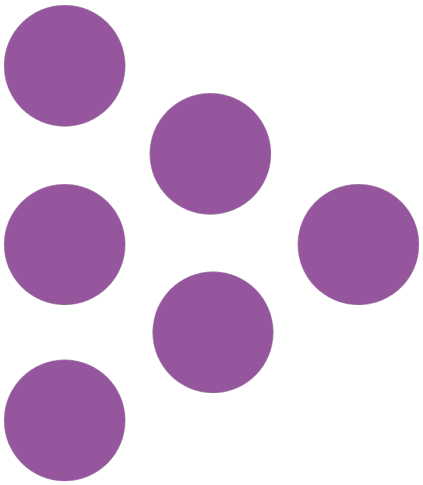
NFER
National Foundation for Educational Research

Report

Action Tutoring Year 6 Mathematics Programme Quasi-Experimental Impact Evaluation

National Foundation for Educational Research (NFER)



Action Tutoring Year 6 Mathematics Programme Quasi-Experimental Impact Evaluation

Ruth Staunton and Stephen Welbourne

Published in April 2026

By the National Foundation for Educational Research,

The Mere, Upton Park, Slough, Berkshire SL1 2DQ

https://www.nfer.ac.uk/

ISBN: 978-1-916567-61-0

How to cite this publication:

Staunton, R. and Welbourne, S. (2026). *Action Tutoring Year 6 Mathematics Programme Quasi-Experimental Impact Evaluation.* Slough: NFER

NFER
National Foundation for Educational Research

Contents

About the evaluator

This project was independently conducted by a research team from the National Foundation for Educational Research (NFER). The NFER is the leading independent provider of education research and holds the status of Independent Research Organisation (IRO) from UK Research and Innovation (UKRI). Our unique position and approach delivers evidence-based insights designed to enable education policy makers and practitioners to take action to improve outcomes for children and young people. Our key topic areas are accountability, assessment, classroom practice, education to employment, social mobility, school funding, school workforce and systems and structures. As a not-for-profit organisation, we re-invest any surplus funds into self-funded research and development to further contribute to the science and knowledge of education research.

www.nfer.ac.uk

@TheNFER

Contact details:

Dr. Stephen Welbourne, Research Director

National Foundation for Educational Research (NFER)

The Mere, Upton Park, Slough SL1 2DQ

Email: enquiries@nfer.ac.uk

Tel: +44 (0)1753 574123

Acknowledgements

The research team would like to thank the Action Tutoring team for their support throughout this project.

1. Background and Study Rationale

There is robust evidence that small group tuition (defined as one tutor working with two to five pupils in a group) is effective at improving pupil attainment. The average impact demonstrated in primary school pupils is four months additional progress (EEF, 2021). There is also a body of evidence that the advantages of small group tuition may be particularly relevant for disadvantaged pupils (Dietrichson et al., 2017; Torgerson et al., 2018). However, the evidence around the effectiveness of different delivery models is less secure. Action Tutoring were one of the 33 approved tuition partners for the National Tutoring Programme (NTP), part of the government's COVID-19 catch-up response for schools and were subsequently included in the evaluations of the NTP. The evaluation of the first year of the NTP (Poet et al., 2022) considered the impact of delivery mechanisms such as group size and time of delivery but this was across all tuition partners and Action Tutoring's volunteer-led model was not specifically explored. In 2014 a QED evaluating the impact of Action Tutoring's secondary school programme demonstrated a positive association between the number of tutoring sessions attended and GCSE grades (Lucchino, 2016). A further secondary school evaluation (Dorsett, 2017) found positive impacts on GCSE grades for students who attended Action Tutoring sessions and had scored below level 4 at key stage 2.

This evaluation sought to quantify the impact of Action Tutoring's Year 6 maths programme on disadvantaged pupil attainment in maths at Key Stage 2 for the academic year 2023-24.

1.1. Intervention

Action Tutoring partners with state schools to provide volunteer-led tutoring to disadvantaged pupils. They ask schools to select up to 20 pupils (65% of whom have to be Pupil Premium (PP), and with the remaining places prioritised for pupils whom the school knows would not otherwise be able to pay for private tutoring). Along with this focus on disadvantage, the programme is designed for pupils who are middle attainers, with more than 68% classed as 'working towards expected standard' at Key Stage 2 by their classroom teachers. These are the only specified selection criteria for pupils. Pupils sit a baseline assessment before tutoring begins and this is used by the tutors to inform a personalised application of the Action Tutoring Curriculum, designed by subject specialists. All tutors are volunteers. Typically, a pupil will have access to weekly tuition for between 15-20 weeks but attend an average of 13 sessions (hours). Tuition takes place on school premises and group sizes vary between 2 and 3 pupils per tutor. In each school an average of 18 pupils in the same year group are involved in the programme. Presently Action Tutoring are working with 50+ primary schools who are accessing maths tutoring for Year 6.

1.2. Research Questions

This evaluation was deigned to answer the following two research questions:

RQ1: What is the impact of school participation in Action Tutoring's Year 6 mathematics tuition on disadvantaged pupils' KS2 maths attainment?

RQ2: What is the impact of Action Tutoring's Year 6 mathematics tuition on disadvantaged pupils' KS2 maths attainment?

2. Design

This evaluation took a Quasi-Experimental Design (QED) approach to answer the research questions, with school-level matching. A pupil-level matching approach was considered but was ultimately dropped in favour of school-level matching for the following reasons:

1. We had previously found the assumption that it is possible to construct a control group of pupils who would have been selected for tutoring, if their school had been using Action Tutoring, to be unsupported. NFER devoted considerable efforts in our evaluations of the National Tutoring Programme (NTP) into designing sophisticated models using NPD variables to predict tutoring selection, but they never performed much better than chance. This suggested that selection for tutoring is either based on criteria not represented by data available in the NPD or is random. From informal communication with schools, we strongly considered the former to be the more likely scenario.
2. Analysis biases introduced by the above issue tend to be negative - meaning that analyses based on this method would underestimate the true size of the tutoring effect. We believe this is because teachers tend to select those pupils who they feel are in danger of falling behind and when these are compared with a poorly matched comparison group it would look like intervention pupils are doing less well even though tutoring may well have had significant benefits.

Therefore, we proposed that a better approach was to conduct an analysis where the variable of interest was whether or not the *school* participated in Action Tutoring's Year 6 mathematics programme. There is still the possibility of bias introduced by selection at the school level, but a recent study (Weidmann and Miratrix, 2021) suggests that reasonable matching at the school level keeps these biases very small. The one downside of this approach was that it is subject to dilution as not all PP pupils in the Action Tutoring schools received tuition in maths (45% of PP pupils, see below). It was possible to correct for the dilution, however, using this knowledge of the proportion of PP pupils in Action Tutoring schools who received tutoring in maths (RQ2).

2.1. Participants

Participants in the intervention are described in the Intervention section above. For the analysis, intervention participants were defined as i) having been eligible for free school meals (FSM) in the previous 6 years (as a measure of disadvantage), ii) in Year 6 during the 2023/24 academic year, and iii) in schools who implemented the Action Tutoring Year 6 mathematics tuition. Comparison participants were pupils who i) had been eligible for FSM in the previous 6 years (FSM6), ii) were in Year 6 during the 2023/24 academic year, and iii) were in a matched set of comparison schools (using school characteristics from the NPD for matching). Although PP rather than FSM is one of the criteria for selection into the tutoring programme, we consider FSM eligibility an acceptable analytical substitute for PP pupils in this analysis using the following reasoning. FSM eligibility in the previous 6 years is one of the criteria for PP funding. In the 2023/24 academic year, 24.3% of primary pupils in England were eligible for FSM[1]. In the 2023/24 financial year 24.5% of primary pupils in England were PP pupils[2]. Despite the difference in time period covered, we believe this is

[1] https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics/2023-24
[2] https://explore-education-statistics.service.gov.uk/find-statistics/school-funding-statistics/2023-24

good evidence that identifying participants using the FSM6 variable included the overwhelming majority of PP pupils in our analysis sample.

2.2. Outcome Measures

The outcome measure for both research questions was KS2 maths score in the 2023/24 academic year. Due to cancellation of KS1 assessments during the COVID-19 pandemic, this cohort of pupils did not have KS1 data to use as a baseline, so their Year 4 Multiplication Tables Check data were utilised for this purpose.

2.3. Sample size

Below we note different pupil samples and parameters used to calculate the MDES at each stage of the analysis (see Table 1).

Table 1 - Parameters used to calculate MDES

	Design		Analysis	
Parameter	**Value**	**Source**	**Value**	**Source**
Number of Action Tutoring schools (delivering Year 6 maths programme)	61	Communication from Action Tutoring	61	Analysed data
Number of matched comparison schools (drawn from NPD, see below for details)	61		305	
Number of tutored Y6 PP pupils	566		567	
Number of PP Y6 pupils (expected) in 61 schools	693	Explore educational statistics[3] website (version from 8th June 2023)	1266	
Intraclass correlation coefficient (ICC)	0.11	EEF 'Improving Power Calculations in Education Trials'[4] publication	0.09	Primary analysis
Pre-post correlation	0 or 0.7		0.605	
Power	80%	Standard practice	80%	Standard practice
Significance threshold	5%		5%	

[3] https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics
[4] https://d2tic4wvo1iusb.cloudfront.net/production/documents/evaluation/methodological-research-and-innovations/Work_Package_2023-WP6_18_09_2023_FINAL.pdf?v=1713850351

As described in the Outcome Measures section, the outcome variable is KS2 maths score and the baseline variable is Year 4 Multiplication Tables Check. As this is the first cohort who had both Year 4 Multiplication Tables Check and Key Stage 2 data, we did not have an estimate for the correlation between these two measures at the design stage. Therefore, we calculated the MDES for the worst-case scenario (correlation=0) and a probable best-case scenario (correlation=0.7 as seen between KS1 and KS2 maths scores). The pre-post correlation at analysis stage was 0.605.

At design stage we used the numbers of pupils in the table above to estimate that 82% (566 out of 693) of Year 6 PP pupils in the Action Tutoring schools would receive maths tuition. However, at analysis stage it was determined that the intervention schools had higher numbers of PP pupils than average schools, so the percentage of Year 6 PP pupils tutored was lower than anticipated at 45%. These percentages have been used to calculate the MDES corrected for dilution in Table 2 below. This can be thought of as the effect that would need to exist in the tutored pupils to allow detection of the stated MDES when analysing all pupils in the analysis population.

Table 2 - Minimum Detectable Effect Size (MDES)

Stage	Pre-post correlation	ICC	RQ1 MDES	Dilution	RQ2 MDES, among tutored pupils (i.e. corrected for dilution)
Design	0	0.11	0.22	18%	0.27
Design	0.7	0.11	0.20	18%	0.24
Analysis	0.605	0.09	0.14	55%	0.32

2.4. Matching

In this evaluation we sought to minimise bias in the analysis by creating a comparison group of schools with characteristics that matched our intervention schools as closely as possible. The comparison pool was limited to maintained English primary schools where no hours of NTP tuition were recorded and where the percentage of free school meals for year 6 pupils was above the intervention group minimum (13%) and below the intervention group maximum (85%). 305 schools were randomly selected from the comparison group pool to give a 5:1 ratio between comparison and intervention schools. We then applied entropy balancing (Hainmueller, 2012) to generate weights for each school that ensured balance in the matching characteristics. Intended matching characteristics were:

- Historic attainment at KS2
- % free school meals (FSM)
- % English as an Additional Language (EAL)
- Income Deprivation Affecting Children Index (IDACI) score
- Number of pupils in the school
- Establishment type
- Mean teacher pay

- Pupil to qualified teacher ratio
- Geographic region
- Urban/rural

Detail on the outcome of the matching can be found in the Findings section below.

2.5. Statistical Analysis

2.5.1. RQ1

The primary outcome of KS2 maths score was used as the dependent variable in a linear mixed effects model. Fixed effects were intervention group, baseline score from the Year 4 Multiplication Tables Check. The random effect was school to account for the hierarchical nature of pupils clustered within schools. Model weights were taken from the entropy balancing and applied at school level. The analysis population was FSM6 pupils. Model coefficients were used to calculate the effect size between the intervention and comparison schools. This is reported with 95% confidence intervals and a *p*-value for difference from 0. The regression model is defined by:

$$\boldsymbol{KS2MathsScore_{ij}} = \boldsymbol{\beta_{oj}} + \boldsymbol{\beta_1 Intervention_j} + \boldsymbol{\beta_2 MTC_{ij}} + \varepsilon_{ij}$$

Where:

- $\boldsymbol{KS2MathsScore_{ij}}$ is the maths scaled score of pupil i in school j from the KS2 exam in the 2023/24 academic year.
- $\boldsymbol{MTC_{ij}}$ is the multiplication table check score for pupil i in school j from the Year 4 exam in the 2021/22 academic year.
- $\boldsymbol{Intervention_j}$ is a variable indicating whether school j participated in Action Tutoring's Year 6 maths programme in the 2023/24 academic year.
- $\boldsymbol{\beta_{oj}}$ is the intercept in school j (modelled as a random effect).
- $\boldsymbol{\beta_1}$ is the coefficient of interest estimating how much difference participation in the programme makes to KS2 maths score.
- $\boldsymbol{\beta_2}$ is a coefficient estimating the association between baseline MTC score and KS2 maths score.
- ε_{ij} is the residual error term for pupil i in school j.

Distributional assumptions of the model were assessed through visual inspection of the residual plots. No evidence of heterogeneity of variance was apparent.

2.5.2. RQ2

The effect size and confidence interval of the effect from RQ1 were divided by the proportion of PP pupils in Action Tutoring schools who received tutoring in maths to calculate an estimated pupil-level effect size.

2.5.3. Subgroup analysis

It was noted during analysis that the number of schools in some of the subgroup categories/regions were very low (≤3). This made the analysis model unstable and prevented many results from being shared due to statistical disclosure control. In discussion with Action Tutoring, we agreed to combine South East and South West, leave London as a category and combine all other regions into one category. This decision was made based on the number of schools in each region and Action Tutoring's knowledge of their tutoring hubs and was made before any analysis results were shared.

The geographic location subgroup analysis repeated the primary analysis but included an interaction term between the intervention group and geographic region. Within each geographic region subgroup, the between intervention group effect size is reported with 95% confidence intervals and a *p*-value for difference from 0.

2.5.4. Effect size estimation

The primary effect size is reported as Hedges *g* calculated using the following equation:

$$ES = (\bar{Y}_T - \bar{Y}_C)_{adjusted} / sd_{pooled}$$

Where:

$(\bar{Y}_T - \bar{Y}_C)_{adjusted}$ denotes difference in means between groups adjusting for baseline workload. In our models this is equivalent to the coefficient $\boldsymbol{\beta_1}$ so this model coefficient was used in the calculation.

sd_{pooled} is the unconditional standard deviation of primary outcome measure pooled across the two groups, calculated as

$$s_p = \sqrt{\frac{(n_T - 1)s_T^2 + (n_C - 1)s_C^2}{(n_T - 1) + (n_C - 1)}}$$

Confidence intervals for each impact estimate were estimated by adding/subtracting from the point estimate the standard errors of the intervention coefficient (β_1) multiplied by the left-tailed inverse of the Student's t-distribution with a probability of 2.5% and the number of degrees of freedom associated with the size of the sample. The confidence around the impact estimates was converted to effect size confidence intervals using the same formula as the effect sizes themselves.

In addition, we have reported the effect sizes as scaled scores and we have converted the Hedge's *g* effect sizes to additional month's progress using the table in EEF's toolkit[5].

[5] https://d2tic4wvo1iusb.cloudfront.net/documents/toolkit/Toolkit_guide_v1.2_-_2023.pdf?v=1676921207

3. Findings

3.1. Matching

We started the matching process with 61 intervention schools and 16,923 possible comparison schools. 8089 of the comparison pool were discarded due to our exclusion criteria that comparison schools must record no hours of tuition for Year 6 pupils through the NTP school-led tutoring. 3379 further schools were discarded from the comparison pool through common support. The table below shows the actions taken to ensure common support for each of the matching variables.

Table 3 - Actions to ensure common support

Matching variable	Action
• % free school meals (FSM)	Schools discarded which were outside of the intervention group range (13% to 85%)
• Geographic region	Schools discarded if they were in a region which did not appear in intervention group (East of England)
• Historic attainment at KS2 • Income Deprivation Affecting Children Index (IDACI) score • Number of pupils in the school • Mean teacher pay • Pupil to qualified teacher ratio	Schools discarded which recorded NA for these variables
• % English as an Additional Language (EAL) • Establishment type • Urban/rural	No schools discarded

A random subset of 305 of the 5516 schools in the comparison school pool were selected as the comparison schools. Entropy balancing was attempted with all matching variables, but the optimisation arrived at a degenerate solution necessitating reduction of the matching variables. Urban/rural, establishment type and number of pupils in the school were removed[6]. Urban/rural, establishment type and region all had categories with very low numbers of intervention schools but since region was the subgroup for this evaluation, we felt it was important to retain as a matching variable.

The pre and post weighting percentages and means are shown in Table 4 below. Where unweighted counts of schools are less than three, disclosure rules require suppression of the true counts, and suppression of the corresponding weighted counts.

[6] Distributions of these variables are still included in the balance tables below.

Table 4 - Unweighted and weighted counts and percentages or means and standard deviations

Categorical variables	Level	Unweighted		Weighted	
		Intervention Group	Comparison Group	Intervention Group	Comparison Group
		N (%)	N (%)	N (%)	N (%)
Total		61 (100%)	305 (100%)	61 (100%)	305 (100%)
Type of Establishment	Academy converter	8 (13.1%)	90 (29.5%)	7.5 (12.3%)	84 (27.5%)
	Academy sponsor led	13 (21.3%)	36 (11.8%)	18.9 (31%)	37.3 (12.2%)
	Community school	>26 (>42.6%)	107 (35.1%)	>10.1 (>16.6%)	113.2 (37.1%)
	Foundation school	3 (4.9%)	9 (3.0%)	9.4 (15.4%)	7.9 (2.6%)
	Free schools	<3 (<4.9%)	3 (1.0%)	<3 (<4.9%)	3.6 (1.2%)
	Voluntary aided school	5 (8.2%)	37 (12.1%)	9.1 (14.9%)	37.4 (12.3%)
	Voluntary controlled school	<3 (<4.9%)	23 (7.5%)	<3 (<4.9%)	21.6 (7.1%)
Region	All other regions	19 (31.1%)	191 (62.6%)	35.0 (57.4%)	175 (57.4%)
	London	29 (47.5%)	32 (10.5%)	10.2 (16.7%)	50.8 (16.7%)
	SW & SE	13 (21.3%)	82 (26.9%)	15.8 (25.9%)	79.2 (26.0%)
Urban/Rural	Rural	<3 (<4.9%)	85 (27.9%)	<3 (<4.9%)	76.3 (25.0%)
	Urban	>58 (>95.1%)	220 (72.1%)	>58 (>95.1%)	228.7 (75.0%)
Continuous Variables	**Statistic**	**Unweighted**		**Weighted**	
		Intervention Group	**Comparison Group**	**Intervention Group**	**Comparison Group**
Average KS2 maths scaled score in 2022/23	Mean	103.9	103.6	103.6	103.6
	SD	3.0	2.9	223.9	28.8
Y6 Percentage FSM	Mean	49.2	35.1	37.5	37.5
	SD	16.7	15.4	85	24.7
Y6 Percentage EAL	Mean	36.5	20	22.8	22.8
	SD	23.7	22.6	48.8	30.8
IDACI Score	Mean	0.26	0.17	0.19	0.19
	SD	0.12	0.11	0.45	0.16
Number Of Pupils	Mean	410	304.1	487.6	314.1
	SD	234.5	187.2	1263.1	242.1
	Mean	47280.4	44696.1	45126.7	45126.8

Categorical variables	Level	Unweighted		Weighted	
		Intervention Group	Comparison Group	Intervention Group	Comparison Group
		N (%)	N (%)	N (%)	N (%)
Mean Teacher Pay	SD	5499.1	4045.9	97590.3	15162.4
Pupil To Qualified Teacher Ratio	Mean	19.4	20.8	20.6	20.6
	SD	3.0	3.6	43.6	5.9

Variables that were included in the matching procedure were very well balanced after matching. Before weighting, intervention schools were more likely to be in London, had higher FSM and EAL percentages and higher IDACI score than comparison schools. After weighting these were equal between intervention and comparison schools. Characteristics not included as matching variables were less well balanced after matching, as might be expected. Intervention schools were more likely to be urban and had more pupils than comparison schools, both of which persisted after weighting.

3.2. RQ1

FSM pupils in the intervention group on average scored one point higher in KS2 maths than those in the control group (equivalent to two months progress). This effect did not reach the accepted threshold for statistical significance of $p<0.05$.

Table 5 – RQ1 model means

Group	N School (weighted)	N pupil (weighted)	KS2 Maths Score (95% CI)
Comparison	304[7]	4332	101.4 (101.0, 101.8)
Intervention	61	1064	102.4 (101.0, 103.9)

[7] One of the comparison group schools was missing baseline scores for all FSM year 6 pupils (although not all year 6 pupils in the school) so was not included in the analysis.

Figure 1 - RQ1 model means

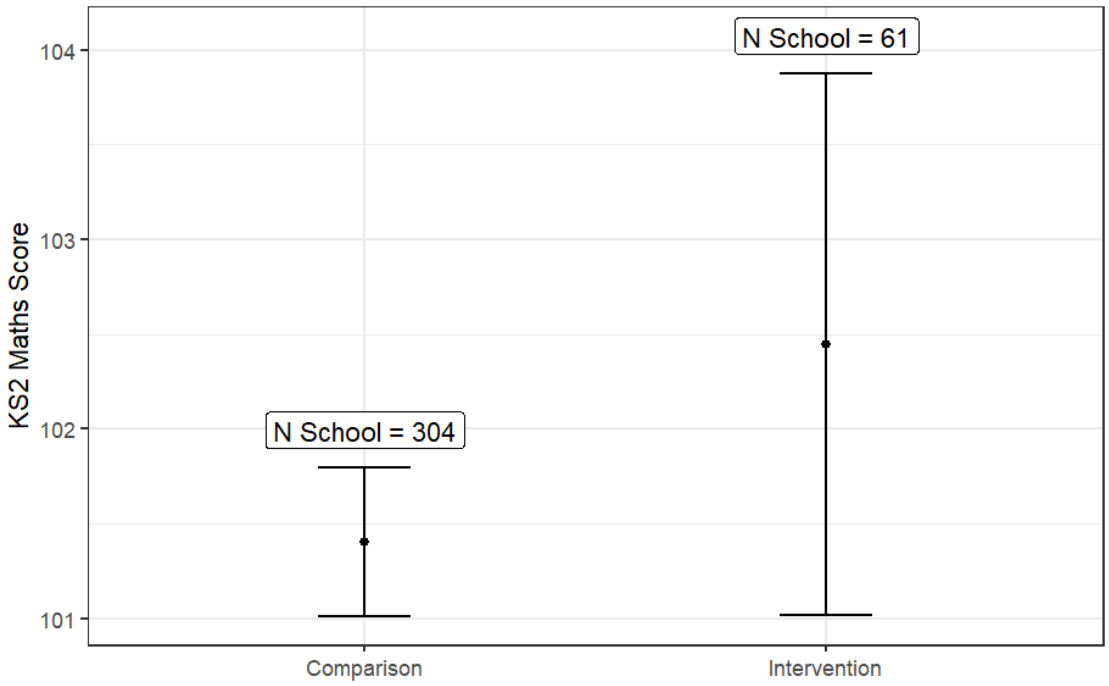


Table 6 - RQ1 effect

Comparison	Effect Size (95% CI)	Difference in Maths Score (95% CI)	Months Progress[5]	*p*-value
Intervention-Comparison	0.13 (-0.03, 0.30)	1.04 (-0.25, 2.34)	+2 months	0.11

3.3. RQ2

As detailed in the Sample size section, the percentage of Year 6 PP pupils in the Action Tutoring schools that received maths tuition was 45% (567/1266). Dividing the primary RQ effect size by 0.45 gives the pupil level effect sizes detailed in Table 7 below.

Table 7 – RQ2 effect

Comparison	Effect Size (95% CI)	Difference in Maths Score (95% CI)	Months Progress[5]	*p*-value
Intervention-Comparison	0.30 (-0.07, 0.66)	2.32 (-0.56, 5.20)	+4 months	0.11

3.4. Subgroup analysis

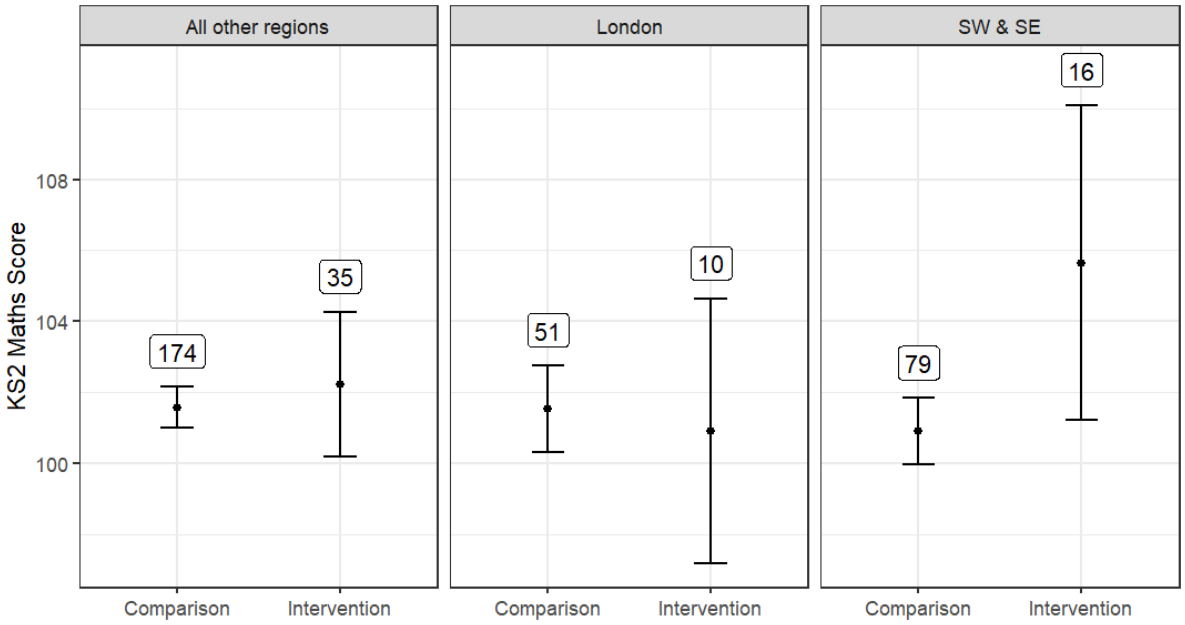
The subgroup analysis population is the same as RQ1 i.e. all year 6 FSM pupils in intervention or comparison schools. In the combined South East & South West subgroup, FSM pupils in

intervention schools had significantly higher KS2 maths scores than FSM pupils in comparison schools. The difference of 5 points on the maths score scale gives an effect size of 0.61, which is considered high impact and is equivalent to +7 months progress. The intervention vs comparison group contrasts did not reach statistical significance in either of the other region subgroups.

Table 8 - Subgroup analysis model means

Region	N school (weighted) Intervention; Comparison	Intervention KS2 Maths Score (95% CI)	Comparison KS2 Maths Score (95% CI)	Effect Size (95% CI)	*p*-value
South East & South West	16; 79	105.6 (101.2, 110.1)	100.9 (100, 101.8)	0.61 (0.08, 1.13)	0.02
London	10; 51	100.9 (97.2, 104.6)	101.5 (100.3, 102.8)	-0.08 (-0.53, 0.38)	0.97
All other regions	35; 174	102.2 (100.2, 104.3)	101.6 (101, 102.2)	0.08 (-0.16, 0.33)	0.80

Figure 2 - Subgroup analysis model means



4. Additional analysis

The following analyses were not included in the study plan and should be considered exploratory additions to the planned work.

4.1. Compliant school subset

Action Tutoring shared data on the number of sessions offered and attended in each of the intervention schools. We used percentage of offered sessions attended to select the top 50% most 'compliant' schools, giving a subset of 31 intervention schools where >84% of offered sessions were attended. We reran the matching and the primary analysis using this subset and results are presented below.

In the compliant subset, FSM pupils in intervention schools had significantly higher KS2 maths scores than FSM pupils in comparison schools.

Table 9 – Compliant subset model means

Group	N School (weighted)	N pupil (weighted)	KS2 Maths Score (95% CI)
Comparison	155	2431	101.8 (101.3, 102.3)
Intervention	63	651	103.2 (102.0, 104.5)

Figure 3 – Compliant subset model means

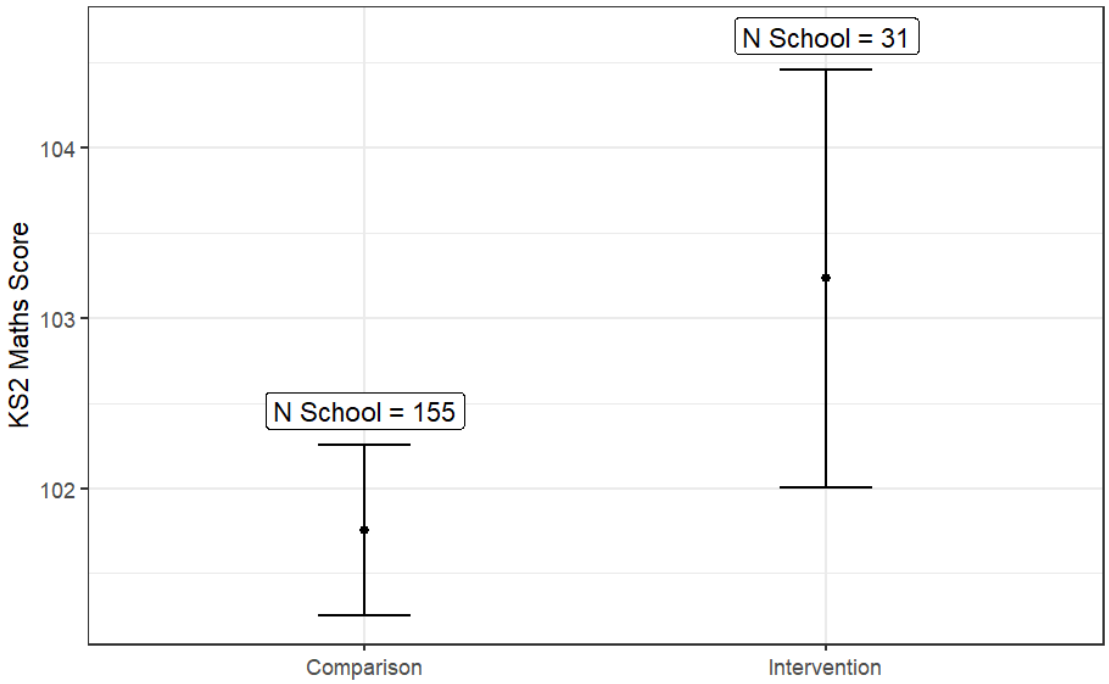


Table 10 – Compliant subset effect

Comparison	Effect Size (95% CI)	Difference in Maths Score (95% CI)	Months Progress[5]	*p*-value
Intervention-Comparison	0.18 (0.04, 0.34)	1.48 (0.32, 2.64)	+2 months	0.01

4.2. Sample size for a trial spanning multiple years

The number of schools in the intervention group was limited in this analysis by the maximum number of schools enrolled in Action Tutoring's Year 6 maths programme. For future evaluations this is likely to still be a limitation, so we have considered the MDES achievable if year 6 pupils from multiple consecutive academic years were included in the analysis. We have looked the MDES if the number of intervention schools stayed the same across the three years, and also if the number of schools increased by 5 each year. The table below shows the MDES using the analysis pre-post correlation (0.605) and ICC (0.09). We have adjusted the number of covariates to 2 to allow 'academic year' as a covariate.

Table 11 - Minimum Detectable Effect Size (MDES)

Number of years included	Number of covariates	Number of intervention schools	Average number of pupils per school	RQ1 MDES
1	1	61	14.4	0.142
2	2	61	28.8	0.131
3	2	61	43.1	0.127
2	2	66	28.8	0.126
3	2	71	43.1	0.117

Although there is a decrease in the MDES by including multiple years of data, unless the number of schools also increases it is not a large decrease. The modest size of this benefit should be considered in any cost-benefit discussions for future evaluations.

4.3. SEND

Interest was also expressed in understanding the differential effect of pupils recorded as SEND or not. Unfortunately, this was not part of the initial analysis plan so the data were not requested from the NPD data base, and therefore the subgroup analysis could not be undertaken in this evaluation. However, any future evaluations may wish to include this approach.

5. Conclusions

Although the direction of the effect is positive and the size of the effect represents a potentially meaningful change (+2 months progress), the effect in the primary research question (all year 6 FSM pupils) did not meet the threshold for statistical significance. The current reach of the Action Tutoring year 6 maths programme was a limitation on the number of schools in the intervention group for this analysis and consequently the statistical power of the evaluation. While the lack of significance means we cannot definitively state that a positive impact exists for disadvantaged year 6 pupils in the Action Tutoring maths programme, it should not be taken as evidence of no effect. The size and direction of the programme effect is evidence of promise and the post hoc power analyses demonstrate that significance for an effect of this size would be achievable with a modest increase in sample size and/or multiple years of year 6 data.

A sizeable positive significant effect was demonstrated in one of the region subgroups (South East & South West) despite the lower power inherent in subgroup analyses. Further exploration into any distinct practices engaged in by tutors working in these regions may generate beneficial knowledge for all small group tutors. The 'compliant schools' subset analysis also demonstrated a significant positive effect. This suggests that intervention schools where pupils attended more of the offered sessions had better KS2 maths outcomes than equivalent comparison schools. The compliant schools are not over-represented in the South East & South West region so these conclusions can be considered independently.

Two risks to the analysis identified at the outset of the trial were i) difficulty selecting an appropriate comparison pool and ii) dilution of the effect either because not all PP pupils were selected for tutoring, or because not all offered tutoring sessions were attended. During analysis the first of these risks was not realised and a sufficient comparison pool was available. As described in the matching section over 5000 schools were left in the comparison pool after applying our criteria and common support. The second risk was initially addressed through RQ2 to counter the dilutionary effect of not all PP pupils being selected for tuition. The dilution due to PP was larger than anticipated before the analysis. The estimated effect on target individuals was consequently large (effect size = 0.30, equivalent to +4 months progress) but, as a direct conversion of the RQ1 findings, still did not meet the 5% significance threshold. In addition, to address an alternative dilution avenue, we added the 'compliant subset' exploratory analysis described above.

This analysis is one of the first to use Multiplication Tables Check as a baseline for KS2 maths score. In the pre analysis sample size calculations we assumed the pre-post correlation could be anywhere between 0 (worst case scenario) and 0.7 (best case scenario). In the analysis the pre-post correlation was 0.605 which is a good amount of variability explained and demonstrates that MTC can be a relevant baseline for KS2 maths score in future research.

Future work in understanding the impact of the year 6 Action Tutoring maths programme could solidify the evidence of promise demonstrated in this work. Increasing the number of schools in any evaluation would improve the power to detect smaller but still meaningful effects at the population level (all year 6 FSM pupils). Although a pupil-level matching approach is not recommended for the bias concerns detailed above, identifying the pupils who received the intervention could allow a compliance style analysis to decompose the impact of the programme into direct effect and spillover effect.

References

Dietrichson, J., Bøg, M., Filges, T. and Klint Jørgensen, A.-M. (2017) 'Academic interventions for elementary and middle school students with low socioeconomic status: a systematic review and meta-analysis', *Review of Educational Research*, 87(2), pp. 243–282. Available at: https://doi.org/10.3102/0034654316687036.

Dorsett, R. (2017) *Action Tutoring's small-group tuition programme An impact evaluation for 2015/16 participants*. London: University of Westminster.

EEF (2021) *Small group tuition*. Available at: https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/small-group-tuition (Accessed: 16 April 2026).

Hainmueller, J. (2012) 'Entropy balancing for causal effects: a multivariate reweighting method to produce balanced samples in observational studies', *Political Analysis*, 20(1), pp. 25–46. Available at: https://doi.org/10.1093/pan/mpr025.

Lucchino, P. (2016) *Action Tutoring's small-group tuition programme. An impact evaluation using statistical comparison groups*. Available at: https://actiontutoring.org.uk/wp-content/uploads/2021/10/action_tutoring_impact_evaluation.pdf (Accessed: 16 April 2026).

Poet, H., Lord, P., Styles, B., Oppedisano, V., Zhang, M. and Dorsett, R. (2022) *Evaluation of year 1 of the Tuition Partners programme: impact evaluation for primary schools*. Available at: https://d2tic4wvo1iusb.cloudfront.net/documents/projects/Evaluation-Report-Primary-school-impact-evaluation.pdf (Accessed: 16 April 2026).

Torgerson, C., Bell, K., Coleman, E., Elliott, L., Fairhurst, C., Gascoine, L., Hewitt, C. and Torgerson, D. (2018) *Tutor Trust: affordable primary tuition. Evaluation report and executive summary*. Available at: https://dro.dur.ac.uk/26952/1/26952.pdf?DDD29+vrfd57+d700tmt (Accessed: 16 April 2026).

Weidmann, B. and Miratrix, L. (2021) 'Lurking inferential monsters? Quantifying selection bias in evaluations of school programs', *Journal of Policy Analysis and Management*, 40(3), pp. 964–986. Available at: https://doi.org/10.1002/pam.22236.

NFER
National Foundation for Educational Research

Evidence for excellence in education

Public

The Mere, Upton Park, Slough, Berks SL1 2DQ
T: +44 (0)1753 574123 • enquiries@nfer.ac.uk

www.nfer.ac.uk

NFER ref. ATMQ

ISBN. 978-1-916567-61-0