



Does continued participation in STEM enrichment and enhancement activities affect school maths attainment?

Pallavi Amitava Banerjee

To cite this article: Pallavi Amitava Banerjee (2017) Does continued participation in STEM enrichment and enhancement activities affect school maths attainment?, Oxford Review of Education, 43:1, 1-18, DOI: [10.1080/03054985.2016.1235031](https://doi.org/10.1080/03054985.2016.1235031)

To link to this article: <http://dx.doi.org/10.1080/03054985.2016.1235031>



Published online: 28 Nov 2016.



Submit your article to this journal [↗](#)



Article views: 314



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

Does continued participation in STEM enrichment and enhancement activities affect school maths attainment?

Pallavi Amitava Banerjee

University of Exeter, UK

ABSTRACT

Science, technology, engineering, and mathematics (STEM) skills are very valuable for economic growth. However, the number of young people pursuing STEM learning trajectories in the United Kingdom was lower than the predicted demand during the last decade. Several STEM enrichment and enhancement activities were thus funded by the government, private, and charitable organisations to improve understanding of and raise pupil interest in these subjects. One possible way of measuring the impact of these activities in supporting pupil understanding of maths was to track the proportion of young people obtaining a 'good' grade in standardised national tests such as the GCSEs. Attainment is of course only one possible outcome of education but certainly a very important one because students are more likely to continue studying subjects in which they score higher. This makes maths attainment even more important as it is a pre-requisite for admission to STEM degree courses. This longitudinal study makes use of the National Pupil Database to assess the impact of these schemes on maths attainment of participating schools. Following up 300 intervention schools for five years the study shows the intervention group did not do any better than the comparator. The paper suggests further directions for research and offers recommendations for practice.

KEYWORDS

Mathematics; attainment; school; enrichment; enhancement; STEM

Introduction

During the last decade, science, technology, engineering, and maths (STEM) education was increasingly seen as a key contributor in providing a highly skilled workforce for the continued economic development of the United Kingdom (SEMTA, 2010). Research commissioned by the Science Council in 2010 suggested that by 2017 over 58% of all new jobs will require STEM skills (Garnham, 2011). This demand for STEM skills was over and above the national demand at that time (but see Smith & Gorard, 2011; UKCES, 2015). A good predictor of the ease of meeting this future demand was the number of students studying STEM subjects post-16 (Ofsted, 2011a). It was shown that there were likely to be fewer STEM graduates than had been projected as a future requirement (Ofsted, 2011b). This was worrying as it could thwart the UK's economic progress since STEM skilled workers play a crucial part in scientific

advancement, and in keeping up with competitive standards as global leaders (Royal Society, 2011).

Some basic STEM skills can be fostered by studying a single STEM subject for level 3 qualifications, however progression to STEM careers can only be met with an undergraduate degree in a related subject. SCORE (2010) concluded in addition to other course-specific criteria that admission tutors highly value maths qualifications when offering places into any STEM undergraduate degree programme. This is because competence in maths is a very important element in preparing young people for these degrees (Hodgen, Pepper, Sturman, & Ruddock, 2010). An evidence-based forecast of whether this predicted STEM demand could be met was reflected in Ofsted reports (2011a, 2011b). The reports showed student progression rates to specialist science and maths courses were very low. Although enrolment to these courses in colleges had improved in recent years, it was not sufficient. Amongst those who opted to study STEM subjects post-16 many of them were ineligible for pursuing STEM undergraduate courses as they had a single level 3 STEM qualification which often excluded maths (Hodgen, Marks, & Pepper, 2013; Ofsted, 2011a, 2011b).

A difficulty was thus foreseen in meeting the anticipated STEM skills demand (Behr, 2011). Factors affecting student subject choices are deeply embedded in a specific social and educational framework. These cannot be easily shifted and there are no ready policy measures to change the picture (Banerjee, 2016b; Rodeiro, 2007). To combat this problem, a wide range of schemes was designed and implemented (POSTNOTE, 2011) with three main objectives: (a) to develop a better understanding of science and maths; (b) to link science and maths as done in the classroom to STEM done in the real world; and (c) to break the myth held by young people—'STEM is meant only for the brainy' (DeWitt, Archer, & Osborne, 2013), thereby, encouraging pupil participation in STEM subjects. Sustained efforts have been made since at least the start of 2000; how successful have these been?

Policy background

The Trends in International Mathematics and Science Study (TIMSS) uses internationally comparative assessments dedicated to improving teaching and learning in maths and science for students around the world. This study is carried out every four years at the fourth and eighth grades. Detailed information about maths and science curriculum coverage and implementation, as well as teacher preparation, resource availability, and the use of technology, is collected. TIMSS outputs about trends in maths and science achievement over time inform educational policy in the participating countries. The research reports showed English pupils' actual achievement in maths had improved between 1995 and 2007 and had plateaued between 2007 and 2011. TIMSS also showed students with positive attitudes towards these subjects attained higher (Sturman, Burge, Cook, & Weaving, 2012; Sturman et al., 2008). This was perhaps the justification for the plateau. The proportion of young people in the UK with a positive attitude towards maths was around 10 percentage points below the international average in 2007 (UK Parliament, 2011).

Given the growing demand and limited supply of STEM skilled people in the UK this was worrying as pupil aspirations by the age of 14 give a good indication of their willingness to continue with STEM when they get older (Archer et al., 2012; Ing & Nylund-Gibson, 2013). Interventions early on in the academic life of a pupil were thus likely to be more effective.

Policymakers thus targeted STEM attrition in schools, with the rationale of retaining more students in science and maths in secondary school—supposedly a low-cost, fast, and efficient way of producing the STEM professionals the nation would need. Several funded STEM enhancement and enrichment activities were run to increase pupil confidence and competence in science and maths. These approaches hoped to redress educational outcomes by increasing knowledge and improving understanding. Attainment is of course only one possible outcome of education but certainly a very important one. First, students achieving higher are more likely to continue studying these subjects. Second, a good level 3 maths qualification is essential for admission to most STEM undergraduate degree courses.

Low attainment in maths is common among some social groups. Several contextual indicators are now known to be linked to attainment. Schools with a higher percentage of disadvantaged pupils for example do not always rank very high in performance tables. It is important to address both of these issues, efforts towards which have been done by the introduction of the brilliant initiative evaluated here—STEM enrichment and enhancement activities. Beginning in 2000 the initiatives, schemes, and budget have all increased considerably (DCSF, 2004). However, major studies or surveys of participating schools and students, looking at the long term impact these schemes have in improving achievement in STEM subjects, are relatively scarce (Wynarczyk, 2008; Wynarczyk & Hale, 2009). There have of course been short term programme evaluations which capture pupils' and teachers' experiences during and immediately after the programme but these have little or nothing to do with solid evidence about whether students' lives were changed. There is a growing need to evaluate what works to be able to build on the best ones to achieve better results for the same amount of money (DfE, 2014; DfES, 2006). Analysing large scale secondary datasets, this new research evaluates the impact some of these enrichment and enhancement activities have had on raising school attainment and narrowing the achievement gap. In this paper, achievement gap is defined as the observed, persistent disparity of educational measures between the performances of groups of students, especially groups defined by socio-economic status (SES) as considered by the Department for Education UK.

The intervention—STEM enrichment and enhancement activities included in the study

STEM enrichment and enhancement providing organisations deliver several activities throughout the year. These are in the form of hands-on activities, engaging and fun sessions, inspirational talks delivered by STEM Ambassadors and people successful in STEM careers, maths challenges and fun sessions, and often as faculty mentoring programmes run by higher education institutions. All activities considered here were delivered as after-school clubs, competitions, or out-reach programmes. The common elements linking these programmes were the objectives and practical element involving active participation of students in some kind of set-up beyond in-school teaching. The same activity providers delivered a variety of activities under a common underlying theme to engage students. The providers had a list of registered schools whom they catered to year after year (Appendix). Further, during initial discussions with the heads or point of contact of these organisations one of the goals all of these providers hoped to achieve was to help young people develop a better understanding of maths and to help them attain higher. Several organisations run such programmes in England; the criteria used for screening activity providers for this research project were:

- (1) The activities were designed to improve the understanding of students in maths.
- (2) The schemes were delivered from the beginning of KS3 to the end of KS4 in England.
- (3) All chosen initiatives claimed to improve educational outcomes.
- (4) None of these activities were one-day events; they were delivered at different times across the year in the form of advanced follow-up versions of the previous event.
- (5) All of these programmes reported data which could be used to estimate an effect size.
- (6) Outcome effectiveness of these interventions could be measured in terms of GCSE performances.
- (7) All chosen programmes had sustained participation of schools.
- (8) Programme leaders were willing to co-operate and share data for the research project.

Methods

Have school performances been affected as a result of their pupil engagement in STEM initiatives? Do participating schools have a higher share of young people obtaining higher grades in maths? These questions are addressed via a quasi-experimental study—‘quasi’ because the researcher was not delivering any intervention, and the cases were not allocated randomly. Such evaluations can provide information about naturally occurring events, behaviour, attitudes, or other characteristics of a particular group. Also, these studies are helpful in demonstrating associations, for example here between STEM initiatives and maths attainment, without disturbing the educational setting or introducing a bias.

The educational performances of about a 1000 identified intervention schools were compared with all other schools year-wise in a snapshot. From these 1000 schools, 300 schools were identified which enrolled the 2007/08 cohort of their Year 7 pupils in STEM schemes every year from the beginning of KS3 till the end of KS4 when this cohort took GCSEs. GCSE maths results were the outcome measure for assessing the impact of STEM initiatives.

Data

Based on the screening criteria, 10 activity providers (anonymised) in England were a part of this study and provided names of participating schools (Table 1). Eight of these were

Table 1. Number of participating schools registered with STEM activity providers.

| STEM activity provider | Activity zone | 2007/08 | 2008/09 | 2009/10 | 2010/11 | 2011/12 |
|--|---------------|---------|---------|---------|---------|---------|
| 1 | North East | 68 | 69 | 69 | 72 | 73 |
| 2 | North East | 56 | 80 | 45 | 65 | 71 |
| 3 | Dudley | 20 | 20 | 20 | 20 | 22 |
| 4 | London | NA* | NA | NA | NA | 23 |
| 5 | London | NA | NA | NA | 20 | 57 |
| 6 | London | NA | 48 | 58 | 276 | 56 |
| 7 | England | NA | NA | NA | NA | 3 |
| 8 | England | NA | 105 | 135 | 106 | 101 |
| 9 | Tyne & Wear | NA | 34 | 48 | 28 | 24 |
| 10 | England | 293 | 338 | 335 | 348 | 348 |
| Total participating schools ⁺ | | 421 | 633 | 653 | 852 | 696 |

Notes: *NA=not available; ⁺Duplicates excluded.

government organisations, one an educational charity, and one received public funding. Programme delivery to 11–14 year olds at these organisations was observed to get an idea of what the actual activity entailed. The providers shared details of the programmes and instruction materials with the researcher. For a more detailed discussion of the research protocol please see Banerjee (in press-a).

The study made use of school and pupil level annual census datasets and performance tables from the National Pupil Database. The National Pupil Database (NPD) contains detailed information about individual pupils in schools and colleges in England. In a longitudinal study such as this, it is possible to track the learning trajectory of a child through the NPD even if the student changes schools as long as all of the schools attended are in England. This reduces attrition.

Data made available for this research project by the NPD were the most up-to-date at that point in time and covered from 2007/08 to 2011/12 for GCSEs and 2013/14 for A-levels. The standard KS4 extract requested for this project combined KS4 attainment with prior attainment at KS1, KS2, and KS3, and spring census data from the current academic year (and previous six years), undertaken by schools. All school and pupil level data used for analysis here are amended data. All special schools, pupil referral units, and independent schools were excluded from the study. State maintained schools included were academies, city technology colleges, voluntary aided, voluntary controlled, and foundation schools.

Repeated cross-sectional design—school-level data

A repeated cross-sectional research design was used to assess the impact of STEM initiatives on school maths performances. Each provider had a set of schools registered for each academic year which meant a range of STEM enrichment activities were delivered to all pupils in these schools throughout the academic year. However, schools were free to continue or discontinue their registration with the provider for the next academic year. Some schools were registered with one, two, three, and at times four activity providers. All schools registered with at least one STEM activity provider for each academic year were shortlisted. This group was termed as the intervention group. The comparator group was the population of all other secondary schools excluding special schools and those schools for which attainment data were not available (such as independent schools). Thus for this design the comparison between intervention and comparator was year-wise. This is because there were almost always new schools joining each academic year and a proportion who decided to discontinue. Schools were included in the intervention group for the years they were registered for this snap-shot. Through correlation techniques and comparison of population means, maths attainment figures were compared between groups year-wise from 2007 to 2012.

Longitudinal design—school-level data

Three-hundred state maintained secondary schools were identified from the intervention group in the repeated cross-section design dataset. The Year 7 cohort of 2007 of all of these schools had continuously participated in STEM activities every year beginning from KS3 until the end of KS4 when they took their GCSEs. The number of organisations the schools registered with varied from a minimum of one to a maximum of four out of the 10 organisations being considered in the study. This meant students from these longitudinal intervention schools were exposed to an age-appropriate advanced version of STEM activities every following year. The final dataset carried details of school census, attainment data, and

participation in STEM schemes. Mean school GCSE performances for the longitudinal intervention group were then mapped before and after intervention in 2007 and 2012, respectively. Population means, correlation coefficients, and achievement gaps were then estimated.

During these five years when the longitudinal intervention was delivered, some schools closed and some new schools were opened. A school was included in the longitudinal intervention group only if it participated consistently each year. Thus if a school participated for some years but closed even during the last year of data collection it was excluded. Similarly, if a school just opened during the second year of data collection and participated every following year it was still excluded. Some schools converted into academies, and the new URN was checked in NPD records and Edubase to ascertain it was the same school. All such schools were included only if they participated each year.

During face-to-face meetings some activity providers claimed to know certain schools which had never enrolled with STEM activity providers but these names could not be disclosed by them due to data protection reasons. Thus, in the absence of clearly matched data the national performance was considered. All secondary schools in England following the National Curriculum whose school results were available from performance tables published by the NPD, excluding the intervention schools, formed the comparator for this study. These clearly included schools not involved in any STEM enrichment activities. It also includes some schools that were participating in interventions. This could dampen any effect size, but was the only feasible comparison. Trying to match schools could be worse, since the matched comparator schools might also be unknown treatment schools.

Study design using pupil level data

For the multiple regression analysis, students from these 300 longitudinal intervention schools were followed from the beginning of Year 7 until the end of KS4 (Table 2). If a child moved school and new school details were available from NPD, the student was included in the intervention group only if both old and new schools were known intervention schools. Students who dropped out of education or left the country were not included as their records were not available from NPD. Similarly, new students who joined the cohort any time after the first year of intervention were also excluded even if they were at an intervention school.

Case selection procedures were based on actual participation as far as possible. This was important, despite causing attrition, to ensure a direct effect of longitudinal interventions could be seen in pupil attainment. It is expected that there might have been a few instances when students were absent on the actual day of intervention delivery, though it was not possible to check these cases and is one of the known limitations of this study. However, the huge sample size of nearly 80,000 intervention pupils reduces these considerations.

Table 2. Number of students in sub-groups—STEM activity participation.

| Sub-groups | End of KS4 | |
|---------------------------------|------------|-------------|
| | Numbers | Percentages |
| Comparator | 555,295 | 88 |
| Longitudinal intervention group | 76,462 | 12 |
| Total | 631,757 | 100 |

Indicators used in the study

Maths attainment

School-level data obtained from NPD were merged with Key Stage 4 maths performance indicators, notably the percentage of pupils achieving 5+ A*–C grades including both English and Maths GCSE. This variable was chosen for the study as it was available for all academic years being considered, 2006/07–2011/12. Using the same indicators for all years reduces the chances of error which could possibly arise in matching of variables for comparison in a longitudinal evaluation. A grade of C or above in GCSE maths was thus considered a ‘success’. Figure 1 shows the clear relationship between percentage of pupils eligible for free school meals (FSM) and the percentage gaining C or above in maths in schools. Schools with a higher percentage of FSM eligible pupils had a lower percentage of pupils achieving A*–C in maths, and vice versa. A similar trend was observed each year.

Eligibility for free school meals

One of the indicators used as a measure of pupil’s poverty is eligibility for free school meals (FSM) (Gorard, 2012; Hobbs & Vignoles, 2010; Shuttleworth, 1995). As stated in the guidelines issued by the Department for Education UK, pupil eligibility for FSM is assessed based on several criteria, all of which are self-declared by parents/carers when claiming free school meals. Some examples of these contextual indicators include parent or guardian in receipt of income support or jobseekers allowance, support under Part VI of the Immigration and Asylum Act 1999, the guaranteed element of State Pension Credit, Child Tax Credit (provided they are not also entitled to Working Tax Credit and have an annual gross income of no more than £16,190), working Tax Credit run-on—paid for four weeks after they stop qualifying for

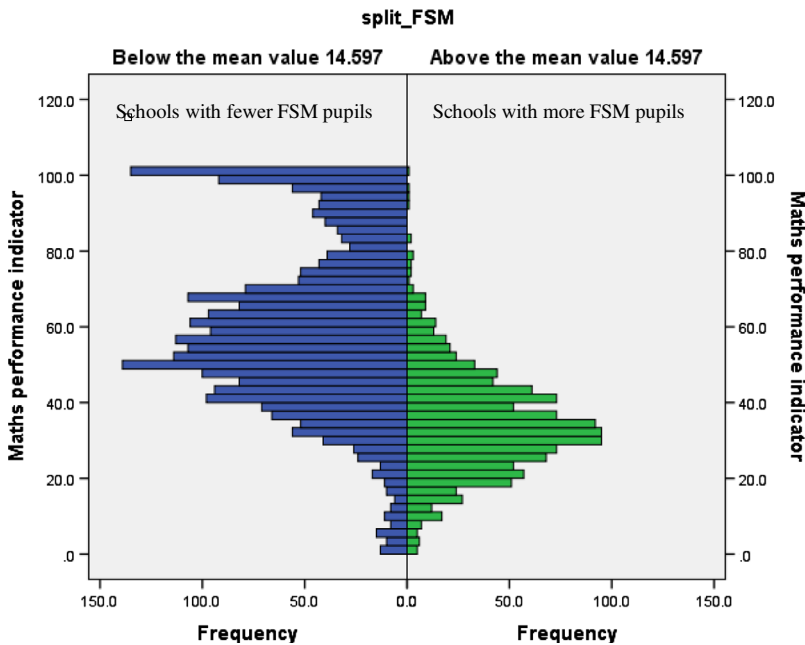


Figure 1. Percentage pupils achieving A*–C in maths by % FSM pupils in school. Source: National Pupil Database, 2009–2010.

Working Tax Credit or Universal Credit. All of these are indicative of a lower socio-economic status and hence FSM eligibility was used as a disadvantage measure here.

Other indicators used in the study

Other contextual indicators linked to academic performances were considered in the study for regression analyses. These were gender, ethnicity, language group, and prior attainment in maths at the end of KS2. This is because the interventions were delivered from the beginning of KS3. Only state maintained secondary schools were considered in this study and the proportion of pupils with a statement of education needs (SEN) in this school was not more than 1–2%. SEN was also included as an explanatory variable.

Data analysis

The study explored the possible impact of STEM initiatives on school GCSE maths performances. The means of percentage of pupils achieving 5+ A*–C grades including English and Maths in schools were compared. The percentage point difference in achievement between intervention and comparator groups was assessed. The achievement gap was estimated using Newbould and Gray's approach (Banerjee, *in press-a*; Gorard, 1999). They define the achievement gap as the difference between performances of intervention and comparator relative to the performance of all entries, minus the entry gap. Here entry gap is the difference in number of entries from the intervention and comparator divided by the total number of all entries for the chosen measure. Since these calculations are from snap-shot data, the entry gap has been taken as zero. Thus, for calculating the achievement gap the following formula was used here:

Achievement gap for cross - sectional design

$$= \frac{\text{Mean attainment}_{\text{intervention}} - \text{Mean attainment}_{\text{comparator}}}{\text{Mean attainment}_{\text{population}}}$$

Pupils belonging to families with a lower socio-economic status have been shown to perform not so well academically. Thus school performance is negatively correlated to percentage of FSM eligible pupils in school. The analysis explores if exposure to STEM activities can break the link between SES and maths performances. Pearson's *R* correlation coefficient was used to study the correlation between school maths performances and percentage of pupils eligible for free school meals (FSM) in the intervention group and the comparator group. It was expected that the correlation coefficient would decrease if STEM interventions had been effective, suggesting that the link between poverty and attainment had been weakened.

Mean maths attainment of 5+ A*–C grades (including EM) was calculated for the longitudinal intervention group before and after intervention—that is in academic year 2007/08 and then in 2011/12. Table 3 shows a breakdown of the various school types included in the longitudinal intervention group.

Table 4 shows the number of schools at the beginning of 2007/08 that were registered for longitudinal interventions. Number of schools decreased by the end of 2011/12 as there were occasions when two nearby schools were merged to form a new academy. Thus for the first year of analysis the number of schools is higher than the last year.

Table 3. School types in intervention group.

| School type 2007/08 | Frequency |
|----------------------|-----------|
| Academies | 2 |
| Community | 198 |
| Foundation | 38 |
| Voluntary aided | 56 |
| Voluntary controlled | 5 |
| Total | 299 |

Table 4. Number of schools—longitudinal intervention and comparator.

| Group | Longitudinal intervention | Comparator |
|---------|---------------------------|------------|
| 2007/08 | 299 | 2720 |
| 2011/12 | 289 | 2409 |

The entry gap was considered for the longitudinal intervention schools. Thus the achievement gap was calculated before and after intervention using the formula:

$$\text{Entry gap} = \frac{\text{Numbers}_{\text{Intervention}} - \text{Numbers}_{\text{Comparator}}}{\text{Numbers}_{\text{Intervention}} + \text{Numbers}_{\text{Comparator}}}$$

$$\text{Achievement gap} = \frac{\text{Mean attainment}_{\text{Intervention}} - \text{Mean attainment}_{\text{Comparator}}}{\text{Mean attainment}_{\text{Population}}} - \text{Entry gap}$$

Correlation coefficients

Pearson’s correlation coefficient *R* was used, the possible values range from –1 to 1. A value of 1 means strong positive correlation—that is the higher the value of predictor variable the greater the value of outcome variable. If *R* is 0 it denotes no correlation—the factors are not linked. A value of –1 denotes strong negative correlation—which means as the value of the predictor variable increases the value of the outcome variable decreases.

Cross-product ratio

This is an estimate of the relative incidence of the outcome (attaining 5+ A*–C grades in GCSE, including EM) associated with exposure (STEM intervention). The cross-product ratio was estimated for mean maths performances in the longitudinal intervention group. For example in a table of the form:

| | |
|---|---|
| a | b |
| c | d |

no change was defined as $ad=bc$ or $ad/bc=1$. Here ‘a’ was the attainment of the intervention group before intervention, ‘b’ after intervention, ‘c’ was attainment of the comparator at the beginning, and ‘d’ at the end of the study. For a detailed discussion please see Banerjee (2016a, 2016b) and Gorard (1999).

Pupil level data multiple regression analysis

Multiple linear regression was used to understand whether attainment can be predicted based on explanatory variables, to determine the overall fit of the model and the relative

contribution of each of the explanatory variables of the total variance explained. For instance, this analysis tried to answer whether participation in STEM activity during KS3 and/or KS4 was a good predictor of maths attainment at the end of KS4. The continuous variable highest standardised points achieved in GCSE maths was used as the attainment indicator. The regression models used pupil background information such as SES, gender, ethnicity, language group, SEN status, participation in STEM initiatives, and prior attainment in maths as the independent predictor variables for the study. Current eligibility for FSM was used as an indicator of SES.

Missing data

Descriptive statistics showed independent and dependent variables have missing data in the range of 9–11%. For KS2 prior attainment in maths and science all missing data were excluded list-wise. This was because using a mean for missing data imputation rendered the data biased between the groups. However, for all other predictors missing data were treated as ineligible. For instance, all missing FSM were treated as FSM ineligible, missing data for SEN were treated as not SEN (Table 5).

Pre-analysis data estimation

In order to ascertain that the data can actually be analysed using multiple regression the following pre-requisites were checked. The dependent variable used for the multiple regression analysis—highest standardised points achieved in maths was measured on a continuous scale. Amongst the explanatory variables FSM eligibility, language group, ethnicity, gender, SEN, and participation in STEM activity were categorical variables. KS2 prior attainments in maths and science were interval variables. There was a linear relationship between (a) the dependent variable and each of the explanatory variables, and (b) the dependent variable and the predictor variables collectively. This linear relationship was confirmed by the residuals, which was the difference between observed values of the dependent variable and the value predicted by the regression equation for each case (Figure 2).

Multicollinearity

Very strong correlation between predictor variables means they are measuring the same thing, for example, current FSM eligibility and FSM eligibility during the last six years. It is then difficult to measure the individual contribution of each predictor and leads to inaccurate estimations of the relationship between predictor and outcome variables. In order to rule out multicollinearity, collinearity diagnostics were estimated. Tolerance, the amount of variance in one predictor variable not explained by other predictors, was checked. This value varies from 0 to 1, where a value close to 1 suggests other predictors do not explain the variance in that variable. A value close to 0 indicates almost all the variance in the variable is explained by other variables. For the predictors considered in this analysis all the variables had a tolerance higher than 0.8 (Table 6).

Table 5. Missing data.

| Missing data | FSM eligible now | Major language group | Major ethnic group | Gender | SEN | KS2 prior attainment in maths |
|--------------|------------------|--------------------------|--------------------|--------|---------|-------------------------------|
| Percentage | 11 | 11 | 11 | 0 | 11 | 9.3 |
| Treatment | FSM ineligible | Excluded cases list-wise | | NA | Not SEN | Excluded cases list-wise |

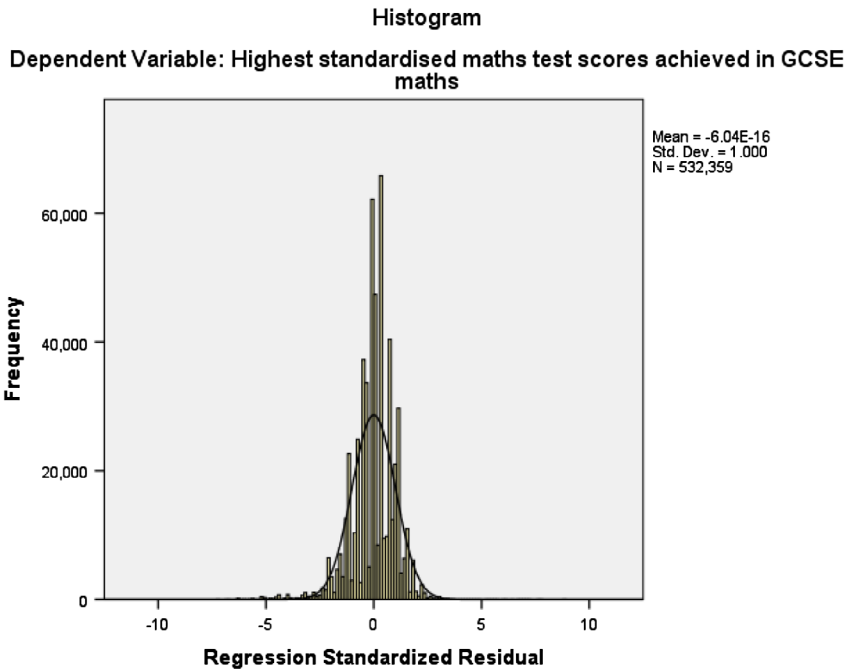


Figure 2. Residuals in regression.

Table 6. Collinearity statistics for regression analysis for mathematics test score prediction.

| Predictor variables | Tolerance |
|----------------------------------|-----------|
| KS2 mathematics prior attainment | 0.79 |
| SEN | 0.79 |
| FSM eligibility | 0.94 |
| Gender | 0.97 |
| Language group | 0.98 |
| Ethnicity | 0.99 |
| STEM intervention | 0.99 |

Results

The headline result for KS4 school performance is based on the mean percentage students achieving 5+ A*-C GCSEs or equivalent, including A*-C in both English and maths. Table 7 shows that there was an upward trend with the average for all schools 10 percentage points higher in 2012 than 2008. The average for intervention schools was higher than for all other schools in every year. This reinforces the point that schools willing and able to volunteer for STEM interventions are not some kind of random sub-set of all schools. They already have higher attainment scores. The gap is probably higher than this in reality since there will have been at least some schools in 2007/08 listed in the comparator group but who were participating in a STEM intervention.

The differences between the two groups were converted to simple differences between percentages, and into proportionate achievement gaps (Table 8). The comparator schools, starting from a slightly lower base figure, gradually caught up with the known intervention schools over time. When looked at in terms of achievement gaps, no clear difference was

Table 7. Mean percentage of pupils achieving 5+ A*-C GCSEs (including English and mathematics).

| Schools | | 2007/08 | 2008/09 | 2009/10 | 2010/11 | 2011/12 |
|----------------------|-------------------|---------|---------|---------|---------|---------|
| Intervention schools | % attainment | 54 | 55 | 60 | 62 | 62 |
| | Number of schools | 390 | 536 | 521 | 601 | 550 |
| Comparator schools | % attainment | 48 | 50 | 55 | 58 | 58 |
| | Number of schools | 2720 | 2566 | 2559 | 2422 | 2423 |
| All relevant schools | % attainment | 49 | 51 | 56 | 59 | 59 |
| | Number of schools | 3110 | 3102 | 3080 | 3023 | 2973 |

Table 8. Annual achievement gap estimation, intervention, and comparator schools.

| | 2007/08 | 2008/09 | 2009/10 | 2010/11 | 2011/12 |
|---------------------|---------|---------|---------|---------|---------|
| % points difference | 5.2 | 4.5 | 4.8 | 3.9 | 3.9 |
| Achievement gap | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |

Table 9. Correlation between attainment and FSM eligible pupils.

| Groups | | 2007/08 | 2008/09 | 2009/10 | 2010/11 | 2011/12 |
|--------------|----------|---------|---------|---------|---------|---------|
| Intervention | % FSM | 13 | 15 | 15 | 17 | 17 |
| | <i>R</i> | -0.6 | -0.6 | -0.6 | -0.6 | -0.5 |
| Comparator | % FSM | 15 | 15 | 16 | 17 | 17 |
| | <i>R</i> | -0.6 | -0.6 | -0.6 | -0.5 | -0.5 |

found over time. This means that, on this headline figure, there is no evidence that STEM intervention schools increased their attainment any faster than the comparator. It is important to note that this is not a question of merely dampening the effect size because only some intervention schools were known. Known intervention schools did not improve faster than another large group of schools, the majority of which did not undertake STEM interventions. There was no positive effect size.

Educational attainment is known to be linked to contextual indicators like a lower SES measured by FSM eligibility (Noden & West, 2009; Strand, 2014). The percentage of FSM pupils was a strong predictor of school maths performances. The links between these are summarised as the correlation (Pearson's *R*) between the percentage in each school reaching the KS4 indicator and the percentage of this type of potentially disadvantaged pupils. Table 9 shows the value of *R* is about the same every year and between the two groups. The more FSM pupils there are in any school, the lower its attainment is on average. This is already well-known. The key points were whether the link was different between groups, and whether there were signs that the intervention groups had somehow reduced the strength of the link over time as a result of the interventions. The answer to both is 'no'. The comparator groups had a slightly higher proportion of FSM pupils, again making the point that the intervention schools were slightly more privileged at the outset. Both groups increased FSM, perhaps as a result of the economic downturn from 2008 onwards. But the link with attainment remained the same in both groups. There is no evidence here that STEM interventions improved outcomes for less advantaged students.

Very similar results were found using the longitudinal intervention group of schools which enrolled the Year 7 cohort of 2007 to STEM interventions every year as they progressed through secondary schools. Again, more children from both intervention and comparator schools attained the 5+ A*-C threshold including English and maths (Table 10). Intervention

Table 10. Comparison of mean maths performances—longitudinal group.

| Schools | | 2007/08 | 2011/12 |
|--------------|-------------------|---------|---------|
| Intervention | Mean % attainment | 52 | 61 |
| | Number of schools | 299 | 289 |
| Comparator | Mean % attainment | 48 | 59 |
| | Number of schools | 2701 | 2409 |

Table 11. Achievement gap between intervention schools and comparator.

| | Before longitudinal intervention | After longitudinal intervention |
|-----------------|----------------------------------|---------------------------------|
| Achievement gap | 0.08 | 0.04 |

Table 12. School FSM intakes and correlation with attainment.

| Sub-groups | SES & <i>R</i> | 2007/08 | 2011/12 |
|--------------------|----------------|---------|---------|
| Intervention group | % FSM | 13 | 16 |
| | <i>R</i> | -0.6 | -0.5 |
| Comparator | % FSM | 15 | 17 |
| | <i>R</i> | -0.6 | -0.5 |

schools achieved higher than all other schools, both before and after intervention. However, while the intervention group showed an improvement of nine percentage points after intervention, the comparator improved by 11 percentage points. This also meant the percentage points difference reduced from four to two (actually 3.97 to 2.43).

The cross-product ratio was also estimated for this table, and was 1. This means no change was noted in attainment before and after intervention. Achievement gap was calculated between the intervention group and comparator using Newbould and Gray's formula. Mean percentage attainment was considered for calculations. The gap lowered after intervention because progress in the comparator's attainment was higher (Table 11).

All of these ways of presenting the findings show that the intervention group of schools did not make more progress than the comparator group. If anything, the comparator group appears to be catching up with the intervention group, for maths outcomes at the school level.

Again, the intervention group started with fewer FSM pupils (Table 12). Both groups show an increase in the percentage of FSM pupils over time. Both had the same level of negative link between the percentage of FSM pupils in each school and school attainment, and both improved this slightly to the same extent. This also indicates a greater improvement was seen in comparator than the intervention schools, and cannot be attributed to the STEM activities.

Multiple linear regression analysis results for GCSE maths attainment

Independent variables linked to maths attainment were included for analysis. This means all variables chosen as predictors were entered into the regression equation and contributed to *R* square. For model 1 below all explanatory variables excluding STEM intervention were included. The model summary suggests that together the explanatory variables KS2 maths prior attainment, SEN, SES, gender, ethnicity, and language group can predict the outcome

Table 13. Models from multiple regression analysis—predicting mathematics test scores.

| Model | Independent variables included | Additional variables included | <i>R</i> | Adjusted <i>R</i> square |
|---------|---|-------------------------------|----------|--------------------------|
| Model 1 | KS2 maths prior attainment, SEN, FSM | None | 0.77 | 0.59 |
| Model 2 | eligibility gender, language group, ethnicity | STEM intervention | 0.77 | 0.59 |

variable of highest standardised point scores achieved in maths with an accuracy of 76.6%. Adjusted *R* square was 0.6 suggesting that the model is good at predicting maths scores. However, for model 2 when STEM intervention was added as an independent variable it did not appear to change the *R* or adjusted *R* square values (Table 13).

Conclusion

Mean maths attainment for all schools increased from 2007 to 2012. A higher percentage of students achieved 5+ A*–C grades or equivalents including English and maths GCSEs in intervention schools every year than the comparator for both study designs. Attainment gap was however exactly the same each year for the repeated cross-sectional design. This meant the improvement in school performances was of a similar order in both intervention and comparator groups. However, the achievement gap between intervention and comparator narrowed down significantly after intervention for the longitudinal design. This was because comparator schools had improved attainment more than intervention schools after intervention.

In most analysis intervention schools had a similar or slightly lower share of lower SES pupils (school-level deprivation measure) than the comparator. A strong negative correlation of similar order was seen in intervention schools as well as the comparator group. This suggests correlation of school attainment in maths with percentage share of FSM pupils was not affected by STEM intervention. If STEM interventions were able to negate the effect of school-level deprivation factors such as SES it would be expected that the values for correlation coefficient would be lower for intervention groups as opposed to the comparator. This suggests there are perhaps other factors linked to this improvement that need to be investigated.

STEM enrichment and enhancement activities were run to motivate by increasing knowledge and improving understanding of these subjects. It is justified to expect that a possible outcome of participating in these activities throughout secondary schools as in the study reported here, can raise maths achievement in standardised national tests. However, the results reported in this paper show this did not happen. The intervention group did not outperform the comparator, in fact the comparator seems to be catching up. The findings provoke the question why is this happening? It also tries to understand why the results obtained are what they seem.

Limitations of the study

The quasi-experimental design used here is perhaps the most practical option for conducting outcome evaluations in the social sciences. By using pre-existing groups, such as individuals already enrolled in STEM enrichment and enhancement activities provided by others, it makes an evaluation possible and avoids the potential ethical concerns involved in

withholding or delaying treatment or substituting a less effective treatment for one group of study participants. The significant limitation of this design is that without randomisation, the study groups may have already differed in important ways that account for some of the group differences in outcomes after the intervention, and which cannot be controlled for by the analysis. In other words, this design provides practical but comparatively weaker evidence of programme effects than one that uses randomisation.

In the absence of randomisation, a matched comparator group can provide a good estimate of the effect of the intervention. A range of effect size estimates was used in the study, showing the difference between the intervention group and the comparator. An ideal matched group for this study would have been schools and pupils from schools who have definitely not participated in any STEM schemes, and such schools clearly exist (STEMNET, 2010). However, it was not possible to identify such schools, and so the compromise selected was to compare the known intervention schools with all other mainstream schools. This is likely to reduce the estimated effect size for the intervention, but should otherwise provide an unbiased estimate of whether the intervention was effective or not.

As is true with any longitudinal study, it is difficult to attribute outcomes solely to the intervention. This is because every child is exposed to a range of societal, familial, and school related factors during the years in secondary school, apart from these STEM enrichment and enhancement activities. Some of the former might ignite a passion for STEM and some others may turn students away. Owing to the long time period involved in this prospective longitudinal study, it is difficult to ascertain with certainty that the educational outcomes in terms of attainment and participation are solely due to the intervention. It is quite possible that several factors have led to raised/improvised attainment and continued/discontinued STEM participation. Of course, this only matters if such other factors are biased in terms of intervention and comparator schools.

Recommendations for practice

There are indications that school performances are gradually improving over the years. A range of factors affect students' subject choices including but not limited to their attitudes and aspirations. It is difficult to link educational outcomes such as attainment to only STEM activities. Several factors act together to improve attainment, and increase and widen participation. This study has focussed only on attainment in secondary schools, though it is equally important to understand how young people's attitudes (Banerjee, 2016a), motivation, or desire to continue with the subject are impacted by these activities (Banerjee, *in press-b*). Similarly this study has considered the population of secondary school pupils in England who took GCSEs; other qualification routes and different age groups of children engaging in these activities should be considered in further research.

Research findings suggest STEM enrichment and enhancement activities have not been phenomenally successfully in improving school performances. These schemes require huge investment of resources—in terms of staff engagement, time, and money. Given the high priority STEM agenda, if these schemes are not working perhaps the money should be saved and used elsewhere. Academic literature identifies robust studies in the UK and elsewhere, which have been clearly effective in improving cognitive development of disadvantaged pupils (for a detailed discussion please see Banerjee, 2015a, 2015b, 2016b), raising academic achievement, and sustaining pupil interests in STEM. This will ensure that with a similar or

reduced investment better results are obtained. Rigorous evaluations are required to understand what works. Some under-researched areas have been highlighted in this paper.

Acknowledgements

This study has been supported by the ESRC impact acceleration award.

Disclosure statement

No potential conflict of interest was reported by the author.

Notes on contributor

Pallavi Amitava Banerjee is a Lecturer in Education at the University of Exeter. Her research interests are in the area of educational effectiveness and improvement. This paper is an output of her PhD research project conducted at Durham University.

References

- Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2012). Science aspirations, capital, and family habitus: How families shape children's engagement and identification with science. *American Educational Research Journal*, 49, 881–908.
- Banerjee, P. A. (2015a). *Impact assessment of STEM initiatives in improving educational outcomes (Doctoral thesis)*. Durham: Durham University.
- Banerjee, P. A. (2015b). Can schemes to inspire tomorrow's scientists close the poverty attainment gap? *The Conversation*. Retrieved from <http://theconversation.com/can-schemes-to-inspire-tomorrows-scientists-close-the-poverty-attainment-gap-35805>
- Banerjee, P. A. (2016a). A longitudinal evaluation of the impact of STEM enrichment and enhancement activities in improving educational outcomes. *International Journal of Educational Research*, 76, 1–11.
- Banerjee, P. A. (2016b). A systematic review of factors linked to poor academic performance of disadvantaged students in science and maths in schools. *Cogent Education*, 3, 1178441. Retrieved from <http://doi.org/10.1080/2331186X.2016.1178441>
- Banerjee, P. A. (in press-a). Review of gender, identity and educational leadership. *Kaleidoscope*.
- Banerjee, P. A. (in press-b). Is informal education the answer to increasing and widening participation in STEM education? *Review of Education*.
- Behr, N. V. (2011). *A-level students in England unprepared for entry to STEM degrees*. The Royal Society. Retrieved from <http://blogs.royalsociety.org/in-verba/2011/02/15/a-level-students-in-england-unprepared-for-entry-to-stem-degrees/>
- DCSF. (2004). *Science, technology, engineering and maths (STEM) mapping review*. Department for Children, Schools and Families. Retrieved from <http://www.dcsf.gov.uk/hegateway/hereform/initiativemaps/index.cfm>
- DeWitt, J., Archer, L., & Osborne, J. (2013). Nerdy, brainy and normal: Children's and parents' constructions of those who are highly engaged with science. *Research in Science Education*, 43, 1455–1476.
- DfE. (2014). *What is the most effective way of increasing young people's take-up and attainment of strategically important subjects, especially science and maths?* (DFE-00093-2014). Retrieved from www.gov.uk/government/publications
- DfES. (2006). *The science, technology, engineering, and mathematics programme report*. Department for Education and Skills. London: HMSO.
- Garnham, D. (2011). *National Curriculum: Now and for the future, SCORE conference report*. Retrieved from <http://www.score-education.org/media/7623/finalcaq.pdf>
- Gorard, S. (1999). Examining the paradox of achievement gaps. *Social research update* (26), Autumn edn. Retrieved from <http://sru.soc.surrey.ac.uk/SRU26.html>

- Gorard, S. (2012). Who is eligible for free school meals? Characterising free school meals as a measure of disadvantage in England. *British Educational Research Journal*, 38, 1003–1017.
- Hobbs, G., & Vignoles, A. (2010). Is children's free school meal 'eligibility' a good proxy for family income? *British Educational Research Journal*, 36, 673–690.
- Hodgen, J., Marks, R., & Pepper, D. (2013). *Towards universal participation in post-16 mathematics: Lessons from high-performing countries*. Nuffield Foundation. Retrieved from www.nuffieldfoundation.org
- Hodgen, J., Pepper, D., Sturman, L., & Ruddock, G. (2010). *Is the UK an outlier? An international comparison of upper secondary mathematics education*. Nuffield Foundation. Retrieved from www.nuffieldfoundation.org
- Ing, M., & Nylund-Gibson, K. (2013). Linking early science and mathematics attitudes to long-term science, technology, engineering, and mathematics career attainment: Latent class analysis with proximal and distal outcomes. *Educational Research and Evaluation: An International Journal on Theory and Practice*, 19, 510–524.
- Noden, P., & West, A. (2009). *Attainment gaps between the most deprived and advantaged schools*. London: The Sutton Trust. Retrieved from [http://eprints.lse.ac.uk/23921/1/Attainment_gaps_between_the_most_deprived_and_advantaged_schools_\(summary\).pdf](http://eprints.lse.ac.uk/23921/1/Attainment_gaps_between_the_most_deprived_and_advantaged_schools_(summary).pdf)
- Ofsted. (2011a). *Improving science in colleges: A survey of good practice*. Ofsted.
- Ofsted. (2011b). *Successful science: An evaluation of science education in England 2007–2010*. Ofsted.
- POSTNOTE. (2011). *The Parliamentary Office of Science and Technology Informal STEM education*, Number 382, June.
- Rodeiro, C. L. V. (2007). A level subject choice in England: Patterns of uptake and factors affecting subject preferences. *Financial Times*. Retrieved from <http://www.cambridgeassessment.org.uk/ca/digitalAssets/114182SurveyReport-Final.pdf>
- Royal Society. (2011). *Preparing for the transfer from school and college science and mathematics education to UK STEM higher education: A 'state of the nation' report*. The Royal Society.
- SCORE. (2010). Admission tutors: Round table meeting report April. Retrieved from http://www.score-education.org/media/3433/admissions_tutor.pdf
- SEMTA (2010). *UK report: Sector skills assessment report for science, engineering and manufacturing technologies*. Watford: SEMTA.
- Shuttleworth, I. (1995). The relationship between social deprivation, as measured by individual free school meal eligibility, and educational attainment at GCSE in Northern Ireland: A preliminary investigation. *British Educational Research Journal*, 21, 487–504.
- Smith, E., & Gorard, S. (2011). Is there a shortage of scientists? A re-analysis of supply for the UK. *British Journal of Educational Studies*, 59, 159–177.
- STEMNET. (2010). *STEMNET report*. Retrieved from <http://www.stemnet.org.uk>
- Strand, S. (2014). School effects and ethnic, gender and socio-economic gaps in educational achievement at age 11. *Oxford Review of Education*, 40, 223–245.
- Sturman, L., Burge, B., Cook, R., & Weaving, H. (2012). *TIMSS 2011: Mathematics and science achievement in England*. Slough: NFER.
- Sturman, L., Ruddock, G., Burge, B., Styles, B., Lin, Y., & Vappula, H. (2008). *England's achievement in TIMSS 2007*. Slough: NFER.
- UKCES. (2015). *High level STEM skills requirements in the UK labour market: Evidence report*. Wath-upon-Dearne & London: UKCES. Retrieved from www.gov.uk/ukces
- UK Parliament. (2011). *Educating the next generation of scientists—Increasing take-up and achievement*. Public Accounts Committee publication and record. Retrieved from <http://www.publications.parliament.uk>
- Wynarczyk, P. (2008). *An overview of initiatives in science, technology, and engineering across England, Scotland and Wales*. Report commissioned by the UK Resource Centre for Women in Science, Engineering and Technology (UKRC) project Women in SET.
- Wynarczyk, P., & Hale, S. (2009). *Improving take up of science and technology subjects in schools and colleges: A synthesis review*. Report prepared for ESRC and DCSF.

Appendix. Ethical issues

The aims of the evaluation research project, planned use of data, strategy for dissemination of research findings, and its likely implications were explained to all STEM activity providers at the onset when the data request was made. Thereafter, the providers were asked if they were willing to be named in the research reports. Names of schools and activity providers have been anonymised as agreed upon.