Department for Education

Longitudinal Evaluation of the Mathematics Teacher Exchange: China-England

Second interim research report

December 2017

Sean Demack, Tim Jay, Mark Boylan, Claire Wolstenholme, Anna Stevens, and Bronwen Maxwell

Sheffield Hallam University

GOVERNMENT SOCIAL RESEARCH
Social Science in Government

Contents

List of Tables

List of Figures

Glossary

Contrast school - A sample of schools with similar characteristics to lead primary schools. These were matched to the lead primary schools in a 20:1 ratio.

CPD - Continuing Professional Development.

Lead primary school or LPS - A school designated by the hub that hosted a Shanghai teacher and in nearly all cases had one or more members of staff visit Shanghai.

Maths Hubs - A network of hubs across England each led or jointly led by a school or college. Maths Hubs work in partnership with neighbouring schools, colleges, universities, CPD providers, maths experts and employers. There were 32 Maths Hubs in England at the start of the exchange and there are, as of November 2015, 35.

NCETM - The National Centre for Excellence in the Teaching of Mathematics.

Mathematics Teacher Exchange - Exchange programme involving 48 English primary schools and schools and teachers in Shanghai.

Executive summary

This report presents the design and initial analyses relating to the impact evaluation of the Mathematics Teacher Exchange: China-England. The Mathematics Teacher Exchange is a highly innovative programme that aims to foster a radical shift in primary mathematics teaching in England by learning from Shanghai mathematics education through a programme of teacher exchange. The evaluation of the Mathematics Teacher Exchange has two strands. Strand 1 evaluates changes in practice in schools involved (lead primary schools), and the influence on others they support in implementing teaching for mastery. Strand 2 evaluates impact on pupil outcomes in lead primary schools. This report describes details of Strand 2 methodology and reports baseline data:

- The main analysis for the impact evaluation will use Key Stage 1 and Key Stage 2 mathematics attainment data from academic years 2012/13 (two years before the exchange project began) through to 2016/17 (three years after the project began).
- Multilevel analyses (pupils clustered within schools) will be employed in order to account for similarities in the ways that pupils experience mathematics education within any given school.
- Each of the 47 lead primary schools participating in the exchange were matched with 20 contrast schools using propensity scores[1]. Any change in mathematics attainment in lead primary schools over the five year period covered by the analysis will be compared with change across the set of contrast schools.
- Assessments of the quality of matching have been carried out. While there is some instability at school level, at pupil level there is good consistency in attainment from year to year. This suggests that the matched contrast group represents a good set of data for comparison as the project progresses.
- Analysis of KS1 and KS2 data for the end of the first year of the exchange shows no change in KS1, and a very small increase in KS2 scores for participating schools relative to contrasts. These early analyses should not be interpreted as evidence for or against impact, however, as data were taken at a very early stage in the intervention.
- Alongside the analysis of mathematics attainment data, there will also be an analysis of pupil attitudes. This report presents pre-test data from a survey instrument measuring general affect towards mathematics, mathematics anxiety, and preference for working alone on mathematical tasks.

[1] Please see section 2 of this report and the appendix; The 1:20 match was complete for the matching year (2013/14) resulting in matching the 47 lead primary schools with 940 matched contrast control schools. However, due to issues of missing data in other academic years this match does not remain complete. For example, only 39 lead primary schools have complete data for the 2012/13; 2013/14 and 2014/15 matched to 718 contrast control school with similarly complete data.

- Limitations of the impact evaluation design are discussed. While the design will allow measurement of associations between the exchange and subsequent mathematics attainment, it will not allow the ascription of cause in this relationship. As the set of lead primary schools participating in the exchange was a non-random sample that is not representative of the full range of English primary schools, it may be that this set of schools would have performed well regardless of the exchange. This means that the interpretation of findings will need to be made with some degree of caution in subsequent reports.
- The impact evaluation will only be a measure of the success of the exchange in raising mathematics attainment, and not a measure of the effectiveness of a particular approach to mathematics teaching. Contrast schools may well be engaging in different forms of teaching for mastery, and may have participated in various forms of related professional development, but did not take part in the 2014/15 exchange.

1. Introduction

The Mathematics Teacher Exchange (formerly titled the England-China Maths Education Innovation Research Project) brings together two key approaches to improving mathematical teaching and attainment in England: benchmarking against, and learning from, high performing countries in relation to curriculum and teaching; and the move towards school-led system improvement, in this case through the Maths Hub initiative co-ordinated by the National Centre for Excellence in the Teaching of Mathematics (NCETM). The exchange is an innovative approach to both learning from Shanghai mathematics education, and to developing an English approach to teaching mathematics informed by what is learnt from Shanghai mathematics education. The Mathematics Teacher Exchange involved 48 lead primary schools located in 32 geographical hub areas. The evaluation of the Mathematics Teacher Exchange has two strands. Strand 1 evaluates changes in practice in schools involved - the lead primary schools. Further, Strand 1 evaluates the work of lead primary schools with other schools and the extent to which Shanghai informed teaching for mastery is being adopted more widely in Maths Hub networks. Strand 2 evaluates impact on pupil outcomes in lead primary schools. It is important to note that the impact evaluation discussed in this document is not an evaluation of mastery approaches to mathematics teaching in general, nor an evaluation of a particular approach to mathematics teaching. It aims to determine whether participation in the England-China exchange in 2014/15 has been associated with increases in mathematics attainment in KS1 and KS2 in participating schools.

This report is the second interim report of the Longitudinal Evaluation of the exchange. The first was published in July 2016[2]. The first report included detail of the exchange programme and its aims and initial findings related to implementation and change in practice. A further interim report in 2017 will report on-going changes in lead primary schools and their work with other schools. Final outcomes of both Strand 1 and Strand 2 will be reported in Spring 2018.

This report adds further details to the impact evaluation methodology for the project. It firstly provides a description of the quasi-experimental design with propensity score matching that will examine attainment outcomes in the lead primary schools in comparison to a contrast group of schools. It then provides analysis of Key Stage 1 and Key Stage 2 (KS1 and KS2, respectively) attainment that provide a baseline for the longitudinal evaluation. This preliminary analysis includes attainment data prior to the start of the intervention and from the end of the first year of the intervention. In addition, outcomes of a pupil attitude survey of Y6 pupils in lead primary schools are reported. These provide the baseline for a longitudinal analysis of how changes in practice

[2] https://www.gov.uk/government/publications/evaluation-of-the-maths-teacher-exchange-china-and-england.

influence pupil attitudes, and other affective factors associated with mathematics teaching and learning. Finally, limitations of the analytical approach are discussed.

The full impact evaluation will be published once analysis of 2016/17 attainment data has been carried out.

2. Analysis Plan

This section presents the planned impact analyses for the Mathematics Teacher Exchange evaluation. Mathematics attainment for five consecutive cohorts of KS1 and KS2 pupils (2012/13 to 2016/17) will be examined. Attainment amongst pupils within lead primary schools will be compared with the attainment amongst pupils within a set of matched contrast schools. This section summarises analyses to date, which use KS1 and KS2 attainment in 2013/14 and 2014/15, and are indicative of the analysis plan for the evaluation as a whole.

Capturing the impact of the Mathematics Teacher Exchange on pupil level maths attainment at KS1 and KS2.

Within the matched design, the counterfactual was captured in a post-hoc way. Prior to the start of this evaluation, schools had already been recruited to host the Mathematics Teacher Exchange. This precluded the possibility of using a Randomised Controlled Trial (RCT) research design that randomly assigned schools to either host the exchange or to be part of a control group.

Following the recruitment of lead primary schools, school census data was used to match 20 contrast schools to each of the schools hosting the Mathematics Teacher Exchange. Propensity scores were used to match lead primary schools participating in the exchange to schools not participating in the exchange. Matching methods attempt to find non-treated observations that are ‘similar’ to the treated observations (in this case, the schools that took part in the exchange), so that they can be compared to each other. Propensity score matching allows a match of treated and non-treated observations while considering a full range of observed characteristics.

In this research the propensity score represents the schools probability of obtaining a certain KS2 performance given the schools’ (observed) characteristic, and it is on this score that schools are matched. The matching was performed using school census KS2 data from the 2013/14 academic year. Attainment of successive cohorts of pupils at KS1 and KS2 will be looked at, covering five academic years from 2012/13 to 2016/17.

Identifying the contrast schools for the lead primary schools using propensity score matching.

At the school level, the lead primary schools were matched to contrast schools using data from the school census. Using the 2013/14 KS2 school census, propensity scores were created to match each lead primary school to 20 contrast schools. Propensity scores were derived from a binary logistic regression model for all primary schools in 2013/14 with KS2 attainment as the dependent variable. Ten explanatory variables were included in the model: KS1 attainment, KS1 to KS2 mathematics value added score, KS1 to KS2 progress in mathematics (boys and girls separately), school size, proportion of pupils claiming free school meals in any of the six previous years (%6FSM), proportion of pupils with English as an Additional Language (%EAL), proportion of pupils with Special Educational Needs (%SEN), proportion of female pupils (%Female), and school Income Deprivation Affecting Children Index (IDACI) score. The variables selected for the propensity score model were based upon the availability within the 2013/14 school census and discussions in the steering group. The propensity scores were then rank ordered and for each of the lead primary schools, 20 schools with the closest scores were selected into the contrast group. This 1:20 matching was completed for the 47 lead primary schools that delivered KS2[3]. This resulted in 940 contrast schools (see Appendix 2 for more detail on the propensity score matching). This approach should ensure that the sample of lead primary schools closely reflected the sample of 940 contrast schools across the 10 explanatory variables. However, this approach does not ensure that the two samples reflect each other on other unmeasured or unmeasurable factors.

The quality of the matches was then tested using data at both school and pupil level:

- At the school level, school census data from 2012/13, 2013/14 and 2014/15 was used to compare the lead primary schools with their matched contrast schools.
- At the pupil level, data from the National Pupil Database (NPD) from 2013/14 and 2014/15 was used to compare pupils in the lead primary schools with pupils in the matched contrast schools.

Looking across the three academic years used to test the matching at the school level, eight lead primary schools had incomplete data for the three academic years[4]. For this reason, these schools (and their matches) were excluded from the comparison. This

[3] One of the lead primary schools was an infant school that shared the same postcode as a junior school - in this instance the junior school only was matched.

[4] This is likely to relate to lead primary schools that became academies. During the transition to academy, schools are not required to submit detail on attainment to the school census.

results in a total of 39 lead primary schools with complete details in 2012/13, 2013/14 and 2014/15 matched to 718 contrast schools[5] with similarly complete details.

Testing the quality of the propensity matches at the school level

Table 1 summarises the school level comparison for the three academic years under observation. See Appendix 4 for more detail on this comparison6. For each of the seven variables, Table 1 shows the mean value for the 39 lead primary schools with complete records, alongside the mean value for the 718 matches. The mean difference between the lead primary and contrast sample is then shown as a Cohen's d effect size statistic.

Cohen's d is a widely used standardised statistic that enables effect sizes to be compared across outcomes on differing scales and across different studies, time points etc. According to the teaching and learning toolkit developed by the Educational Endowment Foundation (EEF)[7], a 'very high impact' is indicated by an effect size of (d=) +0.70 standard deviations or greater; 'high impact' by an effect size between +0.45 to less than +0.70 sds; 'moderate impact' by an effect size between +0.19 to less than +0.45 sds; 'low impact' by an effect size between +0.02 to less than +0.19 sds and below +0.02 sds 'very low or no impact'.

At the school level, the closest match is seen for the 2013/14 variables. This is unsurprising given that the propensity score matching was done using 2013/14 data. Table 2 provides confirmation that the matching process produced similar samples within the 2013/14 academic year.

A year later, in 2014/15, differences between the lead primary schools and their matched comparison are wider. This difference relates primarily to KS2 attainment; pupil context remains relatively stable and similar to 2013/14. Lead primary schools are observed to have higher levels of attainment compared with the matched sample - small differences in 2013/14 are observed to widen across most attainment measures in 2014/15.

A year earlier, in 2012/13, differences between the lead primary schools and their matched comparison are also wider than in 2013/14, and here, the differences are even greater than those observed in 2014/15. These differences also relate primarily to KS2 attainment, while pupil context remained relatively similar to 2013/14. Lead primary

[5] Excluding schools with incomplete school census details for the three years also resulted in dropping the 20 matches for these schools - leaving 39 lead primary schools and 780 matches. A further 62 contrast schools were dropped because of incomplete records. This means that the matches for the 39 lead primary schools are not universally 1:20. For 7 lead primary schools, a 1:20 match was retained but for the rest, the match ranged from 1:14 to 1:19.

[6] This includes a comparison table that ignores the issue of incomplete records - for reference.

[7] https://educationendowmentfoundation.org.uk/evidence/about-the-toolkits/about-the-toolkits/

schools are observed to have higher levels of attainment compared with the matched sample; large differences in 2012/13 are seen to narrow across most attainment measures in 2013/14.

Table 1: Testing the propensity score matching: School level comparison of 39 lead primary schools with their 718 matches

	2012/13			**2013/14**			**2014/15**		
	LPS	Match	d*	LPS	Match	d	LPS	Match	d
KS1 points	15.8	15.6	+0.15	15.7	15.7	+0.02	15.8	15.6	+0.14
KS2 points	29.9	29.2	+0.44	29.9	29.8	+0.04	30.2	29.7	+0.34
KS1-2 VA	100.8	100.3	+0.38	100.5	100.5	+0.01	100.6	100.4	+0.27
Maths VA	101.0	100.4	+0.41	100.6	100.6	0.00	100.7	100.4	+0.23
Schl size	382	323	+0.39	391	331	+0.38	398	338	+0.38
%female	49.3	49.0	+0.09	49.6	49.1	+0.18	49.2	49.3	-0.02
%FSM6	21.9	22.5	-0.04	21.0	22.6	-0.10	25.3	26.9	-0.09

d These are Cohen's d effect size statistics that capture the size and direction of any difference between the lead primary school and the matched contrast school samples. A positive effect size is shown in red - and indicates that, on average, lead primary schools had a **higher** value compared with the matched sample. A negative effect size is shown in blue, and indicates that, on average, lead primary schools had a **lower** value compared with the matched sample. Shaded cells indicate where, according to the EEF teaching and learning toolkit, effect sizes are moderate or higher.*

Notes: Schl = School; VA= value added; FSM = Free School Meals.

Example from Table 1:

For the 2012/13 KS1 points variable, the mean score for the lead primary schools was 15.8 points which compares with 15.6 points for the matched sample. The mean difference can be calculated as being +0.2 points in the original KS1 points score scale. The Cohens d effect size statistic converts this difference so that they are all measured in standard deviation units (+0.15 standard deviations). Whilst it is not appropriate to directly compare mean differences of different variables with different scales, the Cohen's d conversion allows this.

Testing the propensity score matches at the pupil level

Table 2 summarises the pupil level comparison for the two academic years under observation (see Appendix 5 for more detail on this comparison). For each of the KS1 and KS2 variables, Table 2 shows the mean values for the samples of pupils within the 39 lead primary schools and within the 718 contrast schools. The mean difference between the two samples is then shown as a Cohen's d effect size statistic.

As was found in the school level comparison, the differences at the pupil level are smallest in 2013/14. In 2014/15 the differences are seen to widen, but not to the same extent as seen with the school level comparison. This may reflect the more finely grained nature of the pupil level data compared with the aggregated school level comparison.

Table 2: Testing the propensity score matching: Pupil level comparison of lead primary schools with contrast schools, including 2013/14 and 2014/15 data

	2013/14			2014/15		
	LPS	Match	d*	LPS	Match	d
KS1 points	16.4	16.2	+0.08	16.6	16.3	+0.09
KS1 Maths points	16.7	16.6	+0.05	16.9	16.7	+0.07
%female	49.1%	48.9%	-	49.4%	48.7%	-
%FSM6	21.7%	22.0%	-	20.6%	21.2%	-
KS1 points (KS2 cohort)	15.7	15.8	-0.02	15.8	15.7	+0.03
KS2 points	29.9	29.8	+0.03	30.1	29.7	+0.09
KS2 Math FPS	5.1	5.0	0.00	5.1	5.0	+0.09
%female	49.9%	49.5%	-	48.4%	49.3%	-
%FSM6	26.0%	27.5%	-	26.8%	27.2%	-

** Cohen's d is not reported for binary outcomes.*
Notes: d = Cohen's d; FSM = Free School Meals; KS2 Maths FPS = KS2 Maths Fine Point Score.

A positive effect size is shown in red - and indicates that, on average, lead primary schools had a ***higher*** *value compared with the matched sample. A negative effect size is shown in blue, and indicates that, on average, lead primary schools had a* ***lower*** *value compared with the matched sample.*

Table 3: Testing the propensity score matching: Pupil level comparison of lead primary schools with contrast schools. KS1 maths points score, and KS2 maths fine points score

	n	Mean	s.d.	Min	Max	Low Q	Median	Up Q
2013/14								
KS1 Maths Points Score								
LPS	1,908	16.7	3.56	3	27	15	17	21
Contrast Schools	31,099	16.6	3.41	3	27	15	17	21
KS2 Maths Fine Points Score								
LPS	2,044	5.1	0.86	2.5	6.5	4.6	5.0	5.5
Contrast Schools	30,712	5.0	0.83	2.5	6.5	4.6	5.0	5.6
2014/15								
KS1 Maths Points Score								
LPS	1,975	16.9	3.45	3	21	15	17	21
Contrast Schools	31,427	16.7	3.39	3	27	15	17	21
KS2 Maths Fine Points Score								
LPS	1,922	5.1	0.84	2.5	6.5	4.6	5.0	5.5
Contrast Schools	31,515	5.0	0.83	2.5	6.5	4.5	5.0	5.5

Notes: s.d. = standard deviation; Low Q = lower quartile; Up Q = upper quartile.

As shown above, when testing the quality of the propensity matching at both school and pupil levels, the match between the lead primary and contrast school samples was very close in 2013/14. However, assuming that this close match will remain for successive cohorts of pupils and that any observed change can validly be attributed to the Mathematics Teacher Exchange is open to question, and therefore it is unclear whether any observed differences will relate to these (unmeasured) differences or to participation in the exchange. The limitations of the design are discussed in more detail in section 5.

Initial analyses comparing lead primary schools with contrast schools

Table 4 summarises the multilevel analyses of impact of the Mathematics Teacher Exchange on pupil level attainment at KS1 and KS2. The outcome variables for these analyses were KS1 maths points score and KS2 maths fine points score. Data from 2013/14 and 2014/15 are shown; the other three years will be included in future reports.

Table 4 also reports the estimated coefficient and standard errors for the dummy variable, which identifies pupils in the lead primary schools. A positive value for the coefficient indicates that, on average, pupils in lead primary schools attained higher than pupils in the matched contrast schools. The coefficient is then shown as a Cohen's d effect size statistic, with 95% Confidence Intervals (CI). Cohen's d is an effect size statistic that converts the difference between the lead primary and contrast samples into standard deviation units.

Figure 1 and Figure 2 illustrate the findings using error bars to represent Cohen's d effect size statistics with 95% confidence intervals. These 95% confidence intervals are indicative of sample size and allow for some measurement error in the KS1 and KS2 outcomes, but should not be read to indicate 'statistical significance'. This is due to the lack of randomness within the matched research design. The research design precludes being able to validly conclude that impact is 'statistically significant'[8]. Instead, evidence of impact will draw on multiple years of data to construct a descriptive picture of the relative attainment of pupils within LPS schools as compared with the matched contrast control sample of schools. This will be explained in more detail in the next section of the report.

For KS1 maths attainment, one simple bivariate multilevel model is shown that just included, as a predictor, the dummy variable identifying pupils in lead primary schools (= 1) from pupils in contrast schools (= 0). For KS2 maths attainment, estimates for two models are shown; first, a similar simple bivariate model and second, a model that included KS1 attainment as another predictor. In Figure 2, the two KS2 estimates are labelled 'simple' and 'value added' respectively, and in both these models the dependent variable was KS2 attainment. Further details on the multilevel analyses can be found in Appendix 1.

Whilst the final interpretation will need to wait for data from the other three years, in the first two years a greater change is observed at KS2 compared with KS1. At KS1, no change is observed in the effect size estimates and so here it must be concluded that there is no evidence so far that participation in the exchange resulted in an impact on pupil level KS1 attainment. At KS2, the attainment of pupils within lead primary schools is

[8] Statistically significant is a term used when a difference or an association is very unlikely (usually less than 5%) to have occurred as a result of random variation.

seen to increase somewhat compared with pupils in the contrast schools (shown by the increased Cohen's d effect size estimates). Whilst a small KS2 maths attainment advantage amongst the lead primary school sample is seen to increase, the size of this effect is very small, and from the EEF conversion indicates 'low impact'. Given the caution advised earlier in this document, and given that these analyses are at an interim stage of the evaluation; these findings should not yet be interpreted as evidence of the impact of the exchange on pupil level KS2 attainment. A more complete picture will come when the additional three years[9] have been obtained and analysed.

Table 4: Summarising the multilevel analyses into the impact of the Mathematics Teacher Exchange on pupil level attainment at KS1 and KS2, for 2013/14 and 2014/15 data.

		Coeff.	s.e.	d	95% CI for d
Key Stage 1 Maths Points Score					
2013/14		0.12	0.172	+0.04	(-0.06; +0.14)
2014/15		0.15	0.166	+0.04	(-0.05; +0.14)
Key Stage 2 Maths Fine Points Score					
2013/14:	LPS dummy only	0.01	0.044	+0.01	(-0.09; +0.09)
	LPS dummy & KS1 attainment	0.00	0.038	0.00	(-0.09; +0.09)
2014/15:	LPS dummy only	0.09	0.044	+0.10	(0.00; +0.21)
	LPS dummy & KS1 attainment	0.09	0.038	+0.07	(-0.02; +0.16)

Notes: Coeff = Coefficient; s.e. = standard error; d = Cohen’s d; CI = Confidence Intervals.

[9] The additional years to be included are 2012/13, 2015/16 and 2016/17.

Figure 1: Cohen's d effect size statistic estimates of the impact of school participation in the Mathematics Teacher Exchange on pupil level attainment at KS1. Data only for 2013/14, 2014/15 shown.

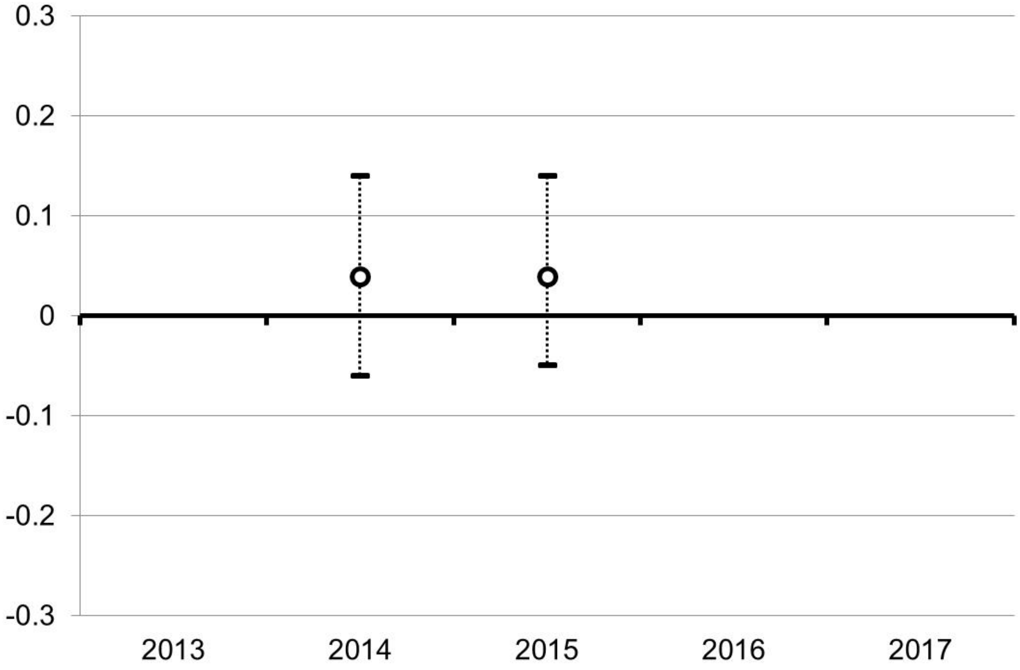
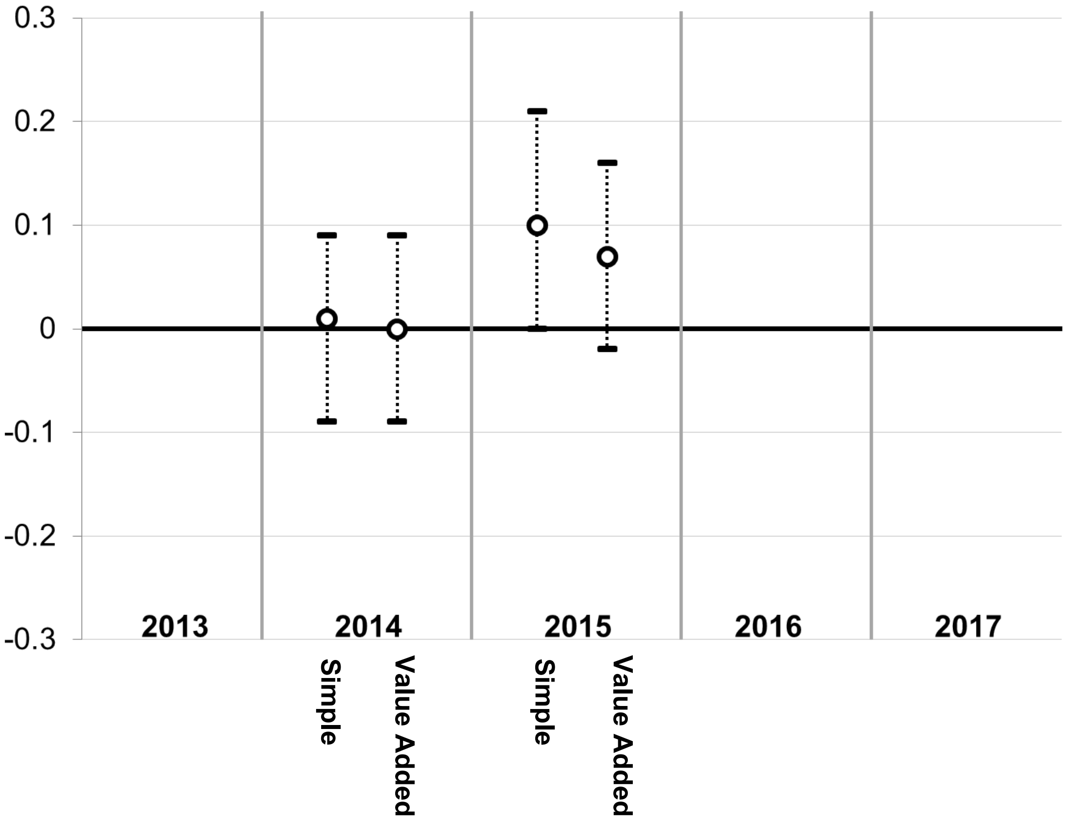


Figure 2: Cohen's d effect size statistic estimates of the impact of school participation in the Mathematics Teacher Exchange on pupil level attainment at KS2. Data only for 2013/14, 2014/15 shown.



Future analyses

Looking ahead towards the analyses that aim to assess pupil level impact of the Mathematics Teacher Exchange, mathematics attainment at KS1 and KS2 will represent key outcomes. The analyses presented in this document relate to impact across the whole sample of LPSs compared with the contrast control school sample at the school and pupil levels. Further 'whole school' analyses of KS1 and KS2 mathematics attainment will be undertaken for three additional academic years (2012/13; 2015/16 & 2016/17) at the school and pupil levels.

If these 'whole school' analyses reveal that the KS1/KS2 mathematics attainment of pupils within the LPS sample increases more sharply than attainment of pupils in the contrast control sample, this would provide evidence to support the conclusion that engagement with the Mathematics Teacher Exchange led to a positive impact on mathematics attainment.

In addition to the 'whole school' analyses, data from strand 1 of the evaluation will be drawn on in order to target sensitivity analyses towards a subsample of schools. For example, strand 1 data will be used to identify LPSs known to have sent KS2 teachers on the exchange. For these KS2 sensitivity analyses, the LPS sample (and their matched contrast control sample) would be limited to schools where this is known to be the case. Similarly, the KS1 sensitivity analyses would be limited to LPSs where it is known a KS1 teacher participated in the exchange.

If these targeted sensitivity analyses show an even greater difference in KS1 and KS2 mathematics attainment between the LPS and contrast control pupil sample, this would provide further evidence to support the conclusion that engagement with the Mathematics Teacher Exchange led to a positive impact on mathematics attainment.

One final consideration relates to the quality of the matching resulting from the PSM design. As reported in this document, the match is good / excellent at the school and pupil levels within the 2013/14 academic year, but problematic the year before. The poor match observed in 2012/13 serves to question the validity of attributing any observed differences in KS1 and KS2 mathematics attainment to participation in the Mathematics Teacher Exchange during the three 'outcome' years (2014/15 to 2016/17).

If the impact analyses conclude a positive impact of participation in the Mathematics Teacher Exchange on KS1 and KS2 mathematics attainment, additional scrutiny is planned by way of a 'stress test'. In brief, these tests would draw on additional academic years to try to improve on the match between the LPS and contrast control samples and if this process is successful the 'whole school' and targeted sensitivity analyses would be replicated using this re-constructed contrast control samples[10]. If a positive impact is also concluded from these follow-on 'stress-tests', this would strengthen the case for validly attributing this impact to participation in the Mathematics Teacher Exchange.

[10] Specific details of this are provides in the separate proposal document.

3. Pupil attitudes to mathematics

Attitudes to mathematics are an important outcome of any teaching programme, especially at primary level. Pupils' attitudes to mathematics are shaped in part by their early experience of education, and are carried with them into secondary education and beyond. There is growing evidence that negative attitudes, and mathematics anxiety, affect children as young as five years old (Ramirez et al., 2013). Mathematics anxiety is of particular concern as this is associated with difficulties in mathematics learning, and with negative effects on attainment (Vukovic et al. 2013). In addition, the Shanghai teaching approach is significantly different from English primary mathematics education, with notable features that potentially might influence pupil attitudes. The first of these features is less curriculum coverage in a single lesson, with a greater emphasis on depth (change in pace) combined potentially with greater interaction between teacher and pupils in a whole class context (change in tempo). The second is an emphasis on the whole class progressing together. These features influenced the survey design.

The effect of the exchange on pupil attitudes to mathematics will be tested using a pre-test/post-test design. Pupils in the year 6 cohort in the 48 participating schools were invited to complete a questionnaire containing items relating to attitudes and anxiety around mathematics teaching and learning during the summer term at the end of the 2015/16 academic year. At this time, no pupils in year 6 in participating schools had had significant experience of teaching for mastery relating to the exchange, and the majority will have had no experience of teaching influenced by the exchange at all. The survey has been repeated with the year 6 cohort in the academic year 2016/17 in order to see if attitudes to mathematics have changed in these schools. This design is limited, in the sense that there is no control group, and it is assumed that the two cohorts of pupils would be expected to have similar attitudes to mathematics. However, it will be of interest to see if there are any changes in attitudes that might be connected with changes in classroom practice following the exchange.

A survey was constructed that made use of tools from previous research. Twelve items were adapted from Brookstein et al. (2011), which addressed three factors: general affect towards mathematics; working together on mathematics and working independently on mathematics. Five items were adapted from Blackwell et al. (2007), which addressed implicit theories of mathematics ability. Six items were adapted from the most recent PISA survey on general attitudes to mathematics learning. Five items were developed specifically for this study, and related to pupils' attitudes towards the pace and tempo of mathematics lessons. All items were designed to be answered using a five-point Likert scale, and were presented to participating children one at a time, in random order, using the Qualtrics survey platform. The full set of questions used is included in this report as Appendix 5.

A total of 1191 surveys were completed online by participants in 36 schools. A principal components analysis (PCA)[11] was conducted, in order to identify the underlying constructs that the questionnaire was measuring. Initial tests (KMO=.918; Bartlett's test of sphericity was significant, $p<.0005$) suggested that the data were suitable for PCA. The analysis revealed three components, which can be labelled as: general affect towards mathematics; mathematics anxiety; and preference for working alone on mathematics (as opposed to working in a group).

Table 5 shows the questions relating to each component. Five questions from the survey do not appear in table 5 as they were not strongly associated with any of the three components. The 23 questions in Table 5 will be included in the survey to be carried out in 2016/17 to determine whether there has been any change associated with the exchange. Table 6 gives descriptive statistics for the three factors, after scores for each factor were standardised to a 0-10 scale. General affect towards mathematics was positively skewed (as one would hope), with a relatively high mean, and a long tail to the left. Both other factors had an approximately normal distribution. A full range of scores was evident for all three factors.

[11] Principal component analysis is a statistical technique used to explore a set of data and establish how a set of variables (in this case pupil's answers to the set of 28 questions) are associated with a set of underlying components (Field, Miles and Field, 2012).

Table 5: Questions generating three attitudes to mathematics components

General affect towards mathematics	1. Mathematics interests me 2. I look forward to mathematics lessons 3. I do mathematics because I enjoy it 4. I am interested in the things I learn in mathematics 5. I like mathematics 6. I believe I am the kind of person who can do well at mathematics 7. I enjoy hearing classmates' thoughts about mathematics 8. I think mathematics is important in life
Mathematics Anxiety	1. I get nervous when I am working on mathematics problems 2. I often worry that it will be difficult for me in mathematics lessons 3. When I see a mathematics problem, I am nervous 4. I often find mathematics questions too challenging in class 5. I sometimes feel like I am not keeping up in mathematics lessons 6. I worry that I will get low marks in mathematics tests 7. No matter how hard I try, I'll never be any good at mathematics 8. We often finish working on a topic before I have fully understood it 9. I sometimes feel nervous talking out loud in front of my classmates 10. I feel confident in my abilities to solve mathematics problems (reverse coding) 11. You have a certain amount of mathematical ability and you can't do much to change it 12. I am not keen to take part in discussions about mathematics
Preference for working alone	1. I enjoy working in groups better than alone in mathematics lessons (reverse coding) 2. I prefer working alone rather than in groups in mathematics lessons 3. I learn more about mathematics working on my own.

Table 6: Descriptive statistics for attitude factors

	Mean	**Median**	**Standard deviation**	**Minimum score**	**Maximum score**
General affect towards mathematics	6.99	7.22	1.94	0.83	10
Mathematics anxiety	3.80	3.75	1.81	0	10
Preference for working alone	4.29	4.17	1.92	0	10

4. Limitations of the research design and conclusions

The longitudinal analysis of national test data is due to continue until the 2017/18 school year. While this analysis will reveal useful information about the effects of participation in this intervention on children's attainment in mathematics, there are some important limitations of quasi-experimental methods, as opposed to true experiments, to bear in mind (see for example Robson, 1993, pp.87-108). A true experiment involves random assignment of cases to conditions. A key implication of this random allocation is that if, at the end of an intervention, some variable or set of variables can be observed to differ between conditions, there is some confidence that these differences have been *caused* by the difference between the conditions. In the case of the present research, it has not been possible to carry out random allocation of cases to conditions[12], and so it will not be possible to infer causation from any difference in attainment during the trial. Schools were chosen to participate in the intervention before the evaluation had been commissioned and were selected by Maths Hub lead schools on a variety of criteria, including levels of attainment achieved by their pupils in national tests in mathematics, on the basis of schools' ability to share good practice with other schools, previous engagement with mastery teaching and East Asian informed practices, and on existing and potential relationships to Maths Hubs. The selection criteria and processes varied across hubs. This means that, should a difference in pupils' attainment in mathematics be observed when the intervention schools and the contrast schools are compared, it will not be possible to be absolutely certain that this difference is due to the intervention itself. While the design of the impact evaluation is such that a large number of potential confounding variables will be controlled for (such as historic attainment, and proportions of children eligible for the pupil premium funding), it is not possible to control for all of them. For example, it is possible that any improvement may be due to schools' role in the leadership of the Maths Hubs, or to a particular interest or enthusiasm around mathematics teaching in participating schools. Data from the parallel process evaluation may help to ameliorate some of this concern, but there will still be a need to remember that findings from quasi-experiments are better interpreted as correlation rather than causation.

A second, potentially more complex, set of limitations concerns the definition of the intervention that is being tested in this study. Interpretation is always easiest when a narrow definition can be used to describe the intervention taking place. In the case of the present study, the intervention is multi-faceted and continually developing over time.

[12] Neither has it been possible to 'double-blind' this trial, but in this respect the present trial is the same as any other trial in an education context. The lack of blinding means that a potential Hawthorne effect (Adair, 1984) or similar should be considered when interpreting any findings. The Hawthorne effect is a phenomenon whereby participants modify their behaviour in response to their knowledge that they are being observed. In the case of the current evaluation, teachers in lead primary schools may change their teaching simply as a result of knowing they are participants in the evaluation, regardless of the particular intervention being tested.

Participating schools engaged in many different activities: visited Shanghai schools; hosted Shanghai teachers; have in some cases allowed participating teachers additional non-contact time to prepare to teach a mastery curriculum and to support other teachers in the school; have had support from the NCETM; have participated in a number of events around teaching for mastery; and will increasingly be supporting other schools to develop a mastery curriculum and pedagogy. Some teachers from these schools are participating in the new Primary Mathematics Teacher Mastery Specialist CPD programme. In parallel with the Mathematics Teacher Exchange, the NCETM is undertaking a textbook trial using textbooks informed by those in Singapore. Some of the lead primary schools are also involved in this and/or with other mastery programmes. Schools' understandings of 'mastery' are also developing through the project; these are not the same now as they were at the outset, and are likely to continue to develop as the project goes on. The NCETM, in supporting the project, have highlighted key aspects for development and so have focused schools' attention on these. The impact evaluation will not be able to differentiate between these various aspects of the intervention to any great extent.

Data from the process evaluation may suggest some aspects are more important than others in determining impact on pupil attainment, and may be able to inform some conclusions about how mastery curricula and pedagogy are instantiated in participating schools. However, the impact evaluation will not be able to deliver firm conclusions in all of these areas. This evaluation will be able to indicate whether, across all aspects of the intervention, participation in the intervention is associated with an increase in pupil mathematics attainment, but it will not be able to differentiate between effects due to those different aspects.

Conclusions

While there are some inherent limitations to the quasi-experimental design described in this report, Section 2 showed that there are reasons to believe that the baseline data represents a good basis for later comparison. There is good consistency in pupil level attainment data from year to year, and so the matched contrast group represents a good set of data for comparison as the project progresses.

References

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child development, 78*(1), 246-263.

Brookstein, A., Hegedus, S., Dalton, S., Moniz, R., & Tapper, J. (2014). *Measuring Student Attitude in Mathematics Classrooms*. Kaput Centre for Research and Innovation in STEM education. Retrieved from: http://www.kaputcenter.umassd.edu/downloads/products/technical_reports/tr4_studentattitude.pdf.

Field, A., Miles, J. & Field, Z. (2012). *Discovering Statistics Using R.* London: Sage.

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development, 14*(2), 187-202.

Robson, C. (1993). *Real world research: A resource for social sciences and practitioner researcher.* Oxford: Blackwell.

Vukovic, R. K., Kieffer, M. J., Bailey, S. P., & Harari, R. R. (2013). Mathematics anxiety in young children: Concurrent and longitudinal associations with mathematical performance. *Contemporary Educational Psychology, 38*(1), 1-10.

Appendix 1: Multilevel analyses of pupil level attainment at KS1 and KS2, using 2013/14 and 2014/15 data.

Key Stage 1 - Maths Points Score

Table 7: Listwise deletion of missing values (39 lead primary schools matched to 718 contrast schools)

	KS1 in 2013/14			
	Null Model		Including LPS Dummy	
	Coeff.	s.e.	Coeff.	s.e.
LPS Dummy	-	-	0.12	0.172
Constant	16.54	0.038	16.54	0.039
School Level Variance (R_{sch}^2)	0.76 (n/a)		0.76 (<0.1%)	
Pupil Level Variance (R_{pup}^2)	10.88 (n/a)		10.88 (<0.1%)	
Total Variance (R_{tot}^2)	11.64 (n/a)		11.64 (<0.1%)	
School level ICC	6.5%		6.5%	

	KS1 in 2014/15			
	Null Model		Including LPS Dummy	
	Coef	s.e.	Coef	s.e.
LPS Dummy	-	-	0.15	0.166
Constant	16.68	0.037	16.68	0.038
School Level Variance (R_{sch}^2)	0.70 (n/a)		0.70 (<0.1%)	
Pupil Level Variance (R_{pup}^2)	10.82 (n/a)		10.82 (<0.1%)	
Total Variance (R_{tot}^2)	11.52 (n/a)		11.52 (<0.1%)	
School level ICC	6.1%		6.1%	

Notes: Coeff. = coefficient; s.e. = standard error; ICC = Interclass Correlation.

Table 8: Raw data (47 lead primary schools matched to 940 contrast schools)

	KS1 in 2013/14			
	Null Model		Including LPS Dummy	
	Coeff.	s.e.	Coeff.	s.e.
LPS Dummy	-	-	0.14	0.165
Constant	16.45	0.036	16.44	0.037
School Level Variance (R_{sch}^2)	0.87 (n/a)		0.86 (0.1%)	
Pupil Level Variance (R_{pup}^2)	11.04 (n/a)		11.04 (<0.1%)	
Total Variance (R_{tot}^2)	11.90 (n/a)		11.90 (<0.1%)	
School level ICC	7.3%		7.3%	

	KS1 in 2014/15			
	Null Model		Including LPS Dummy	
	Coeff.	s.e.	Coeff.	s.e.
LPS Dummy	-	-	0.19	0.155
Constant	16.63	0.034	16.16	0.035
School Level Variance (R_{sch}^2)	0.74 (n/a)		0.74 (<0.1%)	
Pupil Level Variance (R_{pup}^2)	10.85 (n/a)		10.85 (<0.1%)	
Total Variance (R_{tot}^2)	11.59 (n/a)		11.59 (<0.1%)	
School level ICC	6.4%		6.4%	

Notes: Coeff. = coefficient; s.e. = standard error; ICC = Interclass Correlation.

Key Stage 2 - Maths Fine Points Score

Table 9: Listwise deletion of missing values (39 lead primary schools matched to 718 contrast schools)

	KS2 in 2013/14					
	Null Model		Including LPS		Including KS1	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
LPS Dummy	-	-	0.01	0.044	0.00	0.038
KS1 attainment (in 2009/10)	-	-	-	-	0.17	0.001
Constant	5.04	0.010	5.04	0.010	5.05	0.009
School Level Variance (R_{sch}^2)	0.06 (n/a)		0.06 (<0.1%)		0.05 (18.1%)	
Pupil Level Variance (R_{pup}^2)	0.64 (n/a)		0.64 (<0.1%)		0.30 (53.1%)	
Total Variance (R_{tot}^2)	0.70 (n/a)		0.70 (<0.1%)		0.35 (50.3%)	
School level ICC	8.0%		8.0%		13.1%	

	KS2 in 2014/15					
	Null Model		Including LPS		Including KS1	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
LPS Dummy	-	-	0.09	0.044	0.06	0.038
KS1 attainment (in 2009/10)	-	-	-	-	0.17	0.001
Constant	4.99	0.009	4.99	0.010	5.00	0.009
School Level Variance (R_{sch}^2)	0.06 (n/a)		0.06 (<0.1%)		0.05 (17.7%)	
Pupil Level Variance (R_{pup}^2)	0.64 (n/a)		0.64 (<0.1%)		0.31 (51.4%)	
Total Variance (R_{tot}^2)	0.69 (n/a)		0.69 (<0.1%)		0.36 (48.6%)	
School level ICC	8.2%		8.1%		13.1%	

Notes: Coeff. = coefficient; s.e. = standard error; ICC = Interclass Correlation.

Table 10: Raw data (47 lead primary schools matched to 940 contrast schools)

	KS2 in 2013/14					
	Null Model		Including LPS		Including KS1	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
LPS Dummy	-	-	0.02	0.046	0.00	0.036
KS1 attainment (in 2009/10)	-	-	-	-	0.17	0.001
Constant	5.00	0.010	5.00	0.010	5.03	0.008
School Level Variance (R_{sch}^2)	0.08 (n/a)		0.08 (<0.1%)		0.05 (2.6%)	
Pupil Level Variance (R_{pup}^2)	0.64 (n/a)		0.64 (<0.1%)		0.30 (53.1%)	
Total Variance (R_{tot}^2)	0.72 (n/a)		0.72 (<0.1%)		0.35 (51.1%)	
School level ICC	10.7%		10.7%		14.2%	

	KS2 in 2014/15					
	Null Model		Including LPS		Including KS1	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
LPS Dummy	-	-	0.09	0.044	0.07	0.035
KS1 attainment (in 2010/11)	-	-	-	-	0.17	0.001
Constant	4.96	0.010	4.96	0.010	4.98	0.008
School Level Variance (R_{sch}^2)	0.07 (n/a)		0.07 (<0.1%)		0.05 (2.4%)	
Pupil Level Variance (R_{pup}^2)	0.64 (n/a)		0.64 (<0.1%)		0.31 (51.2%)	
Total Variance (R_{tot}^2)	0.71 (n/a)		0.71 (<0.1%)		0.36 (49.4%)	
School level ICC	10.2%		10.2%		13.4%	

Notes: Coeff. = coefficient; s.e. = standard error; ICC = Interclass Correlation.

Converting the model coefficients into Cohen's d effect size statistics

An effect size is a statistical estimate of the strength of a phenomenon in standardised units. In the context of this research, the effect size provides an indication of the difference between the lead primary schools and matched contrast schools for the KS1 and KS2 maths attainment outcome measures. Whilst the model coefficients also provide an indication of this, the effect size standardises these coefficients so that they can be compared directly with each other and across other research studies. Without standardisation, the size of coefficient is dependent on the scale and units of the outcome measure and so it is not possible to compare these directly.

The effect size calculated in this report is Cohen's d. The table that follows provides some summary details on the primary outcome measure (KS2 attainment) and how the Cohen’s d estimates were calculated.

Table 11: Raw data (47 lead primary schools matched to 940 contrast schools)

	All Schools	LPS	Matched Schools
Mean	5.0	5.1	5.0
Standard deviation	0.835	0.856	0.833
n	32,756	30,712	2.044
Pooled standard deviation*	0.835		

The pooled standard deviation is calculated using the following formula:

$$s = \sqrt{\frac{(\mathrm{n}_1 - 1)s_1^2 + (\mathrm{n}_2 - 1)s_2^2}{\mathrm{n}_1 + \mathrm{n}_2 - 2}}$$

Where s is the pooled standard deviation, s_1 is the standard deviation for the lead primary school pupil sample and s_2 is the standard deviation for the contrast school pupil sample. n_1 is the number of pupils in the lead primary school sample and n_2 is the number of participants in the contrast school pupil sample.

The estimated lead primary school coefficient for the 2013/14 KS2 Maths Fine Points Score model is then divided by this pooled standard deviation to convert it into a Cohen's d effect size. Similarly, the upper and lower 95% confidence intervals for the estimated lead primary school coefficient is divided by the pooled standard deviation to provide upper and lower 95% confidence intervals for the Cohen's d effect size statistic.

Appendix 2: Propensity Score Matching

Step 1: Building the model

To generate propensity scores a probability model was constructed that used the '% attaining level 5 or above in KS2 maths test' variable to create a threshold (binary) outcome measure - whether the proportion of pupils attaining level 5 or higher in each school was the median or above (more or equal to 42% = 1) or not (less than 42% = 0). Table 13 summarises the logistic regression model used to generate the propensity scores. The purpose of the model is to generate the scores and not for other interpretation.

Step 2: Generating the propensity scores

The model generated predicted probabilities for each school - a prediction of how probable it was that the proportion of pupils in a school reaching level 5 or higher at KS2 maths was 42% or greater that is based upon the detail provided by the 10 explanatory variables. These predicted probabilities take a value between zero and one and are known as 'propensity scores'. It was these scores that were used to match the lead primary schools to potential contrast schools.

Step 3: Using the propensity scores to match the lead primary schools 1:20

The propensity scores are rank ordered, the lead primary schools are found within this ranked list and schools with similar propensity scores to each lead primary school were selected into the matched sample. Each lead primary school was matched to the 20 'closest' schools based on these propensity scores. This was done for 47 of the 48 lead primary schools[13]. For the three junior schools involved in the study, the 20 closest junior schools were selected into the three respective contrast school subsamples. This resulted in the 47 lead primary schools being matched to 940 contrast schools.

Step 4: Testing the matching

The matches was then statistically scrutinised or 'tested'. This was done first by examining the propensity scores. Table 12 below summarises the propensity scores for the 47 lead primary schools and the 940 matched contrast schools. Additionally, Table 12 highlights the three lead primary schools with the 'worst' matches in terms of the propensity scores. The matching was tested further by comparing the lead primary school and contrast samples in terms of attainment and other factors which are reported in the body of this report.

As summarised in Table 12, analyses of the propensity scores revealed an overall close match when comparing the distribution of the 47 lead primary schools with all 940 matches. Analyses at the lead primary school level compared each of the 47 lead

[13] Colmore Infant School had no KS2 attainment detail and is the only infant school in the study.

primary schools with their 20 matches and found that the match held very well here. Matches can be assessed by comparing the propensity score of a lead primary school with the range of scores for the 20 matches in terms of the mean and extreme values. For the 47 lead primary schools the three schools with the largest distance between their propensity score and a comparison contrast school were the three junior schools in the study. This makes sense. As noted earlier, selecting the 20 matches for these schools involved only selecting the closest junior schools. This meant that a wider range of propensity scores were used for these three instances compared with the remaining 44 schools. However, even here, the match looks good with the 'worst' match (Junior School 1) having a propensity score of 0.963 compared with the 20 matches with a mean score of 0.968 and a range between 0.963 and 0.973 (max calliper difference ~ 0.010). In summary, in terms of propensity scores, the two samples compare very closely.

Table 12: Testing the matches: Statistical Summary of Propensity Scores.

	LPS Sample	Matched Sample
n	47	940
Mean (s.d.)	0.68 (0.382)	0.68 (0.378)
Max	1.000	1.000
Min	0.001	0.001
Three 'worst' matches	LPS (n=1)	Matches (n=20)
Junior School 1	0.963	Mean (sd): 0.968 (0.0027) Min: 0.963; Max: 0.973 Max Caliper = 0.010
Junior School 2	0.953	Mean(sd): 0.949 (0.0024) Min: 0.945; Max: 0.954 Max Caliper = 0.008
Junior School 2	0.962	Mean(sd): 0.960 (0.0017) Min: 0.957; Max: 0.963 Max Caliper = 0.005

Notes: s.d. = standard deviation.

Table 13: Binary Logistic Regression Model used to generate propensity scores. Dependent Variable: Whether the proportion of KS2 pupils attaining level 5 or higher in mathematics is less than 42% (median) or greater (0 or 1).

	B (s.e.)	Wald	Exp(B) (95% CIs)
Size of school	0.0003 (0.0002)	2.00	1.00 (1.00-1.00)
%female	0.0059 (0.0099)	0.36	1.00 (0.99-1.03)
%SEN of School Action	0.0079 (0.0083)	0.89	1.01 (0.99-1.02)
%EAL	-0.0063 (0.0017)	13.90	0.99 (0.99-1.00)
%6FSM	-0.0163 (0.0040)	16.40	0.98 (0.98 - 0.99)
IDACI Score	0.3131 (0.3006)	1.08	1.36 (0.76 - 2.47)
KS1 Points Score	2.1912 (0.0536)	1670.43	8.95 (8.05-9.94)
KS1 to 2 Value Added Score (Mathematics)	2.5003 (0.0633)	1561.87	12.19 (10.77-13.79)
% boys ...at least 2 levels prog in maths	-0.0269 (0.0050)	29.25	0.97 (0.96-0.98)
% girls ...at least 2 levels prog in maths	-0.0261 (0.0046)	31.15	0.97 (0.97-0.98)
Constant	-0.2066 (0.0337)	-	-
Initial -2 Log Likelihood	14353.7		
Model Chi-Square	6343.9		
Residual -2 Log Likelihood	8009.7		
Pseudo R-square (McFaddens)	44.2%		

Notes: s.e. = standard error; CI = Confidence Intervals; SEN = Special Education Needs; EAL = English as an Addiitonal Language; FSM = Free School Meal; IDACI = Income Deprivation Affecting Children Index score.

Appendix 3: Testing the matches at the school level

Listwise deletion of missing values

The following eight lead primary schools (and their matches) were dropped from these analyses because of incomplete records: 1.0310 (St Augustine's RC Primary School); 1.0320 (St Gregory's Catholic Prim School); 1.0420 (Histon & Impington Junior School); 1.0710 (Hillside Primary and Nursery School); 1.0820 (Norwich Primary Academy); 1.1010 (Outwood Prim Acad Lofthouse); 2.2210 (Bude Park Primary School) & 2.2610 (St Thomas of Canterbury School (Acad). This reduces the matched sample to 39 lead primary schools with complete details 2012/13 to 2014/15.

Table 14: Comparison of LPS with matched contrast schools, at school level

	2012/13			**2013/14**			**2014/15**		
	LPS Sample Mean (sd)	**Matched Sample Mean (sd)**	**Cohen's d**	**LPS Sample**	**Matched Sample**	**Cohen's d**	**LPS Sample**	**Matched Sample**	**Cohen's d**
KS1 points (KS2 cohort)	15.8 (1.33)	15.6 (1.30)	+0.15	15.7 (1.45)	15.7 (1.23)	+0.02	15.8 (1.40);	15.6 (1.17)	+0.14
KS2 points	29.9 (1.62)	29.2 (1.52)	+0.44	29.9 (1.56)	29.8 (1.33)	+0.04	30.2 (1.55)	29.7 (1.35)	+0.34
KS1 to KS2 VA - overall	100.8 (1.10)	100.3 (1.10)	+0.38	100.5 (1.02)	100.5 (0.99)	+0.01	100.6 (0.94)	100.4 (1.00);	+0.27
KS1 to KS2 VA - maths	101.0 (1.37)	100.4 (1.39)	+0.41	100.6 (1.28)	100.6 (1.25)	0.00	100.7 (1.24);	100.4 (1.27)	+0.23
Size of School	382 (200)	323 (147)	+0.39	391 (201)	331 (153)	+0.38	398 (208)	338 (158)	+0.38
% Female	49.3 (2.50)	49.0 (3.27)	+0.09	49.6 (3.00)	49.1 (3.05)	+0.18	49.2 (7.26)	49.3 (8.58)	-0.02
% FSM (Last 6 years)	21.9 (16.12)	22.5 (16.77)	-0.04	21.0 (15.78)	22.6 (16.47)	-0.10	25.3 (17.29)	26.9 (18.74)	-0.09

Notes: s.d. = standard deviation; VA = Value Added; FSM = Free School Meal.

Appendix 4: Testing the matches at the pupil level

Table 15: Comparison of LPS with matched contrast schools, at pupil level

	2013/14			2014/15		
	LPS Sample	**Matched Sample**	**Cohen's d**	**LPS Sample**	**Matched Sample**	**Cohen's d**
-						
KS1 Points	16.3 (3.26)	16.1 (3.22)	+0.08	16.5 (3.21)	16.3 (3.19)	+0.09
KS1 Maths Points	16.6 (3.54)	16.5 (3.45)	+0.05	16.9 (3.45)	16.6 (3.40)	+0.07
Gender (% Female)	48.7%	48.7%	-	49.6%	48.8%	-
% FSM (Last 6 years)	23.1%	23.6%	-	21.9%	22.6%	-
-						
KS1 Points (for KS2 cohort)	15.7 (3.66)	15.7 (3.55)	-0.01	15.7 (3.50)	15.7 (3.53)	+0.01
KS2 Points	29.8 (4.33)	29.6 (4.28)	+0.03	30.0 (4.23)	29.6 (4.26)	+0.11
KS2 Maths Fine Points	5.0 (0.86)	5.0 (0.85)	+0.03	5.1 (0.84)	5.0 (0.84)	+0.10
Gender (% Female)	49.9%	49.5%	-	48.4%	49.3%	-
% FSM (Last 6 years)	26.0%	27.5%	-	26.8%	27.2%	-

Notes: FSM = Free School Meal.

Appendix 5: Set of questions used for the pupil attitude survey

Question order was randomised by the survey software in order to control for order effects. All questions were answered using a 5-point Likert scale (1 = Strongly Disagree; 5 = Strongly Agree).

1	I think mathematics is important in life
2	I like mathematics
3	I enjoy hearing my classmates' thoughts and ideas about mathematics
4	Mathematics interests me
5	When I see a mathematics problem, I am nervous
6	I sometimes feel nervous talking out loud in front of my classmates
7	I prefer working alone rather than in groups when doing mathematics
8	I learn more about mathematics from talking to my classmates than from talking to my teacher
9	I am not keen to take part in discussions about mathematics
10	I enjoy working in groups better than alone in mathematics lessons
11	I learn more about mathematics working on my own
12	I feel confident in my abilities to solve mathematics problems
13	I believe I am the kind of person who can do well at mathematics
14	No matter how hard I try, I'll never be good at mathematics
15	You have a certain amount of mathematical ability and you can't do much to change it
16	No matter who you are, you can change your mathematical ability a lot
17	Whether or not I do well in mathematics is completely up to me
18	Mathematics lessons go at about the right speed for me
19	I sometimes feel like I am not keeping up in mathematics lessons
20	I often find mathematics questions too challenging in class
21	Sometimes mathematics lessons go too slowly for me
22	We often finish working on a topic in mathematics before I have fully understood it
23	I often worry that it will be difficult for me in mathematics lessons
24	I get nervous when I am working on mathematics problems
25	I worry that I will get low marks in mathematics tests
26	I look forward to mathematics lessons
27	I do mathematics because I enjoy it
28	I am interested in the things I learn in mathematics

Department for Education

Reference: DFE-RR777

ISBN: 978-1-78105-852-7

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

Any enquiries regarding this publication should be sent to us at: Adam.thompson@education.gov.uk or www.education.gov.uk/contactus

This document is available for download at www.gov.uk/government/publications