

Evaluation of the Progress of the First Cohort of Sparx Maths Students Through Years 7 and 8.

Executive Summary

This report describes the performance of the first cohort of students using the Sparx maths product in Key Stage 3. Compared with national norms Sparx students made 67% more progress in year 7 and a further 63% more progress in year 8. Compared with a previous cohort of year 7 students in one of the same schools, they made 67% more progress. 70% of pupils had reached the expected KS3 standard by the end of the fifth half term in year 7 – more than a year earlier than in the national cohort. Progress was not negatively affected by prior ability, FSM status or gender. There was some indication that lower ability students actually made disproportionately more progress in year 8.

Introduction

Since 2011 Sparx has been working in partnership with schools to investigate how young people learn maths and how high-quality content, technology and traditional teaching methods can be blended to provide an individual learning experience that will get the best out of every pupil regardless of ability, socio-economic status or gender. By September 2016, after five years of research and content development, Sparx maths was released to year 7 students, and at the time of writing has been running for over two years. This paper is a retrospective analysis of the performance of those pupils through the first two years of KS3.

It is important to acknowledge from the outset that this data does not enable any definitive conclusions about the effectiveness of the Sparx product, as that would require a large scale randomised controlled trial across 100 or more schools, which is something that Sparx is currently considering how to achieve. However, it is still possible to get some idea of whether there is evidence that the product is promising or not by comparing the progress of students from the first Sparx cohort with progress nationally and also with progress from a pre-Sparx cohort at the same school.

Comparison Data

Since the abolition of compulsory KS3 testing it is no longer possible to track national pupil progress between KS2 and KS3. The best source of comparison data comes from a 2011 Department for Education report (DFE-RR096¹) which tracked the progress of 70,000 pupils from 10 local authorities through Key Stages 2 and 3. This report uses data from KS3 standardised assessment tests (SATS) measured on a national curriculum levels scale, and it measures progress in terms of sublevel units (where each level is divided into three equal

¹<https://www.gov.uk/government/publications/how-do-pupils-progress-during-key-stages-2-and-3>

sublevels). This data is augmented with termly teacher assessments measured on the same scale. In order to make a sensible comparison with this data, we have adopted the same NCL scale for the Sparx cohort, but it would equally have been possible to express progress using some other nationally recognised scale (e.g. GCSE levels or standardised scores). The key relevant findings from the DFE report are summarised below.

- Overall pupils improved by an average of 1.2 sublevels in maths in year 7 and 1.04 sublevels in year 8.
- Pupil progress was affected by prior attainment, with pupils who were behind in the previous Key Stage less likely to make progress than those at the expected level or above.
- Boys were less likely to make progress than girls.
- Attainment gaps between pupils eligible for school meals and those not eligible widened throughout KS3.

In addition, we wanted to make a comparison with a cohort of data from year 7 in one of the same schools but prior to using Sparx. We were interested in whether the Sparx cohort had a similar or differing pattern of results.

Assessment Strategy

Pupils in the Sparx cohort were assessed every term with two short 20 minute tests assessing ability on topics that had been in focus during the period. The first of these tested mathematical skills, while the second addressed more general problem-solving abilities. The same tests were administered at the beginning and end of each termly period to assess progress. Differentiation was achieved by having four different levels of each test targeted at different ability ranges. Adjacent levels had some questions in common, which allowed the use of an item response theory approach to put all students onto a common ability scale regardless of which variant of papers they sat. These termly assessments were used to calculate a student ability score at each the beginning and end of each term. These in turn were averaged over time to produce a rolling estimate of student ability.

In addition to the termly tests most pupils also sat some annual assessments. At the start of year 7 all the pupils sat a KS2 paper taken from 2016. At the end of year 7 they also sat a KS3 standardised assessment from 2006, while in year 8 a randomly selected half of the cohort sat another KS3 standardised assessment taken from 2007. These standardised tests were used to calibrate the ability score to the same scale as NCL levels and these calibrated scores were in turn used to impute data for any pupil who had not sat one of the annual exams. This was achieved using a linear regression where the independent variable was NCL level and the predictor was ability derived from the half-termly tests. The ability scores proved to be very accurate predictors of NCL levels right across the range with an R Squared of 0.837. Figure 1. shows the predicted versus actual NCL values for pupils in each year.

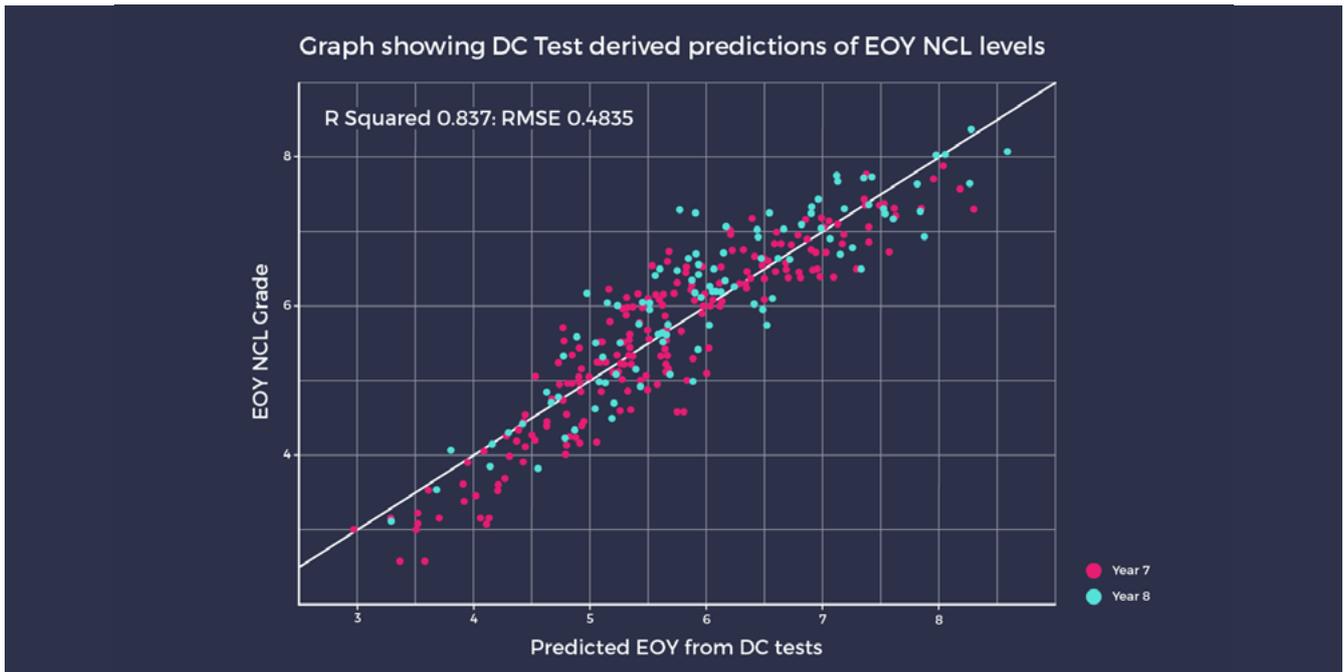


Figure 1. Relationship between predicted NCL scores and actual NCL Grades

All subsequent analyses for yearly progress used the ability measured by these start and end of year tests with missing data imputed from the appropriate half-termly assessment point. Finer grained analyses of termly progress used data from the half-termly tests but expressed on an NCL scale rounded down to the nearest sub level.

Participants in this study included all members of the first Sparx cohort who had participated in at least three of the half-termly assessments in both year 7 and year 8 (212 pupils in total). Where these students had missed a termly assessment, the data was imputed to be the average of the other half-termly tests of the appropriate type.

Results

Overall Progress Figure 2. illustrates how much progress students made during each year compared with the national average. Ability groups are defined by the students KS2 scaled scores with low ability defined as < 97 and high ability defined as > 105. In both years, across all ability groups, students in the Sparx cohort made significantly more progress than the average for the comparison cohorts. In addition, this progress was not negatively affected by prior ability. Indeed, in year 8 the lowest ability group makes the most progress. Overall, pupils made 2 sublevels of progress in year 7 and a further 1.7 levels of progress in year 8 compared with 1.2 and 1.04 in the comparison cohorts. That equates to 67% more progress in year 7 and 63% more progress in year 8. The pre-Sparx cohort in year 7 made an average of 1.2 sublevels of progress – exactly in line with the national average.

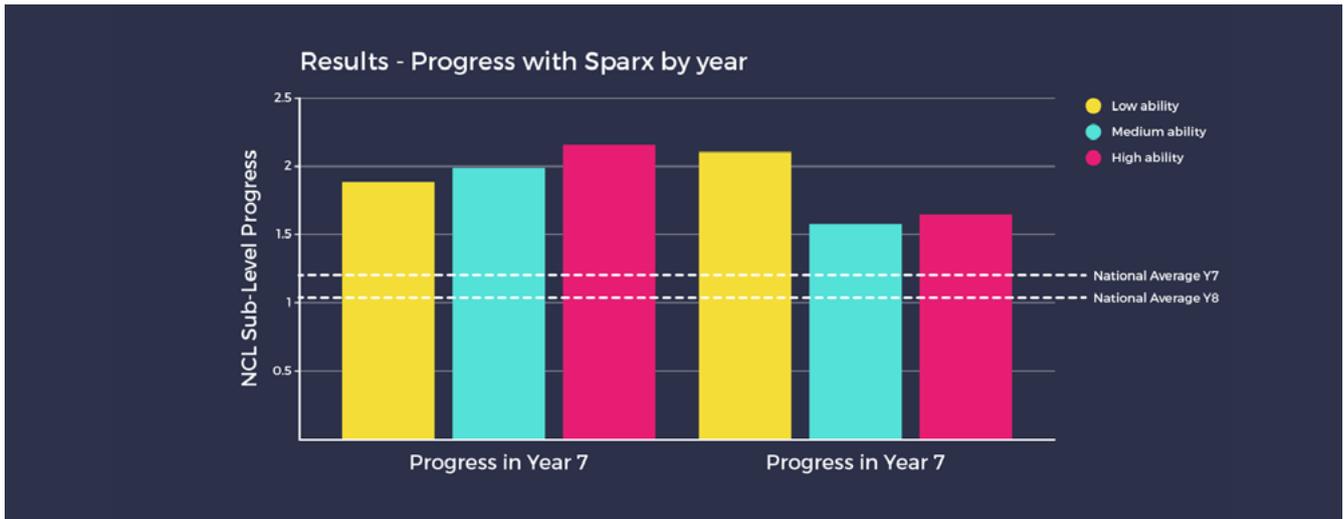


Figure 2. Progress of Sparx students in year 7 and year 8 grouped by ability.

Progress by FSM Status and Gender

A key finding from the 2011 DFE report was that students eligible for free school meals (FSM) made significantly less progress than pupils not eligible for FSM, and boys made significantly less progress than girls. Figure 3. And Figure 4. compare the progress of these groups in the Sparx cohort. It does not appear that there is any obvious difference between either of the groups. To confirm this the observed frequencies from these data were submitted to 4 chi-squared tests (one for each year group and each of the two variables of interest). In all cases the tests were not significant (all p's > 0.05). So, unlike in the national data, pupils in the Sparx cohort do not appear to be negatively affected by gender or FSM status.

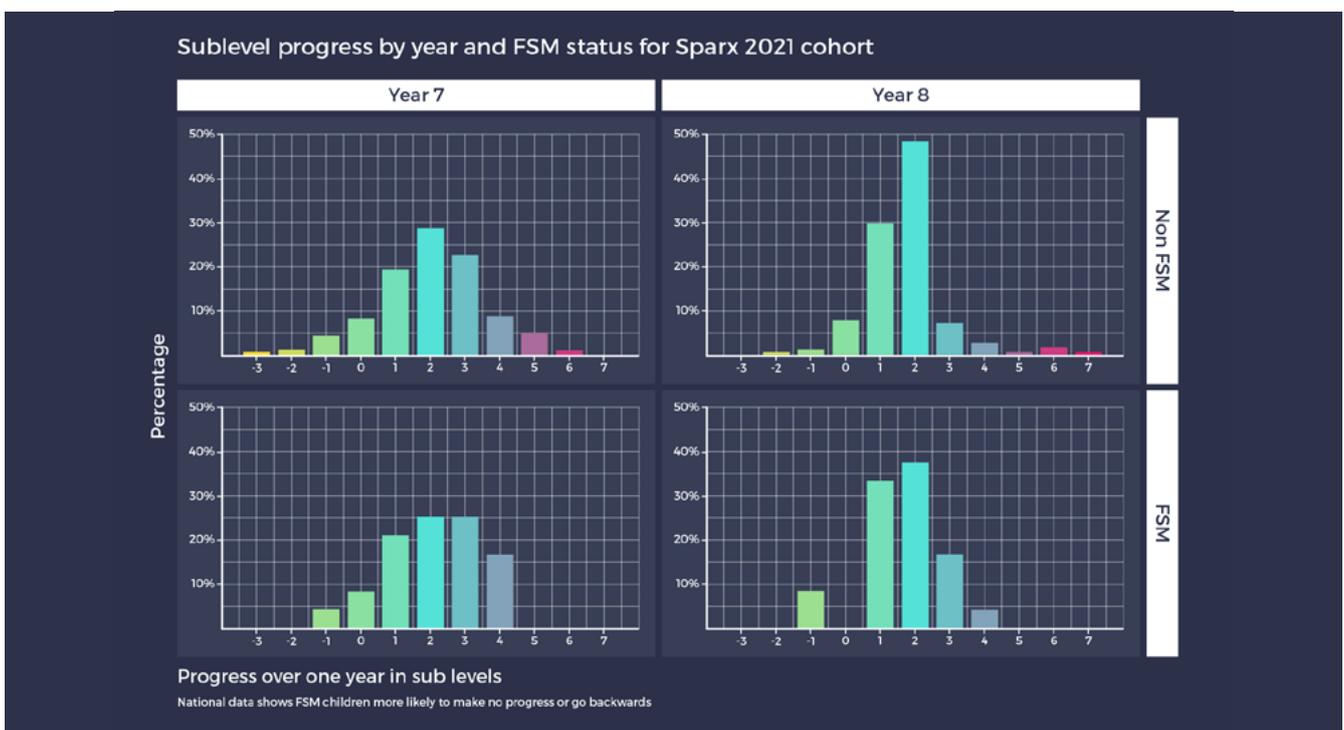


Figure 3. Comparison of progress between FSM and non-FSM pupils.

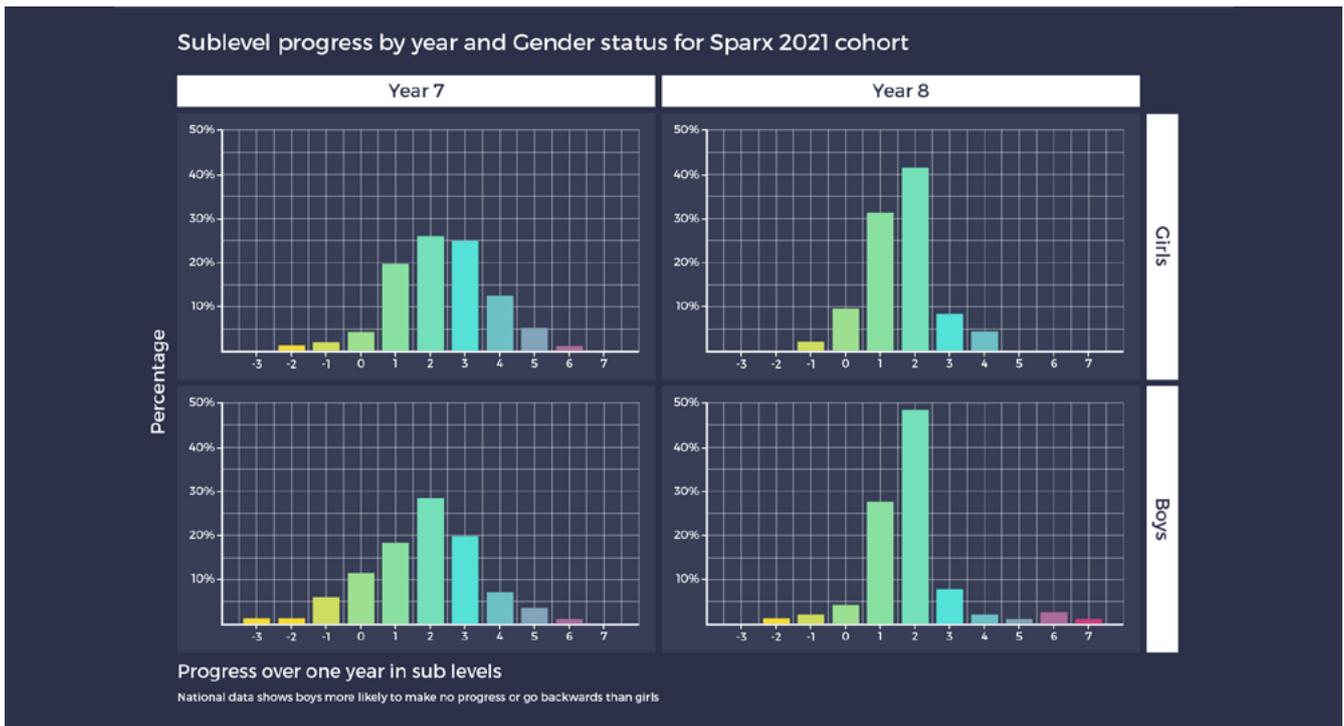


Figure 4. Comparison of progress between boys and girls.

Attainment over Time

One nice feature of this data set is that it is possible to track pupil attainment at a fine level of time resolution so we can compare the level of achievement at the end of each term with the nationally expected levels of attainment at the end of KS3. Figure 5 shows the percentage of students at each timepoint who have reached or exceeded an NCL level 5 (the expected standard for KS3). The following graph illustrates how those two percentages change across time in the Sparx cohort. What is very striking is how quickly pupils in this cohort meet the KS3 target. By the fifth half term in year 7, 70% of pupils have reached the expected KS3 standard, a point that doesn't occur in the national cohort until the end of year 8. By the end of year 8, 12% more pupils in the Sparx cohort have met the KS3 target than in the national cohort.

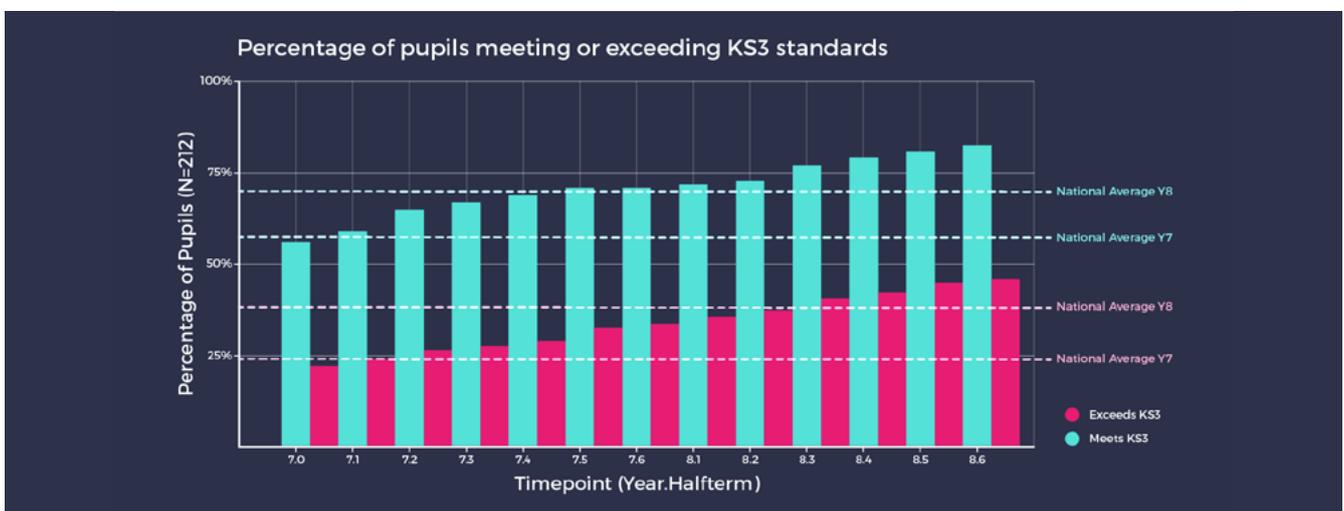


Figure 5. Percentage of students reaching or exceeding KS3 required level over time.

Limitations

As previously discussed, the main limitations of this study are the relatively small cohort size (212) and the fact that it is of necessity an observational study and not a randomised controlled trial. In addition, there are a number of other issues which should be considered:

- 1 **Measure of Progress:** The study uses changes in NCL levels as its progress measure. This measure depends on calibration with old style KS3 SATs exams from 2006 and 2007, which may now be somewhat outdated. This also means that we couldn't use KS2 tests as the start point for measuring progress as there is no established method for mapping between KS2 scaled scores and NCL grades, and the DFE actively discourages attempts to do this. Instead we again used a shortened old KS2 test administered to all students at the start of the year.
- 2 **Repeated Testing:** This particular cohort has been subject of considerable testing and it is possible that some of their progress may reflect improved ability at doing exams in general rather than pure maths ability.
- 3 **Comparison Measures:** While the comparison data set is the best available including data from over 70,000 pupils, it is limited in that it is now quite old (the data was collected between 2007 and 2011) and so it is possible the improved progress seen in our cohort merely reflects a general improvement across the country. The data also involved a large component of teacher assessment so it may not be as accurate as if it had been purely exam based. However, these teacher assessments were anchored by national KS2 and KS3 tests and as long as the error in the teacher assessments was random it is not likely to have affected the population level estimates of progress.

Conclusion

This paper set out to compare the progress in maths of 212 pupils in the first Sparx cohort with data taken from the DFE report on expected progress in KS2 and KS3. Overall Sparx students made 67% more progress in year 7 and 63% more progress in year 8. When compared with the last pre-Sparx cohort of year 7 students they also made 67% more progress. By the fifth half term of year 7 the percentage of students achieve level 5 or above matched that found in the national cohort at the end of year 8 - more than a year earlier. Progress was relatively consistent across prior ability groupings, with the lowest ability students making the same progress as higher ability students in year 7 and slightly more progress in year 8. This contrasts sharply with the national picture where lower initial attainment was associated with lower progress. In the Sparx cohort there was no evidence of progress being negatively affected by socioeconomic status or gender; this again is in contrast to the national data which shows that boys and students on free school meals make slower progress. While the relatively small size of the cohort, and the lack of a randomised control trial, mean these results cannot be taken as completely conclusive evidence that the Sparx maths products are effective at improving maths attainment, they do provide strong evidence of promise in this regard, suggesting further more detailed investigation would be rewarding.