

The English Mastery Programme Pilot Study Impact Evaluation

Dr Celeste Cheung, Research and Evaluation Manager
Dr Irina Sirbu, Data Analysis Officer



Executive Summary

The English Mastery curriculum and teacher development programme is a comprehensive intervention that aims to raise attainment for all pupils in key stage 3 English. The programme aims to deepen and accelerate students' reading and writing progress by exposing them to a cumulative, knowledge-rich curriculum and by training subject teachers on the specific curriculum topics. Compared to traditional curricular, more challenging texts are studied and greater emphasis is placed on vocabulary acquisition, cultural capital, background knowledge and explicit grammar instruction.

This report summarises the results of an independent evaluation conducted by the Research and Impact Department at The Brilliant Club. A total of 1700 pupils were involved in this evaluation. Of these, 851 pupils in seven schools received the intervention and completed a GL Assessment's Progress Test in English (PTE) once a year over the duration of the intervention in Year 7, 8 and 9 (treatment group). A group of 849 pupils in seven other schools completed the PTE in Year 9 but received no intervention (control group). Prior attainment (Key Stage 2 reading scores) and demographic data (age, gender, FSM and EAL status) were collected for both groups of pupils.

This report includes three sets of analyses: 1) A cross-sectional comparison of Year 9 English scores between treatment and control groups to examine the overall effects of the intervention; 2) A longitudinal analyses on progression in the treatment group and a comparison of expected progression with national data in both groups; 3) Subgroup analyses to explore whether the intervention had differential effects on pupils with different characteristics.

Key Findings

- The English Mastery Programme had a significant positive effect on pupils' English scores, regardless of pupils' prior attainment and demographic characteristics.
- Pupils in the treatment group scored on average 4.32 percentage points higher than the control group. The effect size of the intervention was medium ($d'=0.28$), equivalent to four months' additional progress (Education Endowment Foundation, 2012).
- Overall, the intervention had similar effects across three years. It was particularly beneficial for pupils with lower prior attainment, and marginally more effective for non-Free School Meals (FSM) pupils.
- The intervention had similar effects for girls and boys, with girls consistently outperforming boys in all three consecutive years.

1. Treatment vs control analysis

Performance at Key Stage 3 English was measured using the GL Assessment's Progress Test in English (PTE). The test assesses pupils' technical English skills and reading comprehension using standardised ability scores (SAS). The treatment group (n=851) completed three GL tests over three consecutive years (in Year 7, 8 and 9). The control group (n=849) completed the GL test only once in Year 9. Key Stage (KS) 2 reading scores were also obtained for 87% (n=737) of pupils in the treatment group and 78% (n=664) of pupils in the control group, along with other demographic information including age, gender, Free School Meals (FSM), and English as an Additional Language (EAL) status.

1.1 Is there a significant group difference between treatment and control group in Year 9 SAS?

The treatment group had significantly higher scores compared to the control group, with a 4.32 percentage point difference.

We observed a significant difference in Year 9 SAS between the treatment and control groups, controlling for the effects of age, gender and EAL status; $F(4, 1670) = 15.50, p < 0.01$ (Figure 1). The Cohen d' effect size of the group difference is 0.28, which indicates a medium effect that equates to four months' additional progress (EEF, 2012). The treatment group scored 4.32 percentage points higher than the control group (Figure 1).

The rationale for including only age, gender and EAL status as covariates in the above analysis was because these were the variables where significant differences were observed between the treatment and control groups (Table 1). On average, the treatment group was three months older than the control group. There was a higher proportion of boys and a higher proportion of pupils with EAL status in the control group. No group differences were observed for FSM or KS2 reading scores therefore these variables were not included as covariates in the main analysis (Table 1).

School-level break down of Year 9 SAS for both groups are presented on Figure 2. School and pupil characteristics for each school can be found in appendix A (treatment group) and appendix B (control group).

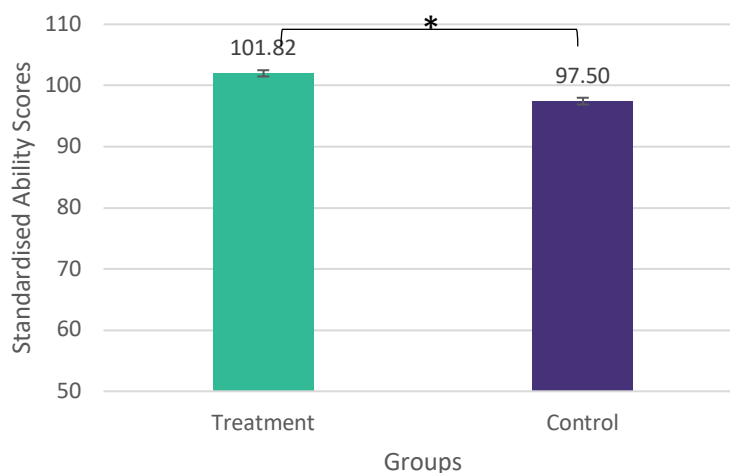


Figure 1. Mean differences (standard error bars) between treatment and control group on standardised ability scores.

Table 1. Descriptive statistics and group comparisons for demographic variables.

	N	Treatment	Control	χ^2/t	p
Age (months)	1700	171.17 (3.58)	167.87 (9.15)	9.76	<0.01
Male (n, %)	1698	367 (43%)	497 (58%)	36.77	<0.01
FSM	1687	504 (60%)	481 (56%)	2.65	0.10
EAL	1675	413 (50%)	536 (63%)	31.88	<0.01
KS2 reading	1421	4.65 (0.75)	4.63 (0.79)	0.42	0.68

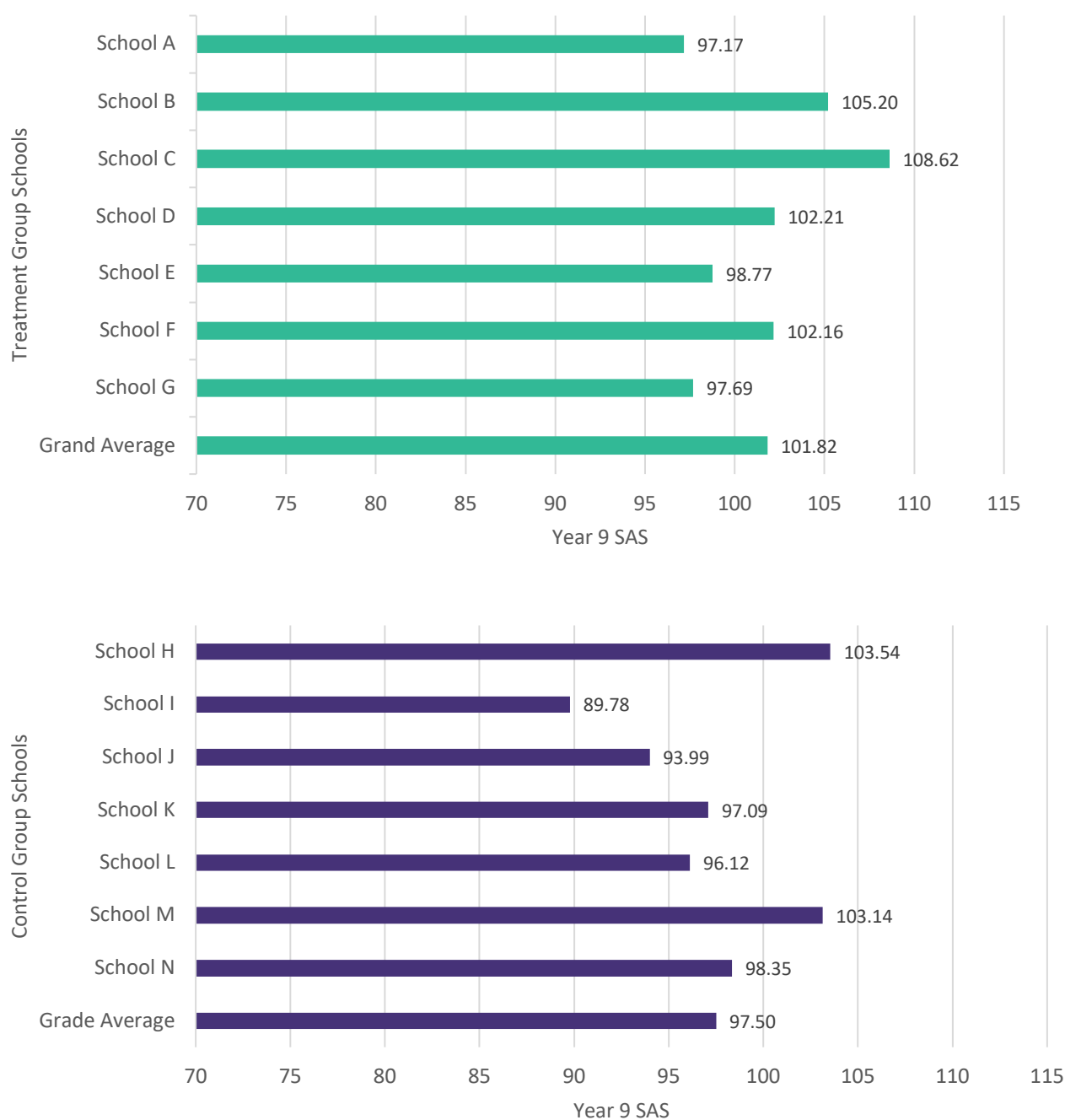


Figure 2. Year 9 SAS for schools in the treatment (top panel) and control groups (bottom panel).

2. Longitudinal analysis on progression

The following analysis included 707 pupils in the treatment group only who had complete data for three consecutive years of the intervention (Y1-Y3).

2.1 Did the intervention have similar effects across three years?

The impact of the intervention doubled between the first and third year of the intervention

The intervention had a similar degree of impact between Y1 and Y2 and between Y2 and Y3; $t(706)=4.52$, $p<0.01$, $d'=0.09$; $t(706)=4.34$, $p<0.01$, $d'=0.08$, respectively. The overall impact of the intervention (Y1 vs Y3) was additive, as the effect size was twice as large for this comparison ($t(706) = 7.97$, $p<0.01$, $d'=0.17$) (Figure 3).

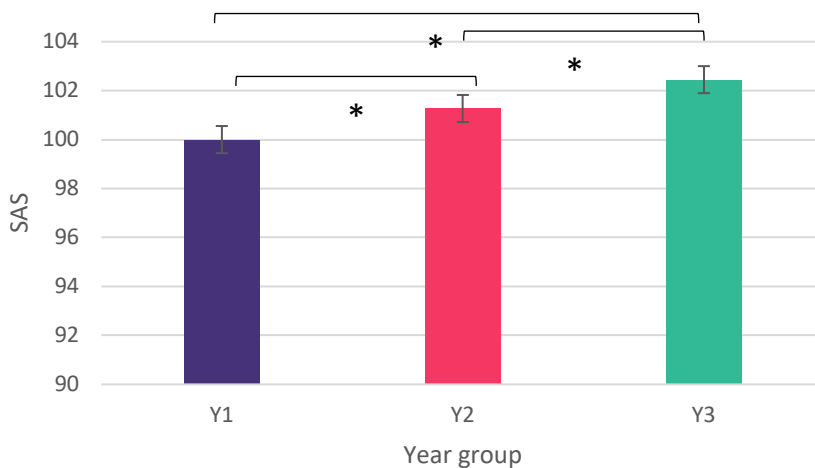


Figure 3. GL test SAS scores in the treatment group over three consecutive years.

2.2. Comparison with national trends

Of all the pupils in the treatment group with Year 9 data ($n=851$), 77% ($n=656$) also had complete data from KS2 through to Year 9. The percentile rank gains were larger over the first year of the intervention (7 percentiles) compared to the second (2 percentiles) and third year (2.5 percentiles) in the treatment group pupils. Percentile rank trajectory for each school in the treatment group can be found in Appendix C.

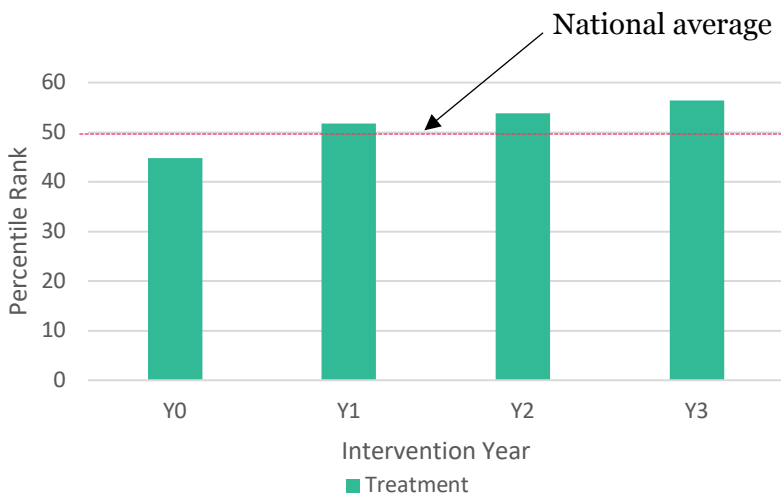


Figure 4. Percentile rank in the treatment group over between KS2 (Y0) and Year 9 (Y3).

This graph shows the percentile rank of pupil performance throughout the intervention. Performance was below average before the intervention (Y0), which increased to above average in Y1, 2 and 3 of the intervention. By the end of the intervention, pupils were 6 percentiles above the national average.

Comparing pupils' actual performance against their expected GCSE levels based on national data, we found that, pupils in the treatment group increased their expected GCSE performance by around half a grade, from an expected grade 4 to a grade 5 (Figure 5). The percentile rank change varies between schools (Figure 6), but on average we observed a change of 11 percentiles from Key Stage 2 and Year 9 in the treatment group. For most schools, pupils' expected GCSE levels increased from a grade 4 to a grade 5, except for School C where the expected grade increased from a grade 5 to a grade 6. (Figure 5).

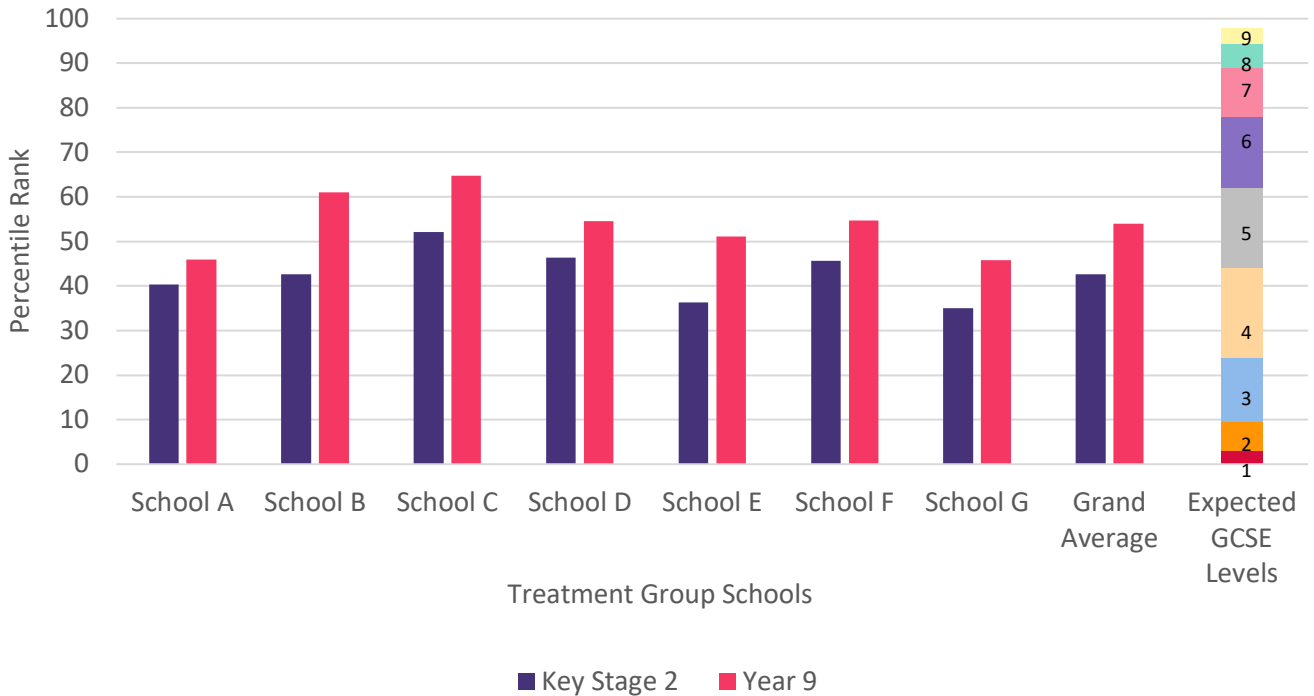


Figure 5. Pupils' Key Stage 2 and Year 9 performance against their expected GCSE levels in the treatment group schools.

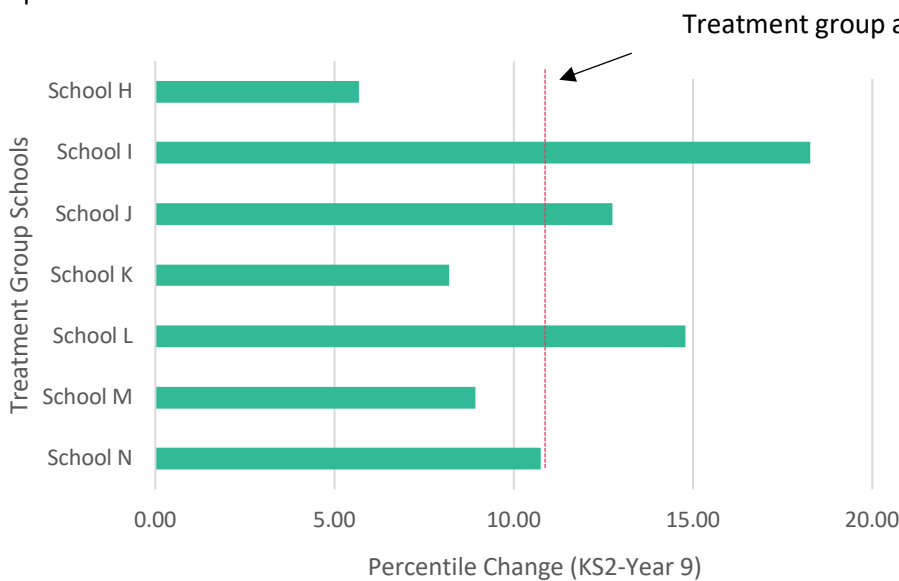


Figure 6. Percentile change in the treatment group schools.

In the control group, on average, pupils' expected GCSE grade also increased from a grade 4 to a grade 5. However, this was driven mostly by one school (School I) where we observed an increase of more than one grade (Figure 7). Pupils in the control group showed an increase of 8 percentiles from Key Stage 2 to Year 9 (Figure 8), which was smaller than in the treatment group.

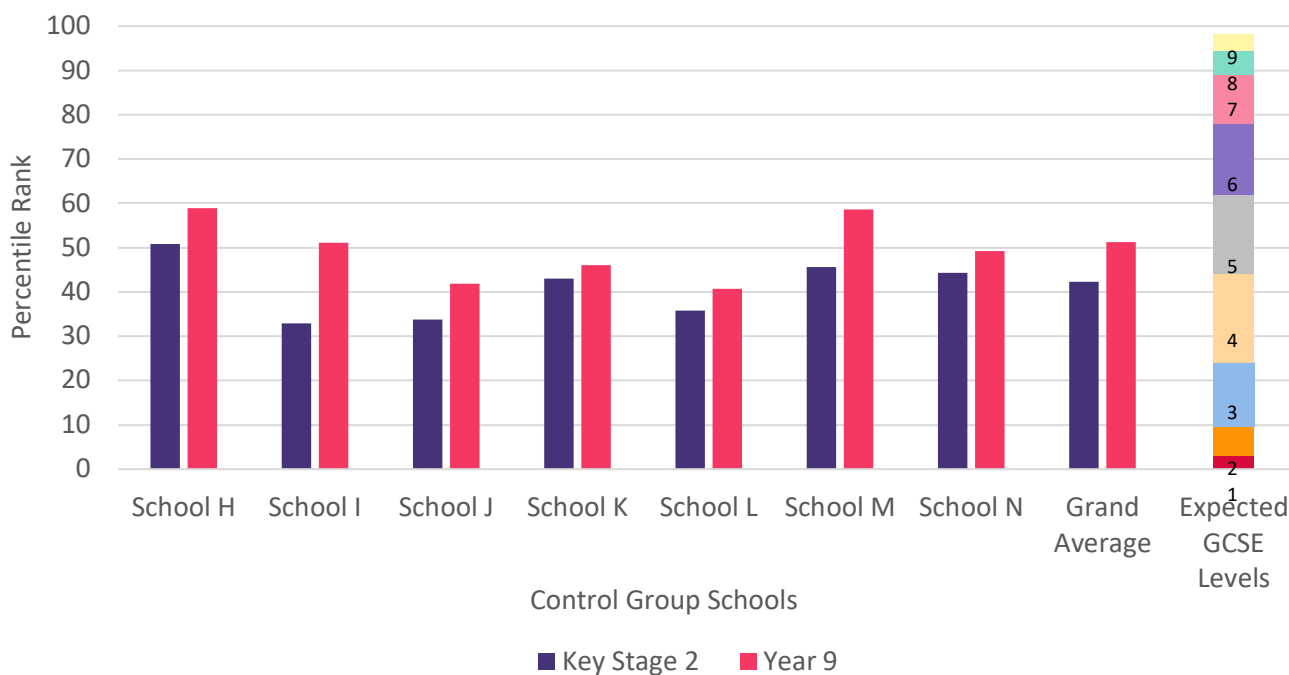


Figure 7. Pupils' Key Stage 2 and Year 9 performance against their expected GCSE grades in the control group schools.

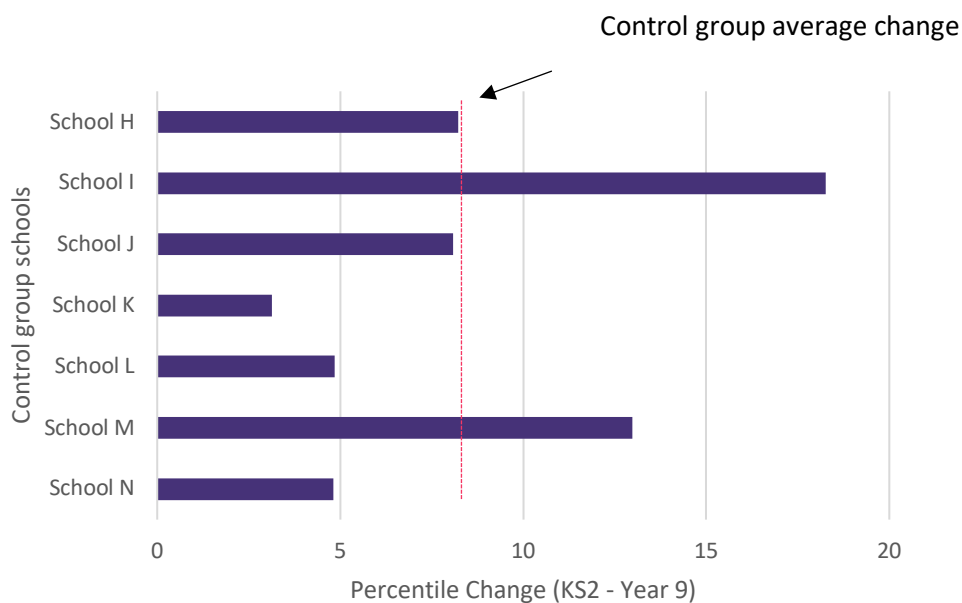


Figure 8. Percentile change in the control group schools.

3. Differential effects of intervention

We carried out additional analyses to explore whether the intervention had differential effects based on pupils' characteristics. For the questions below, we analysed the PTE scores from Year 7 to Year 9 (Y1-Y3 of intervention) in the treatment group only.

3.1 Is the intervention more effective for particular ability groups?

Pupils with lower prior attainment in reading made more progress relative to other pupils, but they did not catch up to the levels of their peers with higher prior attainment

The following analysis included pupils with complete data from KS2 to Year 9. Pupils were divided into three groups (terciles), based on their KS2 reading scores. Pupils in the bottom tercile (33%, n=244) scored between 1.5 and 4.54 in their KS2 reading, pupils in the middle tercile (n=220) scored between 4.55 and 5.10 and pupils in the top tercile (n=192) scored between 5.11 and 5.90.

We conducted further analyses to examine the trajectory in more detail for each KS2 tercile. The results suggest that pupils in the bottom tercile made more improvements from the start to the end of the intervention (Y1 vs Y3) compared to pupils in the top tercile (Figure 9), which was driven by the difference observed in the second and third year of the intervention. The gap in performance between the top and bottom terciles narrowed by 2.5 percentage points, from a mean difference of 23.6 percentage points in the first year of intervention (Y1) to 21.12 to the third (Y3).

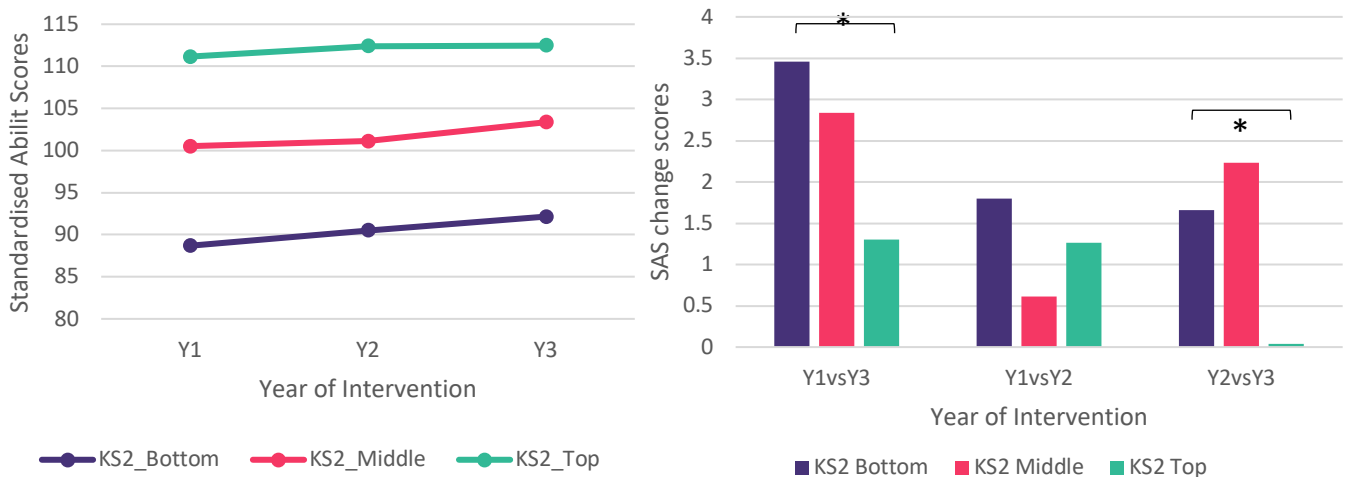


Figure 9. Performance trajectory and change scores between pupils with different levels of prior attainment at different years of intervention

To ensure that the effects observed above were not due to confounding demographic characteristics (the proportion of non-FSM pupils were higher in the top tercile, but the three terciles were not significantly different on age, gender or proportion of EAL pupils), we re-ran the analyses including FSM as covariate and the results remained unchanged. Pupils with the lower prior attainment in English (irrespective of FSM status) benefited more from the intervention. Pupils with the higher prior attainment in English responded well to the intervention between Y1 and Y2 of the intervention but were not responsive to the intervention in the latter year.

3.2 Intervention effects on FSM vs non-FSM pupils

The intervention had positive effects on all students regardless of FSM status, but it had marginally stronger effects on non-FSM pupils.

Non-FSM pupils scored significantly higher than FSM pupils at all three timepoints ($F(1, 695) = 21.59, p < 0.01$); Figure 5, and the average gains made were higher in the non-FSM pupils, indicated by a marginally significant interaction effect $F(1, 695) = 3.84, p = 0.05$ (Figure 10). This was driven by a marginally significant group difference in change score between the start and end of the programme (Y1 vs Y3) ($t(695) = 1.96, p = 0.05$), which was not significant for the other two years ($p > 0.20$) (Figure 10).

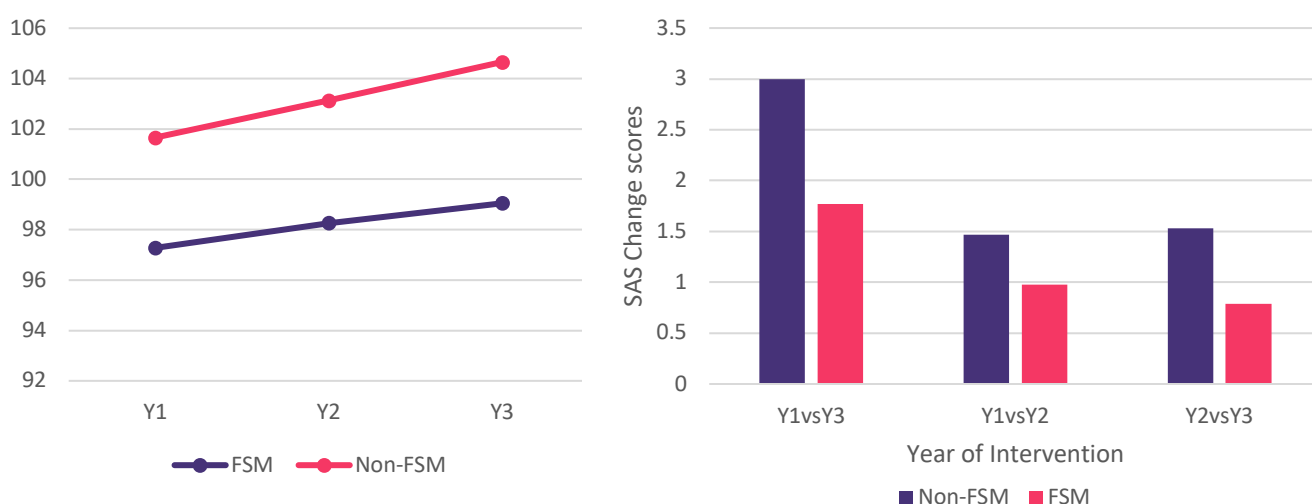


Figure 10. Performance trajectory and change scores between FSM and non-FSM pupils in the treatment group

3.3 Intervention effects on EAL vs non-EAL pupils

EAL pupils performed better than non-EAL pupils in the control group but not in the treatment group, the intervention was equally beneficial for both EAL and non-EAL pupils.

No significant group differences between EAL and non-EAL pupils for any of the three timepoints ($F(1, 692) = 0.56, p = 0.45$). As both groups showed similar performance trajectory across the three years of intervention, no further analyses were conducted.

Although in the treatment vs control group analysis the results suggested that EAL pupils performed better than non-EAL pupils, results from this analysis on the subgroup of pupils did not confirm this finding. The results also suggest that the intervention benefited both EAL and non-EAL equally.

3.4 Intervention effects on boys vs girls?

Both boys and girls benefited equally from the intervention. Girls consistently outperformed boys throughout the intervention.

Girls scored significantly higher than the boys across all three years ($F(1, 704) = 7.38, p < 0.01$; Figure 7). However, the intervention had similar effects across both groups, as indicated by a lack of interaction effect ($F(1, 704) = 0.60, p = 0.44$). Girls made marginally more progress compared to boys. This was driven by the difference in the first two years of the intervention (Figure 11), but no significant difference was observed ($p > 0.05$).

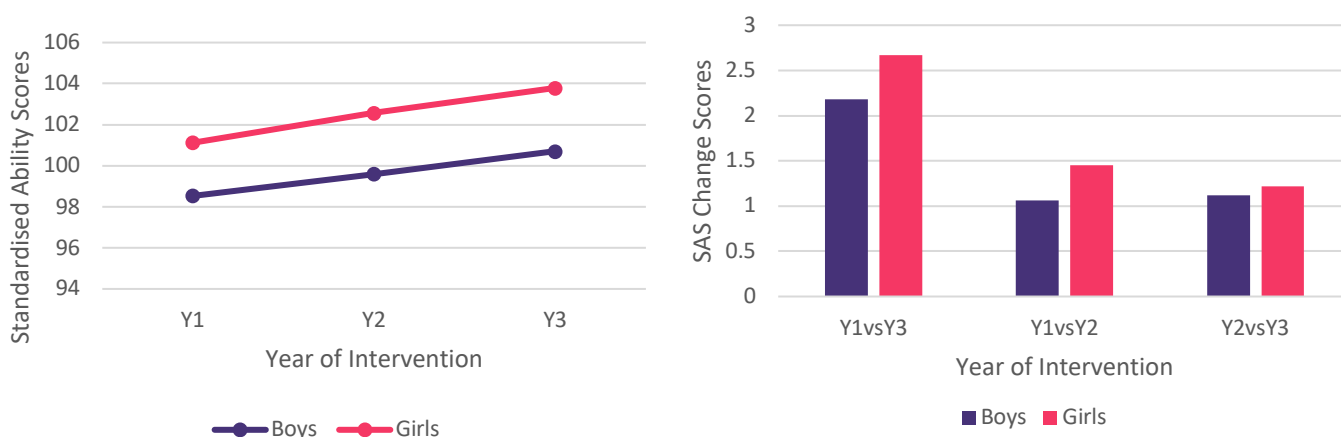


Figure 11. Performance trajectory and change scores between boys and girls

3.5 Is SAS predicted by school Ofsted performance?

Ofsted school judgement was associated with pupil's English performance in the control group but not in the treatment group.

Overall, SAS was associated with Ofsted judgement ($F(2, 1697) = 22.12, p < 0.01$). Schools with 'outstanding' judgement had significantly higher scores compared to schools with 'good' or 'requires improvement' judgement (Appendix A and B). The result remained after controlling for pupil-level demographic characteristics ($F(6, 1661) = 16.35, p < 0.01$).

In the treatment group, pupils from schools with Ofsted judgement 'good' did not score significantly differently in Year 9 SAS from their peers from schools with an 'outstanding' Ofsted judgement ($t(731) = 0.45, p = 0.65$). In the control group, pupils from schools with an 'outstanding' Ofsted judgement performed higher than schools that were 'good' or 'requires improvement', with the overall effect being statistically significant ($F(2, 846) = 25.80, p < 0.01$). Ofsted judgement was more strongly correlated with pupils' English performance in the control group schools than the treatment group schools (Figure 12). However, this is unlikely to be due to the intervention itself, as this difference in correlations is also present in KS2 data.

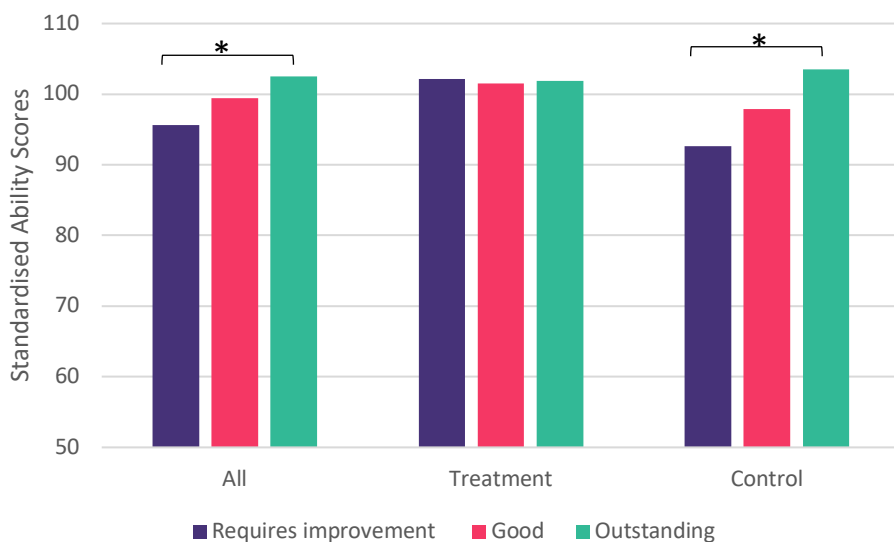


Figure 12. Performance by gender in all pupils, and in treatment or control groups only.

Appendix A. School and pupil characteristics and performance in the treatment group

Treatment group	Average grade on entry	Ofsted judgement	% of FSM pupils	% of EAL pupils	KS2 Score (number of pupils)	Y9 SAS (number of pupils)
School A	4.48	Outstanding	10	76	4.64 (128)	105.20 (142)
School B	4.85	Good	25	27	4.83 (93)	108.62 (110)
School C	4.74	Outstanding	52	38	4.80 (128)	102.21 (144)
School D	4.08	Outstanding	72	72	4.43 (114)	97.74 (115)
School E	4.33	Requires Improvement	35	53	4.74 (107)	102.16 (118)
School F	4.19	Good	65	35	4.50 (87)	98.88 (110)
School G	4.16	Good	18	44	4.55 (94)	97.44 (103)
Grand Total	4.40	N/A	39.57	49.29	4.65 (751)	102.47 (842)

Appendix B. School and pupil characteristics and performance in the control group

Treatment group	Average grade entry	Ofsted judgement	% of FSM pupils	% of EAL pupils	KS2 reading (number)	Standardised Ability Scores
School H	5.02	Outstanding	19	38	4.81 (162)	103.54 (178)
School I	3.97	Requires improvement	34	18	4.22 (71)	89.78 (169)
School J	4.16	Good	29	62	4.62 (77)	97.03 (87)
School K	3.99	Good	44	61	4.59 (56)	96.12 (66)
School L	4.44	Requires improvement	42	65	4.71 (77)	98.35 (84)
School M	3.90	Good	71	10	4.42 (105)	93.83 (136)
School N	4.40	Good	71	39	4.78 (122)	103 (135)
Grand Total/Average	4.27	N/A	44.29	41.86	4.63 (670)	96.44 (855)

Appendix C. Percentile rank trajectory over 3 years in all seven schools in the treatment group.

