems, and reasoning about different types of quantities.

The training course is fully online so can be completed at a time convenient for teachers and TAs. The core modules take approximately 2 days to complete, and the first five modules need to be completed in November 2024. This is so teachers and TAs can begin delivering the Mathematical Reasoning programme to Year 2 pupils in place of one of their regular maths lessons each week either before or immediately after the Christmas break. Teachers and TAs will also have access to three live interactive webinars with a Teacher Leader (TL) over the course of the trial (total time 1 day), as well as personalised support from a TL via an online forum. Schools will retain access to the training and to the programme throughout the school year 2024-25.

About the Effectiveness Trial

The Education Endowment Foundation (EEF) is funding the National Foundation for Educational Research (NFER) to evaluate the MR programme, which is delivered by the Year 2 teachers with support from the University of Oxford. The evaluation aims to:

- Determine the impact of the MR programme on the mathematical attainment of Year 2 children using GL Assessment's Progress Test in Maths (PTM).

- Explore how best to implement the programme on a larger scale in future by finding out how well the online training worked at scale.

The evaluation will run as a Randomised Controlled Trial (RCT) in the 2024-25 academic year. NFER will randomly allocate 240 eligible schools to one of the two following groups:

1. The intervention group, which will deliver the MR programme; or
2. The comparison group, which will continue with usual practice.

All pupils in participating classes will be assessed by their teacher using GL Assessment's Progress Test in Maths (PTM-6) in autumn 2024 (before randomisation). Once programme delivery is complete, an NFER Test Administrator will assess the children using PTM-7 in the summer of 2025.

Teachers and TAs in all participating schools will be asked to complete brief surveys at the start and end of the delivery period. We will extract pseudonymised data about the pupils' performance in the online games from the programme website.

Teacher responses to the questions on the on-line training course will be collected, anonymised, and shared with the evaluation team.

A small number of schools will also be invited to participate in additional data collection activities, including staff interviews, session delivery observations and focus groups.

Who can take part in this trial?

To be eligible to take part in the current evaluation, schools **must**:

- Be state primary schools in England;
- **Not** be participating in the Maths-Whizz or Maths Mastery EEF trials;
- Should they be randomly allocated to the intervention group:

a. Be willing to release the teacher and TA for each participating class for 2 days to complete the training and a total of 1 day to attend the online webinars;

b. Be able to provide each participating class with the screen required for the teacher to use PowerPoint slides for the whole-class component and individual devices with an internet connection for pupils to use when working on the computer games component of the programme;

- Agree to facilitate the evaluation data collection activities outlined in these documents.

Schools can choose to run the programme in one or more Year 2 classes. Mixed-year group classes are eligible, but only Year 2 pupils will be assessed as part of the evaluation.

Ethical approvals

This study has been approved by the University of Oxford's and NFER's ethical committees. There are robust safeguarding procedures in place to protect children during all stages of this evaluation.

What are the benefits or risks of my school participating in this trial?

We do not anticipate any risks or harm to the schools, staff or pupils participating in this study.

All participating schools will receive pupils' individual endpoint assessment scores. In addition, your school will be contributing to the evidence base about impact of the MR programme on Year 2 mathematical attainment and helping us learn more about the programme's implementation.

Schools allocated to the intervention group will receive all the training and resources for delivery of the MR programme free of charge. Once all the data collection has been completed, schools allocated to the comparison group will receive £500 as an acknowledgement for their participation in this study. These funds can be used at the discretion of your school.

What is the timeline for my school's involvement?

Time	**Event**
Feb - July 2024	• Schools sign a Memorandum of Understanding (MoU) agreeing the details of the project. • NFER contacts schools who have signed the MoU and provides the parent information letter and withdrawal form. • Schools provide NFER with pupil data (name, date of birth, UPN, FSM, gender, class name) using NFER's secure portal.
September 2024	• All schools administer the PTM6 assessment to all pupils in Y2 participating in the evaluation (using booklets supplied by NFER and in accordance with assessment guidance) and return the booklets to NFER using a secure courier. Schools will not incur any expenses as part of the assessment process. • Schools inform NFER of any pupils who have left the schools since pupil data was collected by completing the assessment attendance register. • All participating teachers complete a brief online survey containing questions about their usual practice in maths teaching and their views about mathematical reasoning. • Schools will be informed of dates and times of online training events and be asked to make arrangements to participate in the training, if they are randomly assigned to the intervention group.
October 2024	• All schools are notified of their group allocation (intervention or comparison).
November 2024	**Intervention schools**: • Online launch event for the programme. Head Teachers and Maths Leads are encouraged to participate alongside the teachers and TAs. • Teachers and TAs complete the online training. **Comparison schools**: • Continue with usual practice.
December 2024 – April 2025	**Intervention schools**: • Deliver the MR programme and complete the session delivery log. • Teachers and TAs attend three live training webinars. • A small number of schools (optional) participate in staff interviews, session delivery observations and pupil focus groups. **Comparison schools**

	• Continue with usual practice.
June 2025	**All schools**: • Facilitate visit/s from an NFER test administrator to administer the PTM7 to all participating Year 2 children. **Intervention schools**: • Year 2 teachers and TAs complete a brief online survey focused on their experiences of using the intervention, perceived outcomes, and their views about mathematical reasoning. **Comparison schools**: • Year 2 teachers complete an online survey focused on their usual practice and their views about mathematical reasoning.
July 2025	• Oxford University liaises with comparison schools about payment.
September 2025	• All schools receive pupils' individual endpoint assessment scores.
March 2026	• Publication of the final report.

Use of Data

NFER is the data controller for the evaluation so will decide what personal data is used and how, following the study objectives set by the EEF, the Data Protection Act 2018 and General Data Protection Regulation (GDPR), and will treat all data with the strictest confidence.

NFER's school and parent/carer **privacy notices** set out the legal basis for processing personal data and the ways NFER will use personal data for the purposes of this evaluation. Please ensure that you review these before signing the Memorandum of Understanding (MOU).

No school, staff member or pupil will be named in any report arising from this work, nor will we include any information that might mean that someone else could identify them.

NFER will delete any personal data collected within one year of the final report being published. This deletion process is currently planned for March 2027. Interview recordings will be deleted as soon as the interview notes/transcriptions have been finalised.

The University of Oxford is the data controller for the programme delivery. You can find their privacy notice **here.**

Who should I contact if I have questions?

For further information about the MR programme, please get in touch with the Research Project Manager, Louise Matthews at **reasoningfirst@education.ox.ac.uk**

For further information about the evaluation, please get in touch with the NFER evaluation team at **MR-Y2@nfer.ac.uk.**

Appendix L: School Privacy Notice

School privacy notice for the effectiveness trial of Mathematical Reasoning in Year 2

1 Why are we collecting this data?

The Education Endowment Foundation (EEF) has commissioned the National Foundation for Educational Research (NFER) to carry out an evaluation of the Mathematical Reasoning programme, developed and delivered by the University of Oxford. The purpose of this evaluation is to understand whether the Mathematical Reasoning (MR) programme can help to improve pupils' mathematical attainment when rolled out using an online professional development programme for teachers. Please refer to the School Information Sheet (available **here**) for more details about what the MR programme and evaluation will involve.

The trial will take place over the course of the 2024-25 academic year, with the programme delivered between November 2024 and April 2025.

This document outlines how **staff personal data** will be collected from schools and processed for the purposes of the evaluation. There is a **separate privacy notice for parents and carers** (available **here**), which covers the pupils' personal data collected as part of the evaluation.

NFER is the data controller for evaluation and will make decisions about how and what personal data is used in accordance with the objectives of the study set by EEF. The University of Oxford is the data processor for the evaluation and data controller for the programme delivery. There is a separate privacy notice, produced by University of Oxford, which covers how data is processed for programme delivery and some of their own research programme about programme implementation. This is available on NFER's **project webpage**.

2 What is the legal basis for processing activities?

The legal basis for processing personal data is covered by GDPR Article 6 (1) (f):

Legitimate interests: the processing is necessary for your (or a third party's) legitimate interests unless there is a good reason to protect the individual's personal data which overrides those legitimate interests.

A legitimate interest assessment has been undertaken. The evaluation fulfils one of NFER's core business purposes (undertaking research, evaluation, and information activities). It has broader societal benefits and will contribute to improving the lives of learners by providing evidence about the impact of teaching techniques used in the classroom. Research cannot be done without processing personal data, but NFER will ensure your rights are respected at all times.

3 How will personal data be obtained?

All personal data will be collected directly from participating schools.

The University of Oxford will collect staff personal data via the Memorandum of Understanding (MoU) when the schools sign up to participate in the project. This data will be shared with NFER via NFER's secure portal.

NFER will collect participating teacher and Teaching Assistant (TA) details via a Teacher and Pupil data template during the set-up phase of the project. This data will be shared via NFER's secure school portal. Schools will also share pupil data with NFER in the same Teacher and Pupil data template; further details about the pupil data being collected for the evaluation can be found in the parent and carer privacy notice **here**.

Responses to questions asked as part of the online training course will be collected by Oxford University through the online training course website, which are saved on the University's safe Canvas space, and will be shared with NFER in pseudonymised form using NFER's secure portal.

A member of the NFER research team will observe a number of the support webinars offered by the programme Teacher Leaders, in which teachers and TAs will participate, to better understand the nature of the support being

offered. Webinar participants will be notified of this observation in advance. Oxford University will also provide NFER with a list of which schools were represented at each of the webinars.

All teachers will be asked to respond to two online surveys. TAs in intervention schools will be asked to respond to one online survey at the end of delivery.

A small number of schools will be invited to participate in anonymised case studies. These will involve interviews with the participating class teacher and TA, as well as the Maths Lead, and observations of programme session delivery. Participation in these additional activities is optional.

4 What personal data is being collected by this project?

The full name and contact details of the Headteacher will be collected when a school signs up to participate in the project.

As part of the MoU, the full name, contact details and job role will be collected for the key contact on the project, who will facilitate communications between the school and the University of Oxford, and the school and NFER. The full name and the contact details for the school bursar will also be collected as part of the MoU.

The full name, role and contact details for the class teacher and TA in each participating class, will be collected for the evaluation to facilitate contact with participants.

Oxford University will share pseudonymised Teacher and TA responses to questions asked as part of the online training course with NFER so that NFER can analyse rates of training completion.

The surveys and interviews will ask about staff experiences of delivering the programme, challenges, and facilitators relevant to their school context and self-reported knowledge about and attitudes towards mathematical reasoning.

5 Who will personal data be shared with?

Personal data collected by the University of Oxford will be shared with NFER using NFER's secure online portal.

No personal data will be shared with anyone beyond the project team at NFER and the project team at the University of Oxford.

After the evaluation report has been published, NFER will share anonymised pupil data with the team from the University of Oxford. The University of Oxford team will then match this data with pseudonymised teacher data and school IDs that they will collect and hold. Further details about the research by the University of Oxford team can be found in their Privacy Notice, which is available on NFER's **project webpage**.

No individual will be named in any report for this project.

The online survey data is collected via Tivian which stores survey responses in Azure and AWS data centre infrastructure in the European Union. Tivian's privacy notice is available **here.**

6 Is personal data being transferred outside of the UK?

NFER only stores personal data in the UK. When we use Tivian to administer online surveys, data is stored in the EU.

7 How long will personal data be retained?

NFER will delete any personal data collected as part of the evaluation within one year of publication of the final report. This deletion process is currently planned for March 2027.

Interviews recordings will be deleted as soon as the interview notes/transcripts have been finalised.

8 Can I stop my personal data being used?

NFER handles personal data in accordance with the rights given to individuals under data protection legislation. If at any time you wish us to withdraw your data or correct errors in it, please contact the project team, by email at MR-Y2@nfer.ac.uk.

In certain circumstances, data subjects have the right to restrict or object to processing. They also have the right to make a subject access request to see all the information held about them. NFER will cooperate fully when a subject access request (SAR) is made of them as a data controller. To exercise these rights, please contact NFER's Compliance Officer at **compliance@nfer.ac.uk**.

Requests for access, withdrawal or corrections are not possible after the data have been anonymised.

9 Who can I contact about this project?

NFER is responsible for the day-to-day management of the evaluation. If you have any queries about the project, please contact the NFER project team by email at **MR-Y2@nfer.ac.uk**.

If you have a concern about the way this evaluation processes personal data, we request that you raise your concern with us in the first instance (see details above). Alternatively, you can contact the Information Commissioner's Office, the body responsible for enforcing data protection legislation in the UK, at **https://ico.org.uk/concerns/**.

10 Updates

We keep this privacy notice under review to make sure it is up to date and accurate. Any changes will be noted. We recommend that you revisit this information from time to time. The date when this privacy notice was last updated is shown in the footer at the bottom of this document.

January 2025: This change updates the name and privacy notice of the online survey provider.

Appendix M: Parent information sheet

Effectiveness Trial of the Mathematical Reasoning in Year 2 Programme 2024-2025

Parent and Guardian's information sheet and withdrawal form

Dear Parent / Guardian,

We are writing to you because your child's school has agreed to be part of a research study. The study will look at the effects of the Mathematical Reasoning in Year 2 Programme (MR programme) developed by the University of Oxford to improve children's mathematical attainment. Previous research has found this programme can help improve children's numerical skills. Please see the enclosed programme flyer for more information.

The current study is being run by the National Foundation for Educational Research (NFER) and is funded by the Education Endowment Foundation (EEF). It aims to test if the programme can improve children's numerical skills at a larger scale. It is *not* an assessment of individual children.

About the research design

We will use a method called a 'randomised controlled trial'. This means we will randomly choose some schools to use the MR programme in the 2024-25 academic year, while others will continue with their usual maths lessons. By comparing these two groups, we can see if the MR programme makes a difference to children's learning. This method is considered a fair and effective way to test new educational programmes.

Your child's school will be one of 240 schools randomly allocated to one of the two following groups:

1. The intervention group, which will deliver the MR programme; or
2. The comparison group, which will continue with usual practice.

What will the research involve for my child?

To see if the MR programme has helped improve children's maths skills, schools will be asked to complete the following activities:

- Share your child's full name, date of birth, sex, unique pupil number (UPN) and Free School Meal (FSM) status with NFER.
- Provide information about which of the programme sessions your child attended, as well as the number of programme computer games they played, the scores they achieved and the times at which they played them.
- Complete a maths test in autumn 2024.
- Complete a maths test in summer 2025.

The maths tests we will use are often used by schools. This means there will not be anything unusual for your child. Your child's class teacher will carry it out in 2024 and a trained NFER Test Administrator will do it in 2025. We will also extract data about which of the games that are part of the programme your child played and when, from the programme website. Your child's name is not stored on the programme website, but your child will be given a unique project identification to login to the website.

In addition, NFER will also collect pupil sex, Special Educational Needs (SEN) status, English as an Additional Language (EAL) status and eligibility for Free School Meals (FSM) from the Department for Education's (DfE) National Pupil Database (NPD).

Finally, a small number of classes will be visited by NFER research team to observe how the programme sessions work in practice. It is possible that your child will be in the classroom, but the observation is of the sessions and not of individual

children. No personal data (e.g., names) will be collected. Following the observation, four children from the class will be invited to participate in a short, child-friendly group conversation with the teacher present. The children will be asked about how much they like mathematics and if they liked the MR programme. If your child is selected to participate in one of these conversations, we will be in touch to provide more information about this. We will also ask the children if they agree to participate. The children's names will not be recorded during these conversations.

Further information about the trial is available **here**.

What are the benefits and risks of my child taking part?

We do not anticipate any risks or harm to the school, teachers or children participating in this study. This study has been approved by the University of Oxford's and NFER's research ethics committees. NFER has robust safeguarding protocols in place to protect children during the research activities and schools will know what these procedures are.

By participating in this study, your child will be helping us to:

- find out if the MR programme can help improve mathematical attainment for Year 2 pupils, and
- improve how the programme is delivered, which will benefit other children in the future.

How will the findings be used?

The evaluation report will be shared with schools and freely available on the NFER and EEF websites. In addition, the University of Oxford team will write a brief report to be shared with all participating schools. Schools will also receive the maths test scores for individual children from the tests taken in 2025. This is to help schools understand children's progress so they can provide any support needed.

How will NFER use and protect the data collected?

NFER have robust procedures in place to make sure that we comply with the General Data Protection Regulation (GDPR), which sets out how data can be collected and stored securely. We will use number codes to link the children's names with their maths test scores and with the information about the games they played. The children's names will not be kept in the same files as their test results. We have summarised how we will collect, use, and protect personal data collected during this study in the enclosed privacy notice (also available **here**). See also 'What will the research involve for my child?'

No schools, teachers or pupils will be named in the reporting, and we will not include any other information that may mean someone else could identify them. No identifiable pupil data will be made public or shared beyond the research team other than the pupils' maths test scores in 2025, which will be shared with their school.

Does my child have to participate?

The MR programme will be used by the schools in the intervention group as part of their normal mathematics curriculum. The children in all schools participating in the research will be asked to complete the maths tests as part of the school's activities. If you do not want your child's results to be used in the research, your child does not have to participate.

How can I withdraw my child from the study?

If you **are happy** for your child's data to be used for this research, **you do not need to do anything**.

However, if you would like **to withdraw** your child's data from the study, **please complete the form below** and return it to your child's school. You can withdraw your child from

the data collection of the trial at any time before 1 September 2025. After that we will not be able to identify your child's results. If you withdraw your child from the project, we will use any data we have collected up to that point in our analysis unless you ask us to delete it.

If you choose to withdraw your child's data from the evaluation, this will not affect your child negatively nor impact their involvement in classroom activities around the project.

Who can I contact for further information?

If you have any queries, please contact us via email **MR-Y2@nfer.ac.uk**.

Yours sincerely,

Kathryn Hurd

Evaluation and Survey Operations Lead

National Foundation for Educational Research

Effectiveness trial of the Mathematical Reasoning in Y2 Programme

Parent Withdrawal Form

You only need to complete this form if you DO NOT wish your child's data to be shared, stored, and used for this research.

I **DO NOT** give permission for data about my child that is collected as part of the **Mathematical Reasoning in Y2 Programme** trial to be shared, stored, or used for research purposes.

Your child's name……Child's class:……

Name of school……

Your full name……

Your telephone number (optional)……

Your signature…… Date……

Appendix N: Parent Privacy Notice

Parent and carers privacy notice for the effectiveness trial of Mathematical Reasoning in Year 2

1 Why are we collecting this data?

The Education Endowment Foundation (EEF) has commissioned the National Foundation for Educational Research (NFER) to carry out an evaluation of the Mathematical Reasoning programme, developed and delivered by the University of Oxford. The purpose of this evaluation is to understand whether the Mathematical Reasoning programme will help to improve pupils' mathematical attainment. More information about what the programme and evaluation will involve can be found **here**.

The trial will take place over the course of the 2024-25 academic year. Teachers will deliver the Mathematical Reasoning sessions once a week to pupils in Year 2 between November 2024 and April 2025.

This document outlines how pupils' personal data will be collected and processed for the purposes of this evaluation. Your child's data will be treated with the strictest confidence and processed in accordance with the GDPR and the Data Protection Act 2018. We will not use your child's name in any report arising from the research.

NFER is the data controller for the evaluation and will make decisions about what personal data is used and how in accordance with the objectives of the study set by the EEF. University of Oxford is the data processor for the evaluation. This means that they will process some of the data collected in order to support the evaluation, but their access is restricted to anonymous data collected from the children.

The University of Oxford is the data controller for the programme delivery and has produced a separate privacy notice to cover how personal data is processed for it. This is available on our **project webpage**. They will also collect data from the teachers and their mentors about the programme implementation, but not about individual children for their own research purposes.

2 What is the legal basis for processing activities?

The legal basis for processing personal data is covered by GDPR Article 6 (1) (f):

Legitimate interests: the processing is necessary for your (or a third party's) legitimate interests unless there is a good reason to protect the individual's personal data which overrides those legitimate interests.

A legitimate interest assessment has been undertaken. The evaluation fulfils one of NFER's core business purposes (undertaking research, evaluation, and information activities). It has broader societal benefits and will contribute to improving the lives of learners by providing evidence about the impact of teaching techniques used in the classroom. Research cannot be done without processing personal data, but NFER will ensure your child's rights are respected at all times.

The legal basis for processing pupils' special personal data is covered by GDPR Article 9 (2) (j) which states that *'processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) (as supplemented by section 19 of the 2018 Act) based on domestic law which shall be proportionate to the aim pursued, respect the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject'.*

We do not believe this processing will cause damage or distress to the pupils. The outcomes of the evaluation will not result in the creation of measures or decisions being made about individual pupils.

3 How will personal data be obtained?

NFER will collect personal pupil data from all participating schools using NFER's secure online portal. All assessment papers will be sent out from and returned to NFER using a secure courier.

NFER will collect background data including gender, Free School Meal (FSM) eligibility and Special Education Needs (SEN) status for the trial pupils from the Department for Education's (DfE) National Pupil Database (NPD) using the DfE's secure exchange portal and the Office of national Statistics' secure access procedures.

In addition, eight schools will be selected to participate in anonymous case studies. This will involve a visit from the NFER research team to observe how the programme sessions work in practice. Your child may be observed as part of this, but no further personal data will be collected. Four pupils from each case study school will also be invited to share their views about the programme by participating in a short child-friendly focus group with the teacher present. If your child is selected to participate in a focus group we will contact you in advance to provide you with more information and ask for your permission to invite your child to take part.

4 What personal data is being collected by this project?

The following data will be collected for the pupils in each class participating in the study:

- Full name
- Unique Pupil Number (UPN)
- Date of birth
- Gender
- Free School Meal (FSM) status
- Progress Test in Maths (PtM) assessment scores (prior to and following programme delivery)
- Attendance at programme sessions

Number of programme computer games played, their scores achieved in the games and the time at which they started playing them

When matching the trial pupils' data to NPD, the following types of information about your child will be collected:

- Gender
- English as an Additional Language (EAL) status*
- Eligibility for Free School Meals (FSM)

Special Educational Needs (SEN) status*

Pupils selected for focus groups will be asked about their attitudes towards mathematics and their experiences of the programme.

*considered to be special category personal data

5 Who will personal data be shared with?

NFER will share your child's full name, date of birth and UPN with the DfE's NPD team to request the additional background data listed above. NFER will also share this data with GL Assessment to enable marking of the PtM assessments. Your child's assessment score in 2025 will also be shared with their school.

Anonymised pupil assessment scores will be shared with the University of Oxford team after the evaluation report has been published. The University of Oxford will then match this data with pseudonymised teacher data and school IDs that they will collect and hold. Further details around the University of Oxford's research can be found in their Privacy Notice, available on NFER's **project webpage**.

Three months after the publication of the evaluation report, pseudonymised[2] matched outcome data will be added to the EEF archive, which is managed by the Fischer Family Trust on behalf of EEF and hosted by the Office of National Statistics. This will enable the EEF and other research teams to use the pseudonymised data as part of subsequent research through the ONS Approved Researcher Scheme. Further matching to NPD and other administrative data may take place as part of this research, as per statement on page 5 of **this document**. At this point, EEF will become the data controller and is responsible for taking decisions about the means and purposes of processing. **FFT Education** (FFT) and the designated EEF archive evaluator will act as data processors for the archive. Further information about the EEF archive is available **here**.

We will not name any of the children in any report for this project, nor will we include any information that might mean someone else may be able to identify them.

6 Is personal data being transferred outside of the UK?

No personal data will be stored or transferred outside of the UK by NFER.

GL Assessment may transfer personal data outside of the UK. However, this is safeguarded by the appropriate contractual safeguard.

7 How long will personal data be retained?

NFER will delete any personal data collected as part of the evaluation within one year of publication of the final report. This deletion process is currently planned for March 2027.

Focus group recordings will be deleted as soon as the interview notes / transcripts have been finalised.

8 Can I stop my personal data being used?

NFER handles personal data in accordance with the rights given to individuals under data protection legislation. If at any time you wish us to withdraw your data or correct errors in it, please contact the project team, by email at MR-Y2@nfer.ac.uk.

In certain circumstances, data subjects have the right to restrict or object to processing. They also have the right to make a subject access request to see all the information held about them. NFER will cooperate fully when a subject access request (SAR) is made of them as a data controller. To exercise these rights, please contact NFER's Compliance Officer at **compliance@nfer.ac.uk**.

9 Who can I contact about this project?

NFER is responsible for the day-to-day management of the evaluation. If you have any queries about the project, please contact the NFER project team by email at **MR-Y2@nfer.ac.uk**.

If you have a concern about the way this evaluation processes personal data, we request that you raise your concern with us in the first instance (see details above). Alternatively, you can contact the Information Commissioner's Office, the body responsible for enforcing data protection legislation in the UK, at **https://ico.org.uk/concerns/**.

10 Updates

We keep this privacy notice under review to make sure it is up to date and accurate. Any changes will be noted. We recommend that you revisit this information from time to time. The date when this privacy notice was last updated is shown in the footer at the bottom of this document.

8 March 2024: Information about the EEF archiving process added.

31 October 2024: Clarification added about English as an Additional Language being classified as special category data.

Appendix O: Teacher outcomes questions

Understanding of maths topics						
How would you rate your understanding of the following maths topics?						
Please select one response per row.	1	2	3	4	5	6
	Very good	Good	Average	Poor	Very poor	I don't know what this is
Doing and undoing (or reversing) an action						
Using part-whole relations to solve problems						
Counting on						
Interpreting tally charts						
The inverse relation between addition and subtraction						
Additive composition						
Measuring lengths						
Quantitative reasoning						
The difference between additive and multiplicative reasoning						
Recalling multiplication facts						
One-to-many correspondence						

Teaching maths in Year 2					
How confident are you in teaching the following topics to your Year 2 maths class?					
Please select one response per row.	1	2	3	4	5
	Very confident	Fairly confident	Not very confident	Not at all confident	I don't know
Doing and undoing (or reversing) an action					
Using part-whole relations to solve problems					
Counting on					
Interpreting tally charts					
The inverse relation between addition and subtraction					
Additive composition					
Measuring lengths					
Quantitative reasoning					
The difference between additive and multiplicative reasoning					
Recalling multiplication facts					
One-to-many correspondence					

Appendix P: Sources, methods and assumptions for calculating the cost of the programme

Appendix P Table 1: Sources, methods and assumptions for calculating staff time spent on the programme

Activity	Source(s)	Methods and assumptions
Training course	Endpoint Teacher & TA Survey	The average (mean) time spent on the training course as reported by those teachers and TAs who reported doing at least some of it. The mean assumes 10.5 hours for those respondents (1 TA, 11 teachers) that selected '10+ hours' as the scale went up in half-hour increments and it is unlikely that these respondents spent significantly more time on the training than their peers. However, it is still possible that this represents an underestimate – particularly in the case of the maximum. It is assumed that there is no need to train new teachers or TAs within the three-year period.
Webinars	*Length of webinars:* TL report *Attendance:* Endpoint Teacher & TA survey	The average (mean) time spent is calculated based on the average length of a webinar multiplied by the mean number of webinars attended. Minimum time spent is the lower value resulting from the following calculations: • Average number of webinars attended multiplied by the minimum reported webinar length • Minimum number of webinars attended multiplied by the average reported webinar length The equivalent calculation was made for the maximum value. These calculations include time for the launch event, assuming that the teacher attended the launch event, as 98% of schools did so. However, the TA time does not assume attendance at the launch event as the IPE findings suggest that this was uncommon. It is assumed that there is no need to train new teachers or TAs within the three-year period.
Liaising with the TL	Case study cost pro formas	The average (mean) of the time spent liaising with TLs, as reported by case study teachers. It is assumed, based on the IPE findings, that the TA did not spend any time liaising with the TL.
Session preparation	*Time spent per session:* Case study cost pro formas *Number of sessions delivered:* Session delivery logs	Average (mean) time spent on session preparation was calculated based on the mean amount of time case study teachers reported spending on session preparation compared to usual lesson preparation practice, multiplied by the average number of sessions delivered. Minimum time spent is the lower value resulting from the following calculations: • Average amount of time spent preparing per session multiplied by the minimum number of sessions delivered

		• Minimum amount of time spent preparing per session multiplied by the average number of sessions delivered The equivalent calculation was made for the maximum value.
Session delivery	*Length of sessions:* Endpoint Teacher & TA Survey *Number of sessions delivered:* Session delivery logs	The average (mean) time spent on session delivery was calculated based on the mean session length reported by teachers multiplied by the mean number of sessions delivered. These calculations assume 2.25 hours for the response option '2+ hours' as the scale went up in quarter-hour increments and it is unlikely that these respondents spent significantly more time on the training than their peers. However, it is still possible that this represents an underestimate – particularly in the case of the maximum. Minimum time spent is the lower value resulting from the following calculations: • Average session length multiplied by the minimum number of sessions delivered • Minimum session length multiplied by the average number of sessions delivered The equivalent calculation was made for the maximum value.

Appendix P Table 2: Sources, methods and assumptions for calculating the cost of the programme

Nature of cost	Source(s)	Methods and assumptions
Programme materials	Oxford University delivery team	NA
Access to the programme training	Oxford University delivery team	It is assumed that no new teachers or TAs need to be trained in the 3-year period.
Printing (certificates)	Case study cost pro formas	The average (mean) was calculated from the reported costs, including where teachers reported zero printing costs.
Cover for training course completion	Endpoint Teacher & TA Survey	Average (mean) cost was calculated based on mean teacher and TA time spent on the training course (in days) multiplied by the average reported cost of cover per day (£218) (CFE Research, 2024). Although the cost of TA cover is likely to be less than that of a teacher cover, the same figure has been used for both due to no available data on TA cover costs (ie likely to be an over-estimate). It is assumed that no additional teachers or TAs have to be trained in the MR programme over the three-year period.
Cover for webinar attendance	Endpoint Teacher & TA Survey	Some schools reported that paid cover could not be secured just for the duration of the webinar but had to be for the whole afternoon. As a result, average (mean) cost of cover for webinar attendance was calculated based on the mean number of webinars attended by teachers and TAs, multiplied by half a day, multiplied by half the average reported cost of cover per day (CFE Research, 2024). It is assumed that no additional teachers or TAs have to be trained in the MR programme over the three-year period.

Cost of personal devices	NA	Average cost assumes 25 pupils per class with devices required for half of the class at a time (12 pupils). Cost per device estimated at £110.50 based on 91 tablets purchased by NFER for use in the NELI Preschool trial conducted over the same period. However, it is possible that schools purchasing a smaller number of devices may have to pay more per device.
Cost of manipulatives	Case study cost pro formas	The average (mean) was calculated from the reported costs from only those teachers that reported purchasing additional manipulatives.

References

Barrett, P. (2007) 'Structural equation modelling: adjudging model fit', *Personality and Individual Differences*, 42(5), pp. 815–824. Available at: https://doi.org/10.1016/j.paid.2006.09.018.

CFE Research (2024) *Use of supply teachers in schools*. Research Report. Department for Education. Available at: https://assets.publishing.service.gov.uk/media/66f2b45966439d663cf12bb0/Use_of_supply_teachers_in_schools_research_report.pdf.

Fornell, C. and Larcker, D.F. (1981) 'Evaluating structural equation models with unobservable variables and measurement error', *Journal of Marketing Research*, 18(1), pp. 39–50. Available at: https://doi.org/10.1177/002224378101800104.

GL Assessment (2015) 'Technical information: Progress Test in Maths (PTM)'. Available at: https://support.gl-assessment.co.uk/media/3035/ptm-technical-information.pdf (Accessed: 29 January 2026).

Hair, J.F., Anderson, R., Tatham, R. and Black, W. (1998) *Multivariate data analysis*. Fifth. Hoboken: New Jersey: Prentice Hall. Available at: https://books.google.co.uk/books?id=-ZGsQgAACAAJ (Accessed: 29 January 2026).

Hu, L. and Bentler, P.M. (1999) 'Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives', *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), pp. 1–55. Available at: https://doi.org/10.1080/10705519909540118.

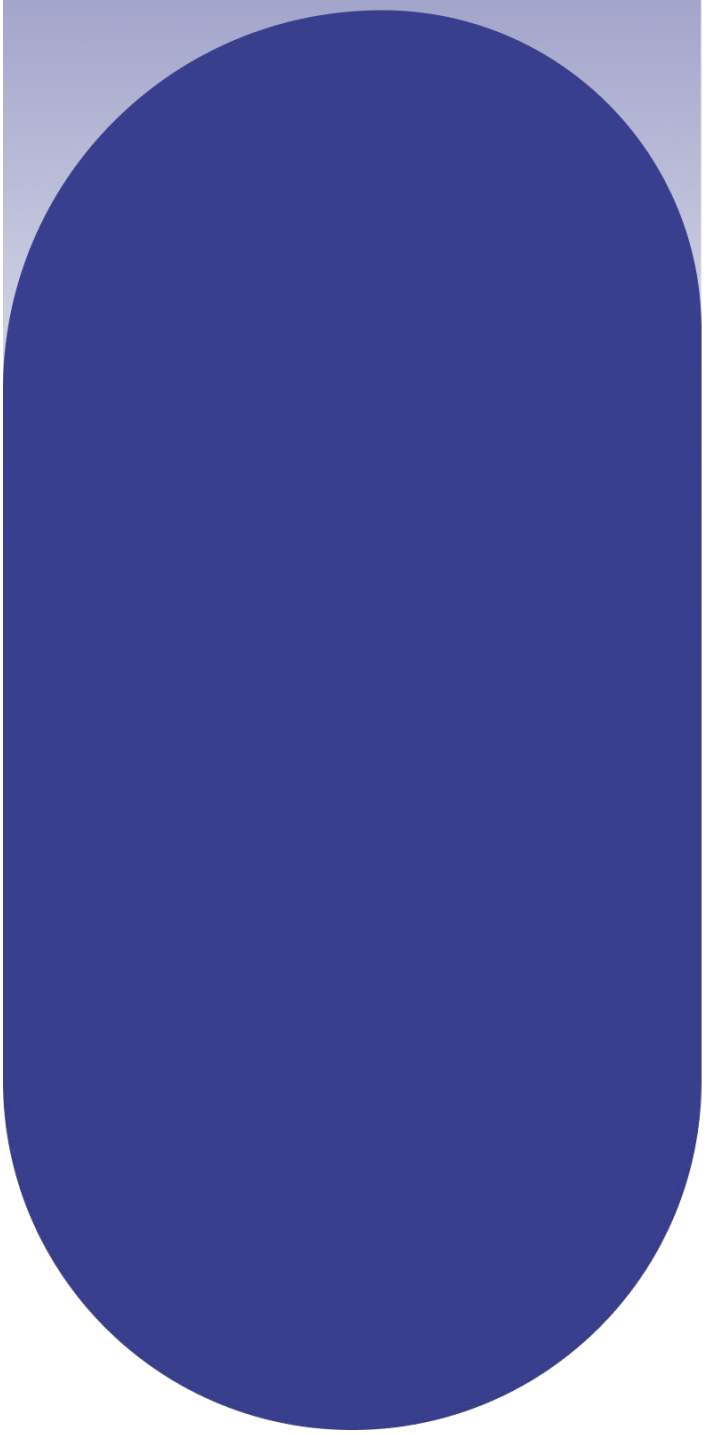
Li, C.-H. (2016) 'Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares', *Behavior Research Methods*, 48(3), pp. 936–949. Available at: https://doi.org/10.3758/s13428-015-0619-7.

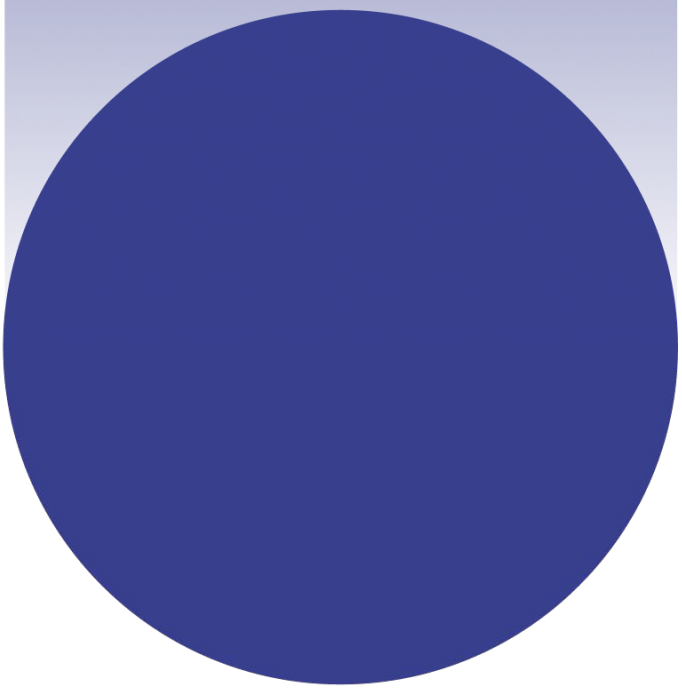
Mardia, K.V. (1970) 'Measures of multivariate skewness and kurtosis with applications', *Biometrika*, 57(3), pp. 519–530. Available at: https://doi.org/10.1093/biomet/57.3.519.

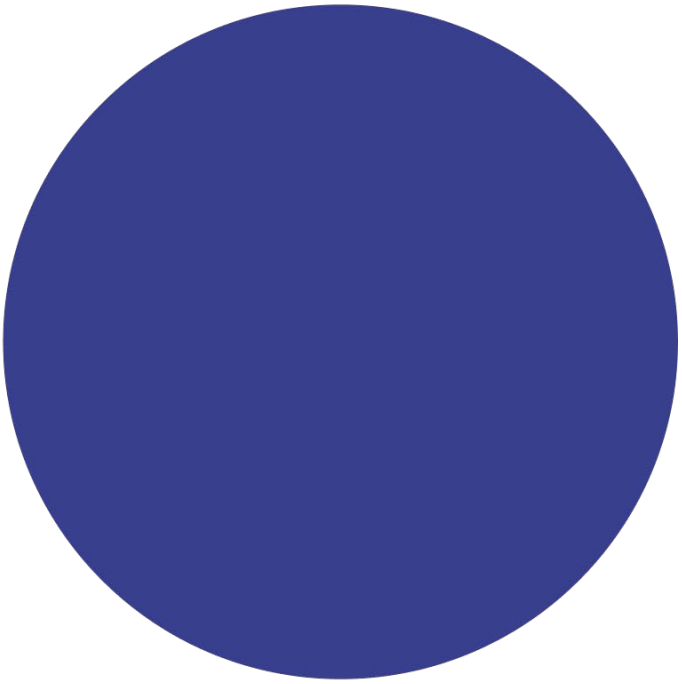
Marsh, H.W., Hau, K.-T. and Wen, Z. (2004) 'In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings', *Structural Equation Modeling: A Multidisciplinary Journal*, 11(3), pp. 320–341. Available at: https://doi.org/10.1207/s15328007sem1103_2.

Nunnally, J.C. (1978) *Psychometric theory*. 2nd edn. New York: McGraw-Hill.

Satorra, A. (2000) 'Scaled and adjusted restricted tests in multi-sample analysis of moment structures', in R.D.H. Heijmans, D.S.G. Pollock, and A. Satorra (eds) *Innovations in multivariate statistical analysis: a Festschrift for Heinz Neudecker*. London: Kluwer Academic Publishers, pp. 233–247. Available at: https://doi.org/10.1007/978-1-4615-4603-0_17.







Education Endowment Foundation

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

https://educationendowmentfoundation.org.uk

@EducEndowFoundn

Facebook.com/EducEndowFoundn