THE GCSE PERFORMANCE OF INDEPENDENT SCHOOL PUPILS: GENDER AND SCHOOL TYPE DIFFERENCES

By John Charnley

SUMMARY

This report is concerned with differential GCSE outcomes in 2007 as between boys and girls in independent schools. It begins by establishing that over all subjects taken, girls achieve slightly more than boys (approximately one-fifth of a grade per subject entry), but the difference is a little less than in GCSE results nationally.

In order to compensate for variations in ability profiles (e.g. of boys and girls taking a particular subject), the report uses thereafter a *value-added* approach, in which the performance indicator is the grade obtained in relation to the average grade achieved by pupils of similar ability. Specifically, the report makes use of data from the *MidYIS* (*Middle Years Information System*) project, which is operated by the Curriculum, Evaluation and Management (CEM) Centre, Durham University. The author gratefully acknowledges the provision by the CEM Centre of the data used in the preparation of this report.

The MidYIS project provides two standardisations, measuring performance respectively relative to national and independent school norms. Considering first the cohort of pupils who took the MidYIS baseline test in Year 7 (in academic year 2002/3), the national standardisation revealed that independent school pupils achieved better value added than their counterparts in maintained schools, this being particularly the case towards the lower end of the independent school ability profile (which, as a whole, is well above the national average). This variation of value added with ability having been identified, it became preferable to use the independent standardisation for comparisons internal to the sector, as it provides a value added measure for independent school pupils that does not depend on their ability (that is to say, the value-added measure has an average value of zero for all pupils taking a particular subject both overall and for pupils in a selected ability range).

Some twenty subjects were selected for detailed analysis. It was found that girls over-performed boys on average by amounts varying from virtually nothing (in Mathematics and Physics, the other Sciences also showing small gender differences) to about half a grade (in French).

Separate comparisons were compiled of the value-added measures for (a) boys in boys' schools and in mixed schools and (b) girls in girls' schools and in mixed schools. In both cases it was found that pupils in single-sex schools achieved slightly but significantly better value added than pupils in mixed schools in most of the subjects selected for analysis. There was considerable variation between subjects, with the "school-type differences" lying in almost all cases between zero and 0.3 grade.

The Year 9 MidYIS test is taken in schools that do not admit pupils before the age of 13+, but it is also available to schools that also use the Year 7 test. Among the pupils who were tested in Year 9 (in academic year 2004/5) but not previously in Year 7, a different pattern of gender difference was found as compared with the Year 7 test sample. In several subjects the gender difference was in the boys' favour, and in several

others the difference was insufficiently large to be considered statistically significant. It was found that boys new to MidYIS in Year 9 achieved slightly better value added than those tested in both Years, but there was no corresponding difference for girls. For pupils tested in Year 9 only, the school-type differences in value added (single-sex vs. mixed) were negligible for girls, but significant for boys.

Finally, an attempt was made to apply what had been learned about gender and school-type differences to a consideration of characteristics of the independent standardisation value added measurement which might help schools to interpret the marked and sometimes perplexing differences between the value-added measures from the two standardisations in individual subjects.

The findings of this report are subject to the caveat that they relate only to the GCSE results in one year of pupils in schools that participated in the MidYIS project.

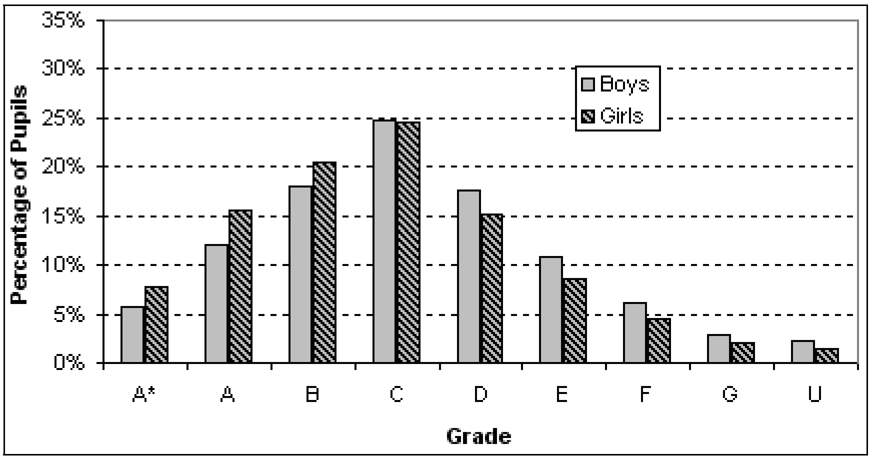
The report was not written at the instigation of the CEM Centre or at the request of any school, organisation or individual.

CONTENTS

INTRODUCTION

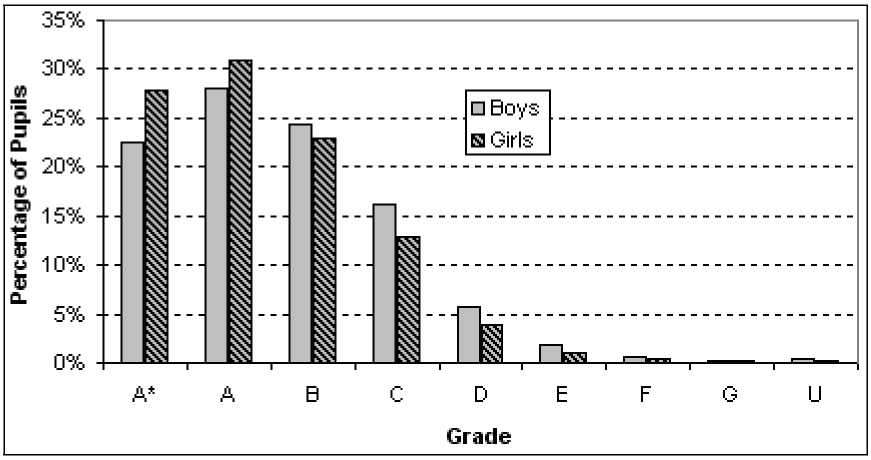
1. Every August, when national GCSE results are published, attention is drawn to the gender differences in the results. The superior average performance of girls is illustrated in *Figure 1* below.

Figure 1: GCSE results 2007 - grade distributions (all subjects) for boys and girls in all schools



2. The average points scores for boys and girls were 4.79 and 5.14 respectively (based on A* = 8, A = 7 and so on). The results represented here include the independent sector. *Figure 2* shows the corresponding grade breakdown for independent schools only and retains the same y-axis scale.

Figure 2: Independent sector GCSE results 2007 - grade distributions (all subjects) for boys and girls



3. The average points scores for boys and girls were 6.34 and 6.60 respectively, a difference of 0.24 grade. The results summarised here are clearly superior to the national average. The data underlying these analyses does not include short course GCSE entries. The independent sector accounted for 8.3% of the GCSE subjects taken in all schools by pupils completing Year 11 in 2007.

4. The principal factor in the determination of GCSE outcome is the ability of the pupil concerned. Consider the comparison of the average GCSE results of two samples (A and B) of pupils whose ability profiles can safely be assumed to be very similar. If the average results of sample A are better than those of sample B, and significantly so, it can be assumed that the superior performance of sample A is for reasons unrelated to ability differences between the two samples.

5. In the absence of any reason to doubt the similarity of the ability distributions of boys and girls in the independent sector, and given the sample size, the above chart and point score averages are probably a reliable indicator that there is a small but not insignificant difference between the 2007 GCSE performances of boys and girls in independent schools that is attributable to factors other than ability differences. Identifying those factors is, of course, a separate issue.

6. Investigations based on smaller pupil numbers are increasingly vulnerable to ability differences between the genders. If the girls in a sample (the entry for GCSE Spanish, for example) have a higher ability profile than the boys in the sample, then the superior performance of the girls may be due in part to their higher average ability, and a comparison of raw results cannot reliably establish the extent to which other factors have contributed to the gender difference.

MIDYIS DATA

7. For the reason outlined in paragraph 6, this paper makes use of data from the *MidYIS* (Middle Years Information System) project, in which pupils' aptitude is measured by a baseline test taken soon after entry to Year 7 or, in 13+ entry schools, to Year 9. (MidYIS baseline test scores are standardised, with a mean of 100 and a standard deviation of 15.)

8. GCSE results in a given subject are then compared with the average result for pupils of similar ability. MidYIS thus measures progress relative to a norm over the period elapsing between the baseline test and GCSE. This is often called "value added", a convenient shorthand. MidYIS data thus offers the potential of a performance indicator for comparative purposes which consists of outcome in relation to ability i.e. is independent of ability.

9. The detailed working of the MidYIS project will not be described here. In summary, a *regression line (or curve)* is computed which, in effect, represents the average grade achieved by pupils within a narrow range of MidYIS baseline test scores. The amount by which a pupil's result exceeds (or falls below) the average is called the *raw residual*, and is a grade fraction.

THE YEAR 7 MIDYIS SAMPLE

10. The MidYIS project offers baseline testing in Years 7, 8 and 9. The number of independent schools involved in the Year 8 project is small, so this paper will focus on the Year 7 project and to a lesser extent (for reasons which will become clear at a later stage) on the Year 9 project.

11. The MidYIS Year 7 data for GCSE 2007 underlying the analyses that follow includes pupils and schools as shown in the following tables.

Table 1: Numbers of pupils in the Year 7 MidYIS cohort

<table>
<tr><th></th><th></th><th>Boys' schools</th><th>Girls' schools</th><th>Mixed schools</th><th>Total</th></tr>
<tr><td rowspan="3">Independent</td><td>Boys</td><td>4151</td><td>--</td><td>4901</td><td>9052</td></tr>
<tr><td>Girls</td><td>--</td><td>7835</td><td>3339</td><td>11174</td></tr>
<tr><td>Total</td><td colspan="2">11986</td><td>8248</td><td>20234</td></tr>
<tr><td rowspan="3">Maintained</td><td>Boys</td><td>3876</td><td>--</td><td>32901</td><td>36777</td></tr>
<tr><td>Girls</td><td>--</td><td>5323</td><td>30929</td><td>36252</td></tr>
<tr><td>Total</td><td colspan="2">9199</td><td>63882</td><td>73081</td></tr>
</table>

Table 2: Numbers of schools represented in the Year 7 MidYIS cohort

	Boys' schools	Girls' schools	Mixed schools	Total
Independent	51	164	148	363
Maintained	32	45	429	506

Table 3: Average pupil numbers by sector and school type

		Single-sex schools	Mixed schools
Independent	Boys	81.4	33.1
	Girls	47.8	22.6
Maintained	Boys	121.1	76.7
	Girls	118.3	72.1

12. In the mixed school column of *Table 1*, the pupil totals exceed the sums of the numbers of boys and girls. This is because of the inclusion in the mixed school totals of small numbers of pupils whose gender was not specified. The figures shown above include only those pupils with MidYIS GCSE

feedback to report: it is known that some schools used the MidYIS baseline test for this year group but did not submit their GCSE results to obtain the "value-added" feedback. It should perhaps be explained that the data available to the author does not permit the identification of schools (or pupils) by name.

13. Notable features of the independent sector figures in the above tables include (a) the uneven gender representation in the sample; (b) the high proportion of pupils – girls especially – in single-sex schools; (c) the higher average year group size of boys' schools as compared with girls' and mixed schools; (d) the uneven gender representation, on average, in mixed schools; (e) the much smaller average year group size in comparison with the maintained sector.

14. The GCSE 2006 MidYIS Year 7 cohort was of very similar size and displayed the same general characteristics.

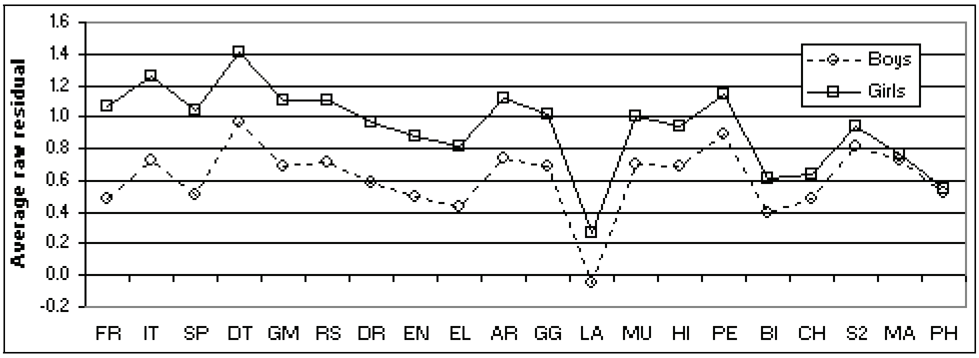
THE YEAR 7 MIDYIS SAMPLE – NATIONAL STANDARDISATION

15. The MidYIS project provides two standardisations of the baseline test and hence two separate computations of the GCSE outcome in relation to ability (or "value-added" outcome).

16. The first of these – the nationally representative standardisation (national standardisation for short) – includes all the pupils in the MidYIS sample from both sectors. The *regression line* (referred to in paragraph 9 above) is computed on the basis of a weighting of the pupil sample which remodels the sample to make it nationally representative. The actual Year 7 MidYIS sample is not nationally representative, principally because the independent sector is over-represented. The weighting ensures that *residuals* are calculated as though the sample were nationally representative.

17. The following chart compares the performance of boys and girls in independent schools as measured by the national standardisation. The performance indicator is the *average raw residual* (ARR) for each subject entry (as explained in paragraph 9).

Figure 3: GCSE 2007 – average raw residuals (ARRs) by subject, independent school pupils



Subjects are arranged here in descending order of the difference between the girls' and boys' ARRs. The meanings of the subject codes should be obvious in most cases. EL is English Literature, GG is geography, S2 is double award science.

18. Regardless of gender differences, the magnitudes of the positive residuals plotted on these charts establish as a background to this research the superior value added achieved on average by independent school pupils as compared with their maintained sector counterparts.

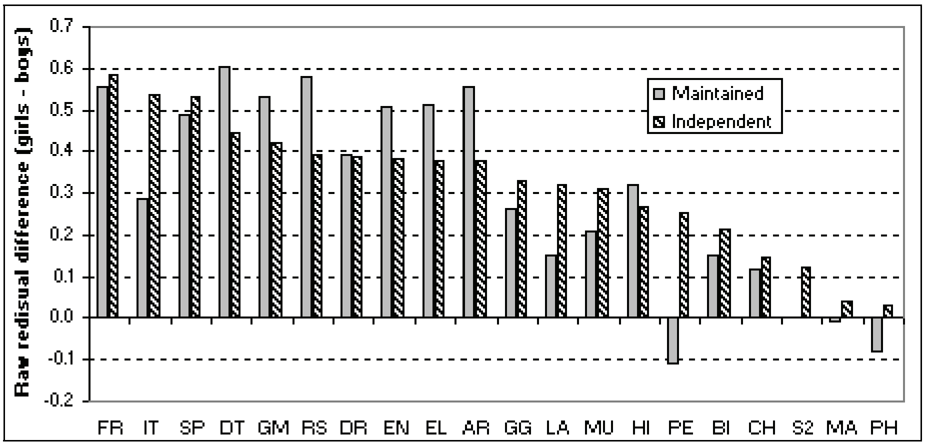
19. That being said, the position in Latin requires clarification. In the 2007 GCSE MidYIS cohort, just over 85% of Latin entries were from independent school pupils, so the value-added computation included relatively few maintained sector pupils. Moreover, the maintained sector representation consisted of pupils whose average MidYIS score and raw results were very little different from candidates in the independent sector. (Nationally, the independent sector accounted for two-thirds of the entries for "Classical Studies", there being no separate statistics for Latin, Greek and Classical

Civilisation.) Broadly similar considerations, though with less dramatic effect, apply to the separate sciences.

20. The independent sector gender differences measured by this method are clear to see. It is probably not unexpected that the gender differences are smallest in Mathematics and the Sciences, and perhaps more surprising that History is towards the lower end of the scale.

21. Raw residuals in the subjects included in *Figure 3* were now computed for maintained school pupils. *Figure 4* below shows gender differences (i.e. the average raw residual for girls minus the ARR for boys) in the two sectors.

Figure 4: Average raw residual differences (girls minus boys) by subject and sector



22. The extent of the similarity between the two distributions is hardly surprising. In fact, the correlation coefficient for the twenty pairs of gender difference values plotted here was 0.80.

Variation of Value Added Across the Ability Range

23. As explained in paragraphs 7 - 9, the MidYIS value-added measurement provides a performance indicator which takes into account pupils' abilities and is thus independent of ability. This is only true for the sample as a whole (or, strictly speaking for the weighted sample on which the value-added measurement is based).

24. For the purposes of the next investigation, the independent sector pupils in the 2007 GCSE MidYIS Year 7 sample were divided into three ability bands (A - C, A being the highest) of roughly equal numerical size, based on their MidYIS baseline test scores. The division was carried out regardless of entry for any particular GCSE subject. Details of the division are shown in the following table.

Table 4: Division of independent sector pupils into three ability bands – GCSE 2007

Band	Score range	No. boys	No. girls	No. pupils
A	>125.1	3295	3510	6807
B	>114.6 <=125.1	2972	3754	6730
C	<=114.6	2785	3910	6697

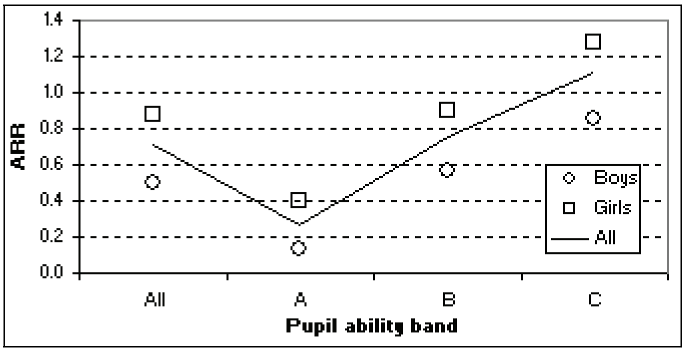
25. An exact division into numerically equal bands is precluded by the numbers of pupils with identical MidYIS scores. It is interesting that neither the boys nor the girls are uniformly distributed between the bands. The final column in this table includes pupils in mixed schools whose gender was not specified in the source data.

26. It is clear from the score ranges shown in the above table that the independent sector representation in the cohort has an ability profile that is well above the national average. Without doubt, the relatively high average ability of independent school pupils is the principal reason for their relatively high average raw GCSE results. However, as illustrated by *Figure 3*, the performance of independent

school pupils is also, on average, better than the performance of maintained sector pupils in relation to ability.

27. Taking English as an example, ARRs were calculated for each of the three ability bands. The results of this exercise are shown in the next chart.

Figure 5: GCSE English – ARRs by ability band, independent school pupils



The line on the chart represents the ARR values for all independent school pupils in each band.

28. This casts a certain amount of doubt on the reliability of the conclusions drawn from *Figure 3* above. It appears that in this subject the value-added measure is not independent of ability for independent school pupils.

29. Since the average value added of independent school pupils measured by the national standardisation cannot now be assumed to be independent of ability, the data represented by *Figure 3* only supports the inferences drawn from it in terms of gender difference if it can be assumed that the ability profiles of the boys and girls entering for a particular subject are very similar. This may be the case, but *Figure 5* suggests that a less vulnerable approach to investigation of independent sector gender differences should be pursued.

30. The variation exposed by *Figure 5* is probably accounted for by two factors. First is the "ceiling effect", the restriction of the value added achievable by the ablest pupils by virtue of the unavailability of grades higher than A*. Secondly, it seems very likely that independent sector pupils in the lower reaches of the sector's ability range perform particularly well in value added terms.

31. A more detailed scrutiny of the data underlying *Figure 5* revealed a correlation coefficient of -0.43 between the residuals and MidYIS scores. Other subjects investigated for the same effect were French (-0.24), Mathematics (-0.41) and Double Science (-0.31), suggesting that the variation of average residuals with ability is likely to be prevalent across the range of subjects.

THE YEAR 7 MIDYIS SAMPLE – INDEPENDENT SCHOOLS' STANDARDISATION

Characteristics

32. In the alternative standardisation, which includes only pupils in independent schools, the baseline test is again standardised to a mean score of 100 and standard deviation 15, but a score of 100 is now the average for pupils in the independent sector.

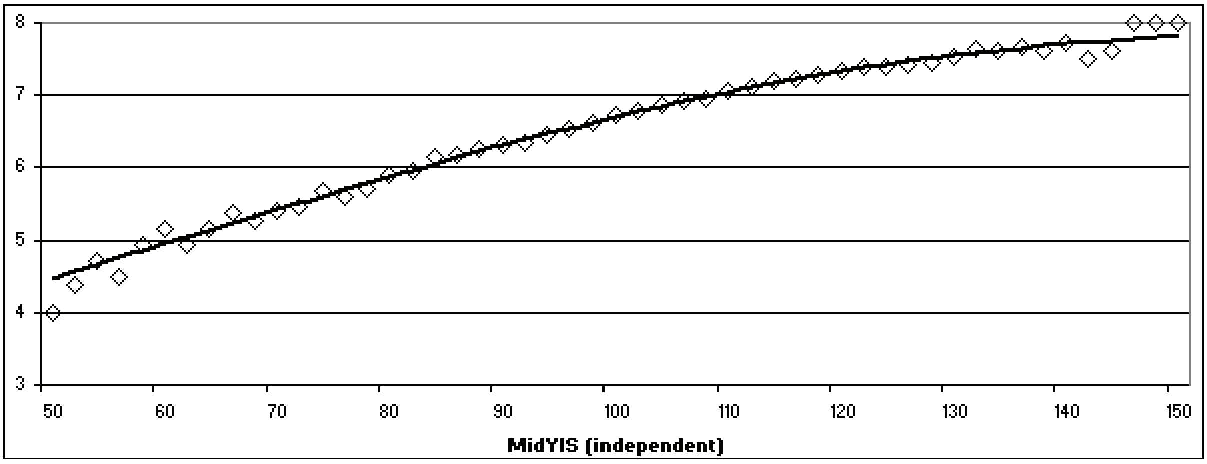
33. The computation of value added differs from that of the national standardisation in two ways.

    (i) The sample is not weighted, with the implication that value added is measured relative to the standard achieved by the pupils in the sample. Strictly speaking, this is not the same as a measurement relative to the sector as a whole unless it is assumed that the sample is broadly representative of the sector.

    (ii) Whereas linear regression is employed in the nationally representative standardisation, logarithmic regression curves are used in the independent schools standardisation. These curves provide a better fit for the data at the upper end of the ability range.

34. The quality of fit for the data provided by the independent standardisation regression lines is illustrated by the chart below, which shows the regression curve for English together with the average grades achieved by pupils within narrow MidYIS score ranges. For example, the plotted average grade for score 100 includes pupils in the score range 99 – 101.

Figure 6: GCSE English 2007 – independent standardisation, regression line and average grades



35. Note that the MidYIS scores here relate to the independent schools' standardisation (independent standardisation for short). The numbers of pupils with scores less than 60 and greater than 140 are, of course, very small indeed.

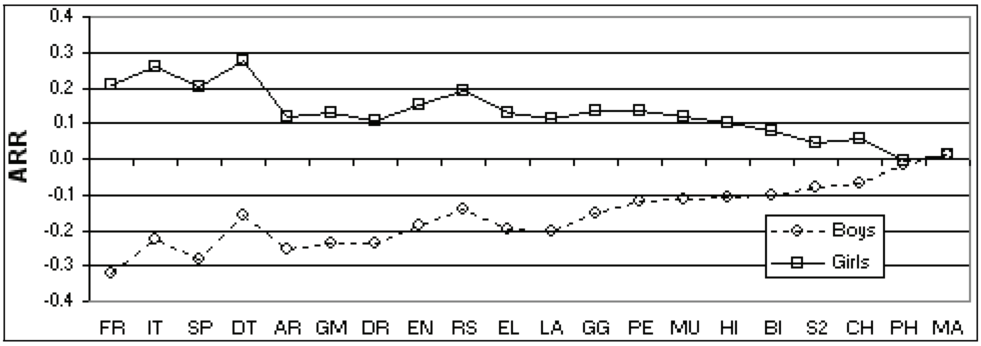
36. The independent standardisation will thus be used as the basis of further comparative investigations that are internal to the sector, on the grounds that there is negligible variation with ability of the average value added measure.

37. This is not to write off the national standardisation, which conveys very important messages to both sectors.

Gender Differences by Subject

38. The following chart is a reworking of *Figure 3*, but with average raw residuals calculated from the independent standardisation. The sample is not weighted, so the ARR for all pupils taking a particular subject will be zero. In other words, if girls have performed better than boys their ARR will be positive, while the boys' ARR will be negative, and the relative magnitudes of the ARRs will depend on the numbers of boys and girls taking the subject. The ARR values shown here cannot be compared with the values computed from the national standardisation.

Figure 7: GCSE 2007 – ARRs by subject, independent schools standardisation



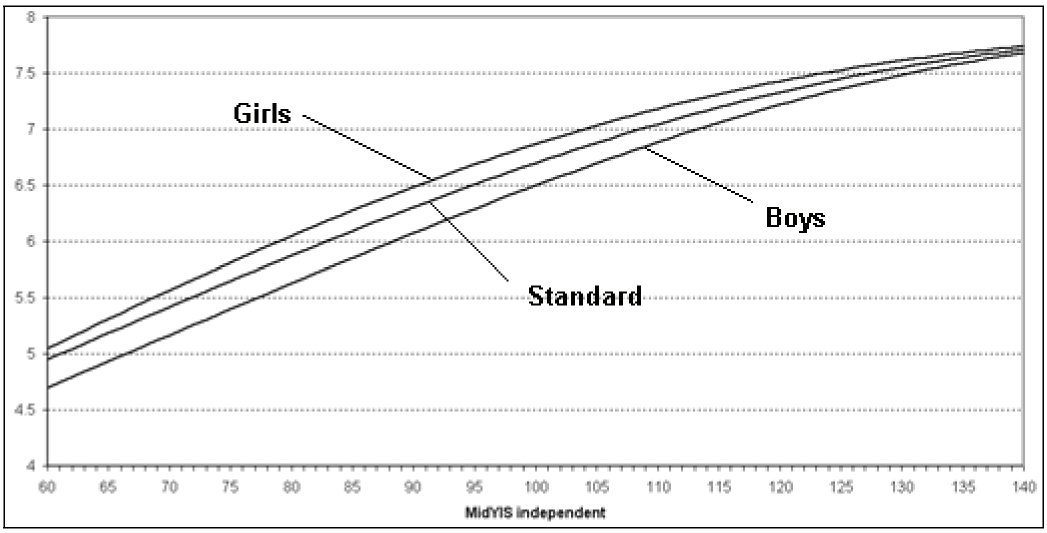
39. As in *Figure 3*, subjects are arranged in descending order of the gender difference.

40. The order of subjects here is not precisely the same as in *Figure 3*. This is a function of the technical differences between the two value-added measurements. In particular, because linear regression is employed in the national standardisation and logarithmic regression curves in the independent standardisation, the relationship between the two separate measurements of raw residual is a non-linear one.

41. The gender differences within a sample taken to be broadly representative of the independent sector are thus readily apparent in the sense that in relation to their ability girls perform better than boys by significant margins in most subjects.

42. As an illustration of the gender gap in an individual subject, separate regression equations for boys and girls were computed and compared with the "standard" regression curve, as shown below.

Figure 8: GCSE English – regression curves for boys & girls separately



43. As expected, given the numerical gender imbalance of the sample, the "standard" regression curve is generally closer to the girls' curve than to the boys' curve. The most important portion of this plot is, of course, in the MidYIS score range 100±15, within which approximately two-thirds of the pupils in the sample will be located.

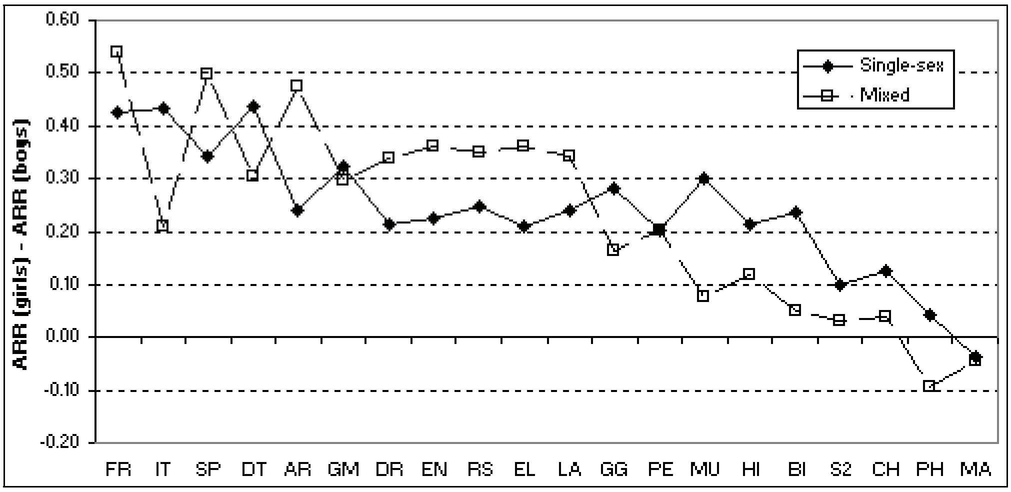
44. An important feature of the above chart is the "category convergence" at the upper end of the ability range, which is a manifestation of the "ceiling effect". The average gender difference is smaller for high ability pupils than it is for lower ability pupils.

Variation of the Gender Gap Between Single-Sex and Mixed Schools

45. The independent sector offers a greater proportionate availability of single-sex education (at least up to GCSE) than does the maintained sector. It cannot necessarily be assumed that the gender gaps in single-sex and mixed schools are similar.

46. The differences between girls' and boys' average raw residuals (independent standardisation) were calculated separately for pupils in single-sex and mixed schools.

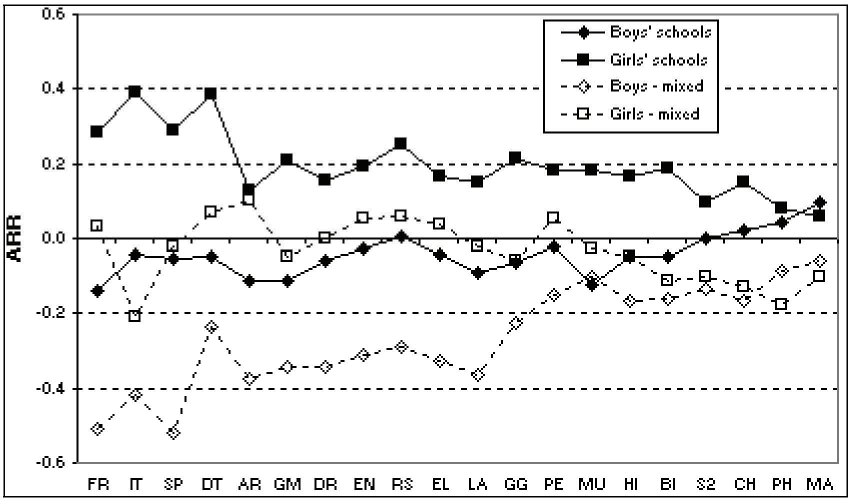
Figure 9: ARR differences by subject in single-sex and mixed schools



47. The order of subjects here is the same as for *Figure 7*. Among the larger subjects, the gender gap is slightly larger in mixed schools in French and English (both), and slightly larger between single-sex schools in the Sciences, History and Geography.

48. It is now very tempting to plot separately the ARRs for boys and girls in single-sex and mixed schools.

Figure 10: ARRs by gender and school type



49. The same order of subjects is retained here. It would appear that in value-added terms (i) the girls' schools were the most successful group, (ii) there was probably little to choose overall between boys' schools and girls in mixed schools, and (iii) the boys in mixed schools achieved the least. This is, of course, an investigation internal to the independent sector: it is probably worth reminding ourselves at this point of the high value added measured in all subjects by the national standardisation (*Figure 3*).

50. It would also appear from this chart that pupils in single sex schools achieved better value added on average than pupils of the same gender in mixed schools. This is probably true, but it would be unwise to rely only on the data underlying the last chart as justification for such an assertion.

51. The value-added measurement on which the last chart is based uses a regression curve for each subject plotted for the independent sector entry for that subject as a whole. Now a comparison of the performance of, say, girls in single-sex and girls in mixed schools is sensitive to the equation of the regression line unless it can be assumed that the two sub-populations have identical ability distributions. This cannot be assumed, either overall or in any individual subject.

Comparing Performance in Single-Sex and Mixed Schools

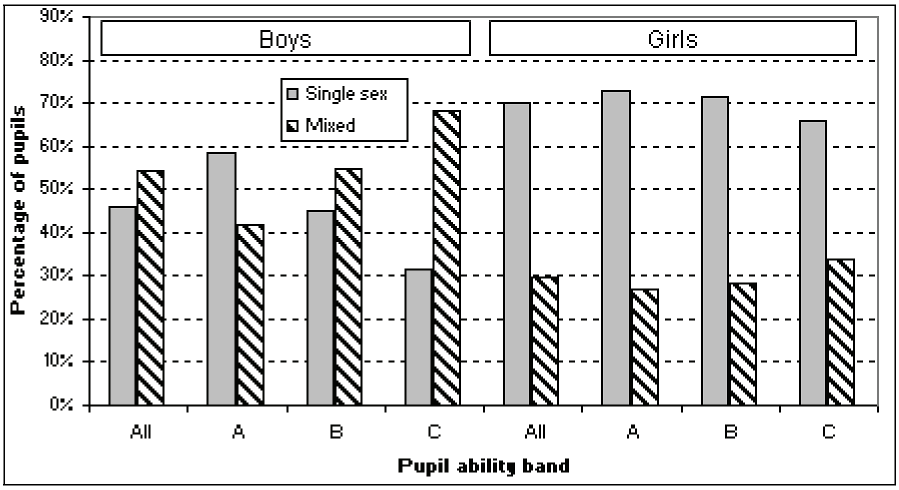
52. For the reasons set out in paragraph 51 above, performance comparisons by gender between single-sex and mixed schools will proceed on the basis of logarithmic regression equations computed separately for boys and girls.

53. For information, the following table and chart show average MidYIS test scores (independent standardisation) by school type overall and by ability band, and the division of boys and girls between single-sex and mixed schools overall and by ability band. The ability band score ranges are the independent standardisation equivalents of those shown in *Table 4*.

Table 7: GCSE 2007 – average MidYIS scores by school type and ability band

		Boys		Girls	
		Boys' schools	Mixed schools	Girls' schools	Mixed Schools
Score range	All pupils	106.0	98.9	101.2	99.1
>107.9	Band A	117.8	116.8	116.9	116.8
> 95.0 <= 107.9	Band B	101.7	101.2	101.4	101.3
<=95.0	Band C	87.0	83.9	85.4	84.6

Figure 11: Percentages of boys and girls in single-sex and mixed schools, Year 7 MidYIS sample



Because percentages are plotted, the chart does not reflect the gender imbalance of the sample as a whole (see *Table 1*).

54. The contrast between the distributions of boys and girls between single-sex and mixed schools is remarkable. This may well be because more boys' schools than girls' schools have become co-educational in the last twenty or thirty years. The chart clarifies the influence of band C boys in mixed schools on the overall ability difference between boys in single-sex and mixed schools. Individual subjects will obviously display different ability variations from those shown in *Table 7* (which is why it is prudent to persist with the use of gender-separated regression equations).

55. For each of the subjects selected for analysis in *Figures 3 and 7*, separate regression equations were computed for boys and girls, using the MidYIS independent standardisation test scores as the x-values. Average raw residuals (ARRs) were then computed for boys and girls in single-sex and mixed schools

and the ARR differences (single-sex – mixed) were calculated.. The following charts convey the results of this exercise:

Figure 12: Average raw residual differences (single-sex minus mixed) for boys

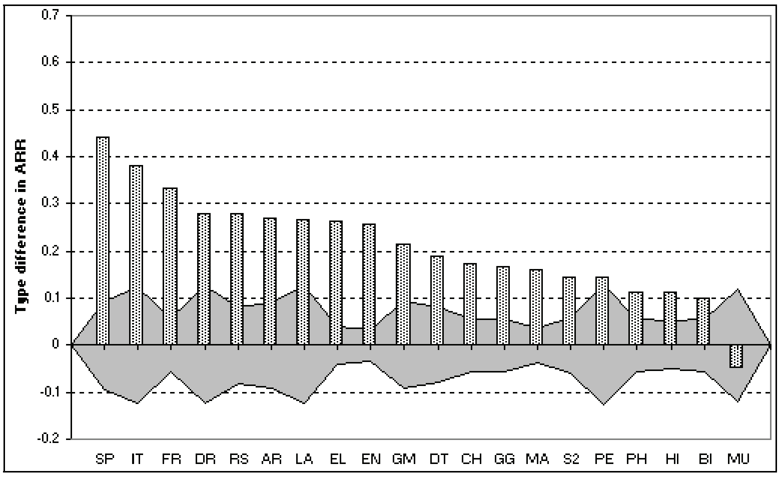
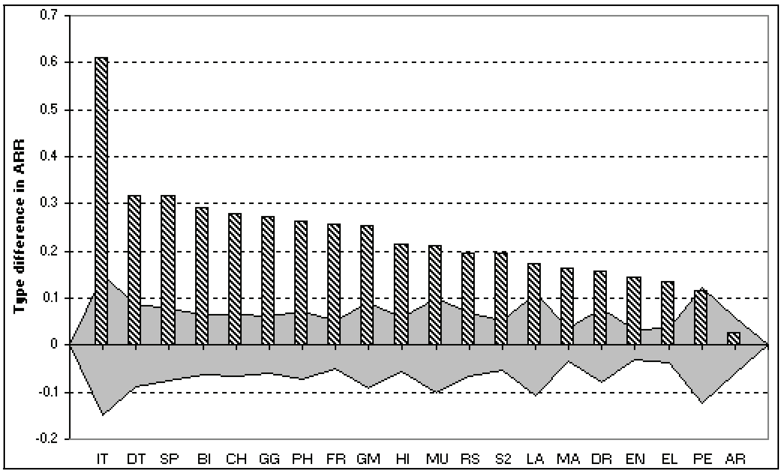


Figure 13: Average raw residual differences (single-sex minus mixed) for girls



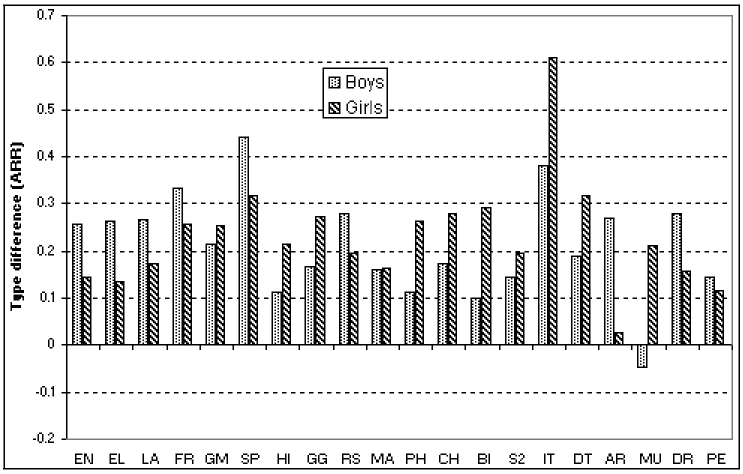
56. In each case, subjects are arranged in decreasing order of the difference between the ARR for pupils in single-sex schools and the ARR for pupils in mixed schools (the "school-type difference" for short). It will be noted that the same y-axis scale is used for both charts.

57. The grey areas on the charts are 95% confidence limits, delineating values which if not exceeded indicate that there is less than 95% statistical confidence that the reported difference has not arisen by chance. It is clear that virtually all of the differences shown on the charts pass this test. The confidence limits are not related to the magnitudes of the reported difference values. In general, they are widest for subjects with small entries and narrower for subjects (such as English and Mathematics) with large entries.

58. Almost without exception, when ability variation is taken into account (which it is by virtue of the methodology employed here), the average performance of pupils of both genders in single-sex schools is better than the average performance of their counterparts in mixed schools by a significant margin. The margins are, however, fairly small in most cases, varying from negligible to about 0.3 grade, with hardly any subjects exceeding this value.

59. The next chart plots the same school-type differences as shown above, but the values for boys and girls are plotted together.

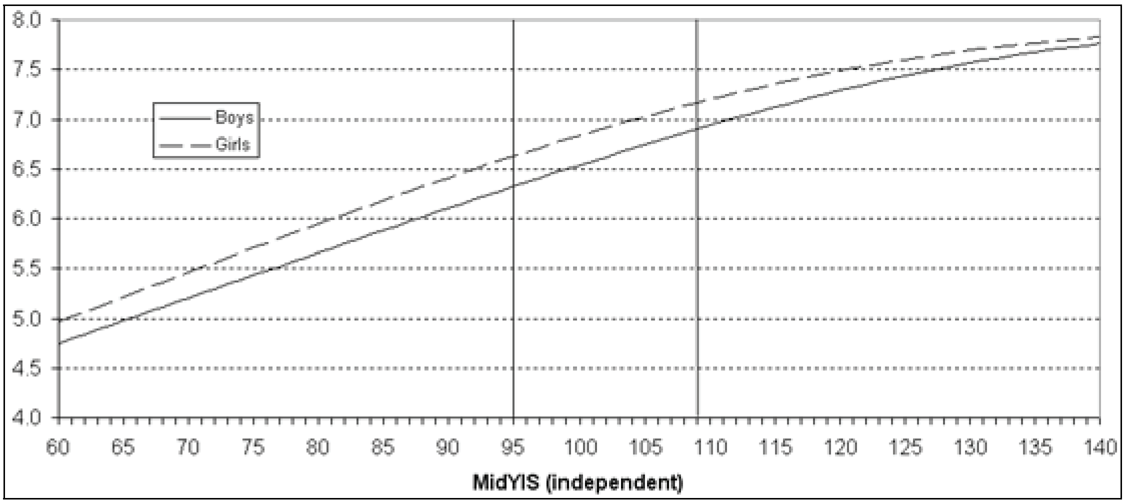
Figure 14: Average raw residual differences (single-sex minus mixed) for girls and boys



60. It is difficult to ascertain whether, overall, the school-type effect is more significant for boys than it is for girls. The type difference for boys is greater than for girls in ten of the twenty subjects considered here, including English (both) and French, to mention subjects with the larger entries. However, the type difference is greater for girls in all four Science subjects, History and Geography.

61. Regression lines were now computed separately for boys and girls using as the output measure the average of each pupil's grades in all the full-course GCSE subjects covered by the data. One or two subjects with small entries not selected for the earlier analyses are included, but a small number of such subjects are not catered for. No short course GCSE results are included. The two regression lines are shown below. The vertical lines delineate the three pupil ability bands as defined earlier.

Chart 15: Average points per subject regression lines for boys and girls

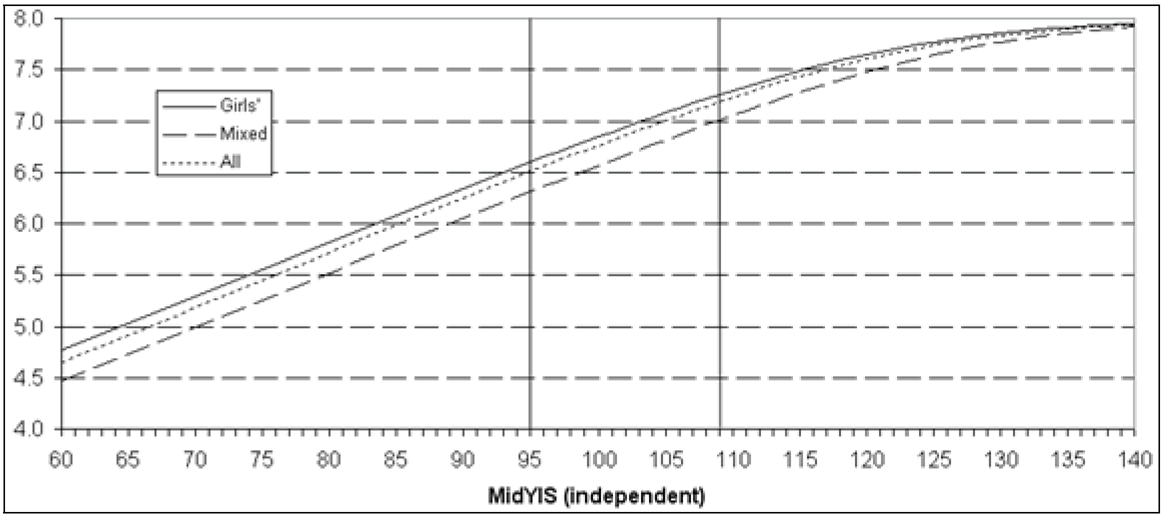


62. The chart is an interesting representation of an estimate of the GCSE 2007 overall gender gap, which is 0.22 grade at score 60, reaching a peak of 0.30 in the score range 97 to 101 and then tailing off to virtually nothing at the top end of the ability range. The convergence of the two curves at the upper end of the ability range is a result of the ceiling effect.

63. Separate average raw residuals for boys and girls in single-sex and mixed schools were then computed on the basis of these regression equations. The result was a school-type difference of 0.21 for boys and 0.20 for girls. Despite the unreliability of this experiment resulting from variations in subjects taken by different pupils and, to a lesser extent, the inclusion and exclusion of particular small-entry subjects, the indication is that the school-type effect has roughly equal overall magnitudes for boys and girls in the Year 7 MidYIS sample.

64. A further question is whether school-type differences vary across the ability range. It is almost inevitable that as measured by this method they will. Consider the following chart, which shows separate regression lines for girls' schools, girls in mixed schools and all girls, the subject being French.

Figure 16: GCSE French – regression lines for girls



65. The vertical lines again represent the cut-off scores for the three pupil ability bands. The regression curve for all girls is closer to the curve for girls in single-sex schools, reflecting the fact that 70% of the girls who took GCSE French did so in girls' schools. The ceiling effect convergence is again evident here, and causes the type difference for band A pupils (0.19) to be less than for band B and band C pupils (0.29 and 0.28 respectively). The effect illustrated here is likely to apply to most if not all subjects. In English, the type differences for girls in bands A, B and C were 0.02, 0.21 and 0.19 respectively, so the ceiling effect convergence would appear to have had even greater influence here.

Commentary on the Findings of the Single-Sex vs. Mixed School Investigation

66. The conclusions established above (and pursued further later in this paper) are not intended as an assertion that single-sex schools provide a better education than mixed schools. All that has been said so far is that both boys and girls in the GCSE 2007 MidYIS Year 7 independent school sample who attended single-sex schools were slightly more successful in GCSE than their counterparts in mixed schools once any variation in ability profiles has been controlled for. It is not an objective of this paper to provide an explanation of what has been established.

67. There is far more to education than examinations, and far more to examinations than GCSE. Nevertheless, GCSE is the first qualification based on public examinations that secondary school pupils acquire. It is also, for many pupils, an important factor in determining where their further education will be acquired and what its content will be. Moreover, the acquisition of effective study habits in the run-up to GCSE is relevant to whatever courses of study pupils follow after the age of 16.

68. The idea of a post-16 study of the relative achievements of pupils in single-sex and mixed institutions is attractive. After all, A-level (or in an increasing number of cases the International Baccalaureate) is the end product of most independent secondary schools. Unfortunately, a study of 18+ relativities with the same reliability as these researches is not possible.

69. In order to allow for ability profile variations, this paper, as explained earlier, uses a "value added" approach to the measurement of performance. Why not, therefore, do the same for 18+ qualifications?

70. The study reported in paragraphs 52 - 63 has the advantage that all pupils whose data contributed to the analyses took the same aptitude test at roughly the same time. The same applies to the 2007 GCSE Year 9 pupils (see below). MidYIS measures value added relative to pupils' aptitude in all cases and so, therefore, does this paper.

71. The Alis (A-level Information System, also managed by the CEM Centre) measures value added in the sixth form. However, it has two (arguably three) different baselines. For historical reasons, the preferred baseline is the average points per subject achieved by a pupil in GCSE, which is undoubtedly the best available "predictor" of A-level grades across the Alis sample as a whole (which accounts for about half the A-level entries in the UK).

72. There are two problems, though, where the independent sector is concerned. The first is that pupils in independent schools achieve value added to GCSE which is markedly above the national average. This inflates the preferred Alis baseline and thus diminishes the value added measured in the sixth form, effects which differ in magnitude between boys and girls.

73. Alis does offer a baseline paper test (the "Test of Developed Ability" or TDA for short) which admits the possibility of measuring value added in the sixth form based on an aptitude measurement as in MidYIS. However, the TDA was taken by only just over half of the independent school students in the Alis project for A level taken in 2006. 286 schools were represented in that cohort, and the TDA was used in 200 of them. However, some schools used the TDA only for pupils joining in Year 12. The computer-adaptive version of TDA is used by some schools, resulting in a degree of fragmentation of the aptitude measurement baseline.

74. Appreciable numbers of schools which are single-sex up to Year 11 have co-educational sixth forms, a factor which reduces the potential value of any investigation of school-type differences in the sixth form.

Gender Differences in Subject Entries

75. While gender differences in outcomes are clearly the major point of this paper, gender and school-type differences between the numbers of subject entries in particular curricular areas may be of some interest. The focus here will be on Science and Modern Languages in the Year 7 MidYIS independent school sample.

Table 8: GCSE percentage entries in Science subjects by gender, school type and ability band

		Single-sex schools			Mixed schools			All schools		
Subject	Band	Boys	Girls	All	Boys	Girls	All	Boys	Girls	All
S2	All	44%	67%	59%	54%	57%	56%	50%	64%	58%
	Band A	34%	58%	47%	41%	47%	43%	37%	55%	46%
	Band B	47%	67%	60%	54%	55%	55%	51%	64%	58%
	Band C	62%	77%	73%	64%	67%	65%	64%	73%	69%
BI	All	52%	29%	37%	39%	38%	39%	45%	32%	38%
	Band A	61%	40%	49%	54%	51%	53%	58%	43%	50%
	Band B	50%	30%	37%	41%	40%	41%	45%	33%	38%
	Band C	34%	17%	21%	27%	27%	27%	29%	20%	24%
CH	All	50%	28%	35%	39%	33%	36%	44%	30%	36%
	Band A	60%	40%	49%	56%	47%	52%	58%	42%	50%
	Band B	48%	29%	35%	41%	38%	40%	44%	31%	37%
	Band C	31%	15%	19%	24%	20%	22%	26%	17%	21%
PH	All	50%	27%	35%	39%	32%	36%	44%	28%	35%
	Band A	60%	39%	48%	56%	47%	52%	59%	41%	50%
	Band B	48%	27%	34%	41%	36%	39%	45%	30%	36%
	Band C	31%	15%	19%	25%	17%	22%	27%	16%	20%

76. As expected, the uptake of the separate Sciences decreases towards the lower end of the ability range, while the percentage entries for Double Award (S2) increase. It is probably not surprising that the percentage uptake of the separate Sciences is lower in girls' schools than in boys' schools.

77. It is interesting that the percentages of girls taking separate Sciences are greater in mixed schools than in girls' schools. This may relate to the slightly greater school-type differences for girls in separate Sciences as compared with Double Award.

78. The figures shown in this table are governed only partially by pupils' choice. In some "mixed Science economy" schools, pupils are directed towards one type of Science curriculum or the other. A good many schools (just over 100 of the363 in the sample) do not offer the separate Sciences. A minority (about 40) offer only the separate subjects but without necessarily requiring pupils to take all three.

79. The next table analyses entries for the three Modern Foreign languages selected for detailed analysis.

Table 9: GCSE percentage entries in Modern Languages by gender, school type and ability band

		Single-sex schools			Mixed schools			All schools		
Subject	Band	Boys	Girls	All	Boys	Girls	All	Boys	Girls	All
FR	All	70%	67%	68%	59%	67%	62%	64%	67%	66%
	Band A	74%	69%	71%	66%	71%	68%	71%	69%	70%
	Band B	68%	67%	68%	62%	70%	65%	65%	68%	67%
	Band C	65%	65%	65%	51%	62%	56%	55%	64%	60%
GM	All	24%	19%	21%	22%	21%	22%	23%	20%	21%
	Band A	27%	24%	25%	30%	25%	28%	29%	24%	26%
	Band B	24%	20%	21%	24%	21%	23%	24%	20%	22%
	Band C	16%	15%	15%	16%	18%	17%	16%	16%	16%
SP	All	27%	30%	29%	22%	29%	25%	24%	29%	27%
	Band A	29%	33%	32%	27%	33%	29%	28%	33%	31%
	Band B	27%	32%	30%	22%	31%	26%	24%	32%	28%
	Band C	23%	24%	24%	17%	23%	20%	19%	24%	22%
All	All	122%	116%	118%	103%	117%	109%	112%	116%	114%
	Band A	131%	126%	128%	123%	129%	126%	128%	127%	127%
	Band B	119%	119%	119%	108%	122%	114%	113%	120%	117%
	Band C	105%	104%	104%	84%	104%	92%	91%	104%	98%

80. The last four rows of this table are produced by adding together the entries for the three subjects analysed, figures in excess of 100% reflecting the extent to which pupils take more than one of these subjects. In practice, slightly higher proportions than represented here will offer more than one Modern Foreign Language, because other languages offered in some schools are not included here.

81. It is not surprising that French remains the most popular of these subjects, or that Spanish is more popular than German. Percentage uptakes tail off slightly towards the lower end of the ability range, the effect being most marked among boys in mixed schools.

82. The percentage uptakes for these three subjects in the maintained sector were French 29%, German 12% and Spanish 8%. It is, of course, no longer a requirement of what remains of the National Curriculum that a Modern Foreign Language should be studied up to Year 11.

THE YEAR 9 MIDYIS SAMPLE

Characteristics

83. The independent sector representation in the sample is set out in the following table.

Table 10: Numbers of independent school pupils in the Year 9 MidYIS cohort

	Boys' schools	Girls' schools	Mixed schools	Total
Boys	4524	--	3891	8415
Girls	--	2131	2354	4485
Total	6655		6249	12904

84. In terms of independent school pupil numbers, the Year 9 project is appreciably smaller than Year 7. It is also the case that 38% of pupils in the Year 9 sample were also tested inYear 7, so the numbers of pupils new to MidYIS in Year 9 were smaller than those shown in the above table. Numbers of pupils new to MidYIS in Year 9 are shown in the next table.

Table 11: Numbers of independent school pupils new to MidYIS in the Year 9 sample

	Boys' schools	Girls' schools	Mixed schools	Total
Boys	2735	--	3026	5761
Girls	--	589	1630	2219
Total	3324		4660	7984

85. It is immediately seen that the gender imbalance of the Year 7 sample is reversed, especially when pupils also tested in Year 7 are discounted.

86. 238 schools were represented here (55 boys' schools, 67 girls' schools and 116 mixed). However, some schools were involved in both the Year 7 and Year 9 projects, and some of these used the Year 9 test only for their "top-up" entries after Year 7. It appears that just 49 schools were new to the Year 9 project (14 boys' schools, 2 girls' schools and 33 mixed schools).

87. The following table comparing average MidYIS scores (independent standardisation) of the two populations making up the Year 9 sample may be of interest.

Table 12: Average MidYIS scores – pupils tested in Year 9 only and in Years 7 and 9

	Single-sex schools			Mixed schools			All schools		
	Boys	Girls	All	Boys	Girls	All	Boys	Girls	All
Y9 only	102.8	98.7	102.1	94.3	98.9	95.9	98.3	98.8	98.5
Y7 & Y9	105.4	103.6	104.6	99.4	101.2	100.2	103.5	102.9	103.2

88. The Year 9 newcomers were clearly slightly less able on average than the pupils who had also taken the Year 7 test. The next table uses the Year 7 independent standardisation to compare average Year 7 MidYIS scores with the average scores of pupils who were also tested in Year 9.

Table 13: Year 7 MidYIS averages overall and for pupils also tested in Year 9

	Single-sex schools			Mixed schools			All schools		
	Boys	Girls	All	Boys	Girls	All	Boys	Girls	All
All Y7	106.0	101.2	102.9	98.9	99.1	99.0	102.2	100.6	101.3
Y7 also tested in Y9	104.9	100.6	102.8	99.0	96.8	98.1	102.7	99.5	101.2

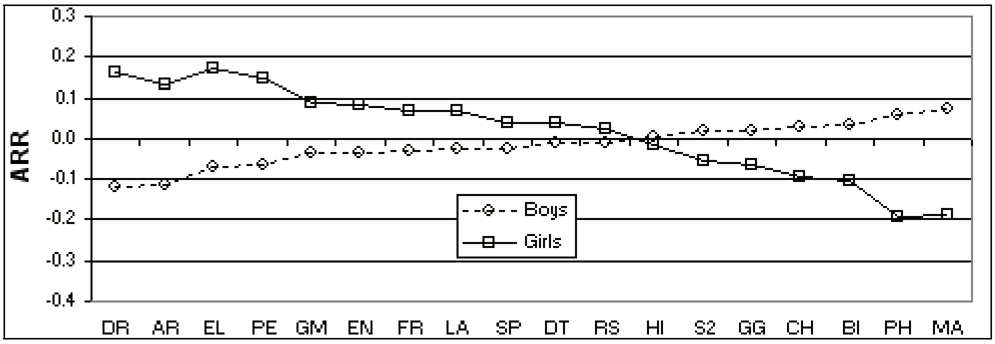
89. It appears that the pupils tested in both years were broadly representative of the Year 7 cohort, and that their average MidYIS scores were very slightly higher in the Year 9 test.

90. It was explained earlier that the computation of value added in the independent standardisation of the Year 7 data does not involve weighting of the sample, so that value added is measured relative to whatever independent school sample subscribes to the project. The same applies to the Year 9 data. It would appear likely, on the basis of the figures presented and bearing in mind other aspects of the ways in which different independent schools organise themselves, that the value-added measurements from Year 7 and Year 9 represent comparisons with slightly different independent school samples.

Gender Differences by Subject

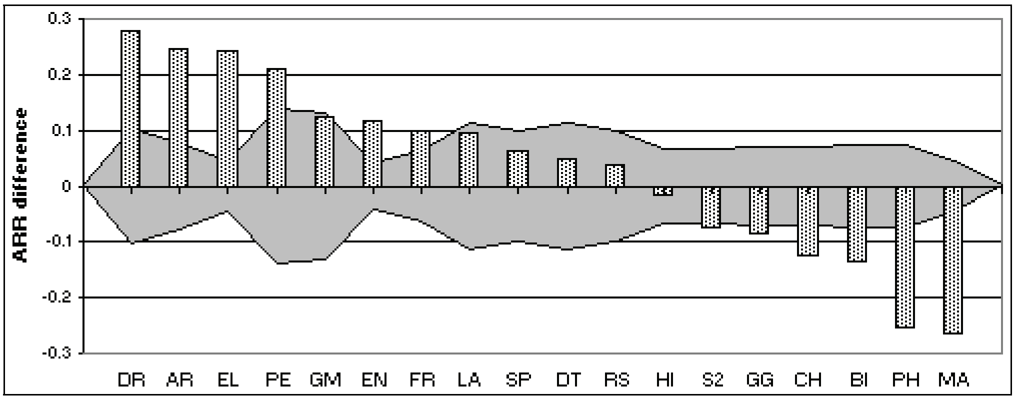
91. This investigation initially excluded the pupils who had been tested in Year 7 on the grounds that they had been covered in the Year 7 analyses. It was therefore necessary to compute a regression equation for each subject including only those pupils new to MidYIS in Year 9. Average raw residual (ARR) values based on these regression equations are plotted on the following chart:

Figure 17: ARRs by subject, Year 9 MidYIS excluding pupils tested in Year 7



92. The method used here dictates that the ARR for all pupils in any subject is very close to zero. The smaller magnitudes of the boys' ARRs reflect the gender imbalance of the sample. It is immediately clear that the gender gaps (i.e. the differences between the ARRs for girls and the ARRs for boys) are appreciably more favourable to boys than the gender gaps in the Year 7 sample (see *Figure 7*). (It will be observed that two GCSE subjects, ICT and Music, that were covered in the Year 7 analysis are excluded here on the grounds that numbers of entries were small.) The Year 9 gender gaps are plotted on the next chart:

Figure 18: Gender differences (girls' minus boys' ARRs), pupils tested in Year 9 only



93. The grey areas are again 95% confidence limits for the ASR differences recorded. It is seen that six of the differences do not pass the 95% confidence test. A similar plot was not provided for the Year 7 sample because the gender gaps were appreciably greater (and in almost every case in the girls' favour), and the numbers of pupils were greater (reducing the size of the confidence limits).

94. In the Year 7 sample, only five subjects produced gender gaps less than 0.2 (Mathematics, Physics, Chemistry, Biology and Double Science), with just the first two of these subjects having gender gaps less than 0.1. Here, there are only four positive values greater than 0.2, and boys show a significant advantage in six subjects.

95. The next investigation reverted to the regression equations provided with the standard MidYIS GCSE feedback. (For this reason, ARR values for the Y9-only pupils will not be exactly the same as shown in *Figure 17*.) The aim was to compare the performance of the two sub-groups in the Year 9 cohort, namely pupils also tested in Year 7 and those not tested before Year 9. ARRs were plotted by subject, retaining the same order of subjects as above and the y-axis scale of *Figure 17*.

Figure 19: ARRs by subject, boys only, tested in Y7 and Y9 and in Y9 only

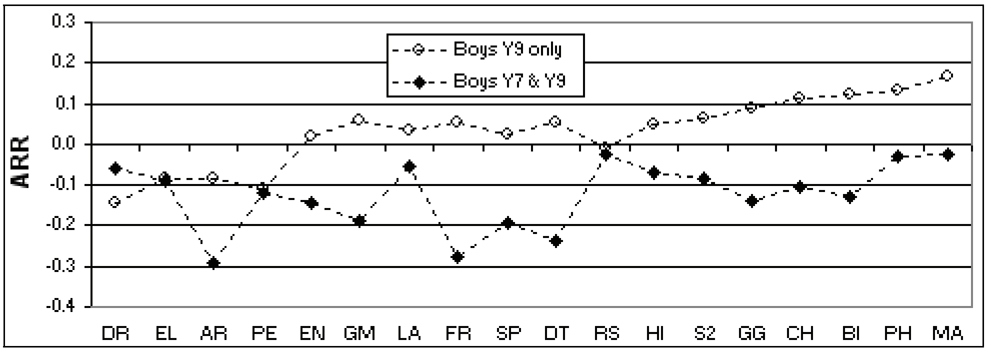
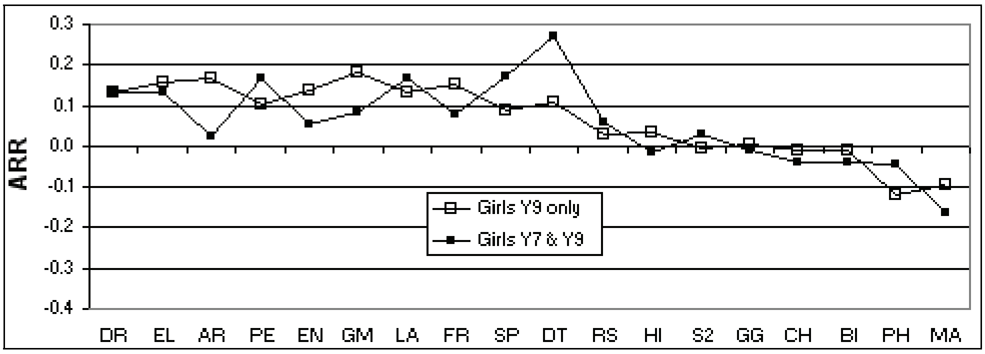


Figure 20: ARRs by subject, girls only, tested in Y7 and Y9 and in Y9 only



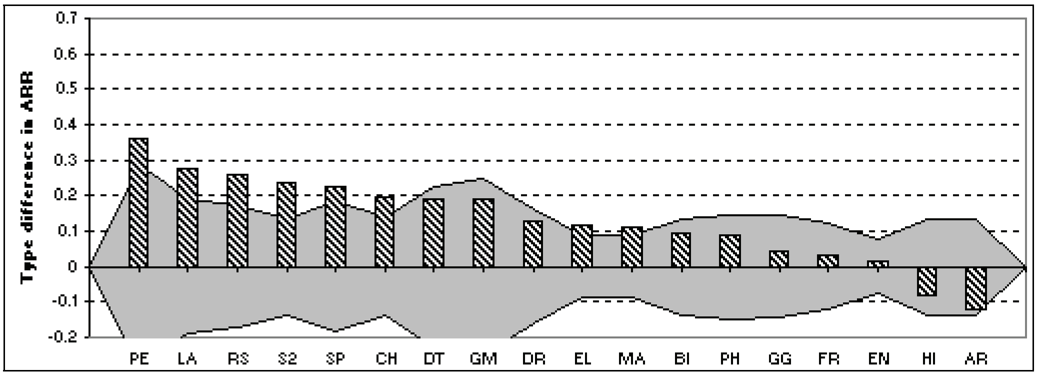
96. The investigation reveals only small differences between the two groups of girls but significant differences between the two groups of boys. In other words, the smaller gender differences recorded for the Y9-only group are explained, statistically, by the slightly better performance of Y9-only boys as compared with boys tested in both years.

97. The reason for this may be related to the proportion of boys who, in their preparatory schools, experience a relatively intensive preparation for competitive examinations taken during Year 8, the results of which represent the final outcome of the preparatory phase. These examinations are principally tests of attainment, and it is highly likely that the attainment of many boys leaving preparatory schools is better than average in relation to ability. In other words, they are already achieving positive value added, even though it is not measured. Perhaps preparation for Year 8 examinations gives these boys a slight edge over others who have enjoyed a regime in Years 7 and 8 that does not have a focus on external examinations.

Comparing Performance in Single-sex and Mixed Schools

98. This investigation reverts to the Y9-only sub-group. To begin with, average raw residuals were calculated by subject, using the Y9-only regression equation. Raw residual differences between single-sex and mixed schools were calculated. The following chart shows the outcome for girls:

Figure 21: Average raw residual differences (single-sex minus mixed) for girls

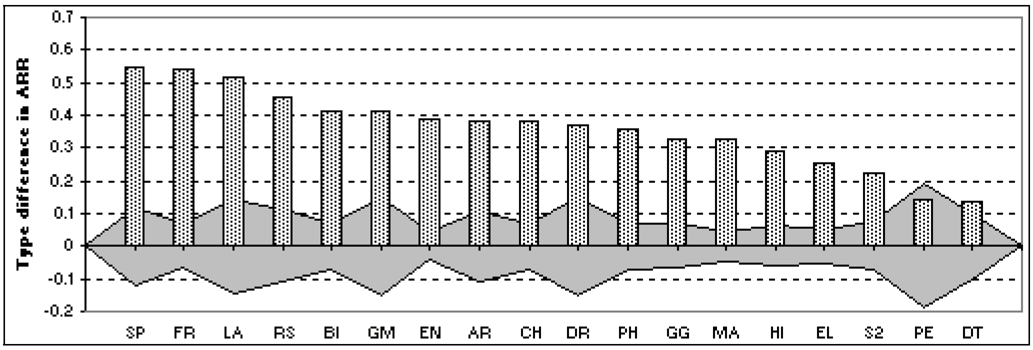


99. The school-type differences are generally smaller than in the Year 7 sample (*Figure 13)*, and there is little correlation between the subject orders on the two charts. More to the point, however, are the magnitudes of the differences in relation to the confidence limits. The situation is not surprising given the small number of Y9-only girls in single-sex schools.

100. For girls at least it is clearly not worth going to the lengths of computing gender-separated regression equations for the subjects in the analysis. It cannot reliably be asserted on the basis of this evidence that girls in this subset of the sample who attended single-sex schools have a significant performance advantage over their counterparts who attended mixed schools.

101. The more detailed analysis, based on gender-separated regression equations for Y9-only pupils, was carried out for boys. The school-type difference chart is shown below:

Figure 22: Average raw residual differences (single-sex minus mixed) for boys



102. Type differences here are generally larger than they were for boys in the Year 7 sample. One is inclined to suggest that in the Year 9 sample boys in single-sex schools did more than slightly better on average than boys in mixed schools. Imagine any likely combination of 10 GCSE subjects: on average, a boy in a single-sex school will be one grade better in three of them than his opposite number in a mixed school.

103. The boys in this sample attended either (a) schools admitting to Year 7 and also to Year 9 and (b) schools that do not admit pupils earlier than Year 9 (although some of these schools have their own preparatory schools or departments). It is possible that the boys' schools in category (b) are making a difference to the type difference represented above. It would be interesting to find out how many of these are boarding schools, or predominantly so.

THE RELATIONSHIP BETWEEN GENDER DIFFERENCES AND SCHOOL-LEVEL AVERAGE VALUE-ADDED DIFFERENCES BETWEEN THE TWO MIDYIS STANDARDISATIONS

The Nature of the Issue

104. Within schools, staff are rightly concerned with the value-added performances of their own subjects and, if they are senior staff, with comparisons between the value-added performances of different subjects. It is sometimes perplexing to school staff to find a subject result in which the value-added measured by the national standardisation looks good or very good but where the same results from the same pupils produce a negative value-added indication on the independent standardisation.

105. It is inevitably the case that in the independent standardisation half the pupils taking a given subject will have negative residuals while the other half will have positive residuals. (If all independent schools try very hard to improve their positions in the independent standardisation value-added league table this will continue to be true.)

106. It is also the case that independent school pupils in the MidYIS sample achieve, on average, better GCSE results in relation to their ability than their maintained sector counterparts, the difference increasing significantly at the lower end of the independent sector ability profile.

107. The preceding two paragraphs constitute the reason why a good value-added performance relative to a nationally representative sample can appear ordinary or even below par in relation to the independent school MidYIS sample. These arguments do not, however, necessarily convince those whose task it is to interpret value-added data in schools.

An Illustration

108. Consider an average mixed school, as defined by the composition of the independent school Year 7 MidYIS cohort. It had 55 Year 11 pupils (33 boys and 22 girls), with an ability profile skewed towards the lower end of the independent sector range.

109. The Director of Studies (DoS, who, it will be assumed, is male) surveys the school's value-added Year 7 MidYIS feedback. For the purposes of comparing its value-added outcomes in different subjects, the DoS will use the *average standardised residuals (ASRs)* provided (to one decimal place only) in the standard MidYIS GCSE feedback. Residuals are standardised to compensate for the varying extents between subjects of "scatter" of points around the regression line. Standardised residuals, which are not grade fractions (unlike raw residuals) must always be used when comparing the value-added averages of different subjects and when comparing value added year on year. They are calculated for each pupil in each subject by dividing the raw residual by the standard deviation of the raw residuals of all the pupils who took the subject.

110. The DoS opens the national standardisation feedback file ASR chart and is gratified to find that the average standardised residual in English is +0.5, which is well outside the 99.7% confidence limit. He sets up a subject report for English, and is mildly surprised to find that many of the best residuals were achieved by the weaker pupils. He is less surprised that some of the abler pupils had only small positive residuals, because the subject scatterplot reminds him that none of them could have scored better than A*.

111. Turning to Mathematics, where the school's average grade was very close to the English average, the DoS finds an ASR of +0.6, which again looks very satisfactory. The less able pupils again seem to have the lion's share of the best residuals. He notices from the subject report that the girls and the boys seem to have achieved about equally in the value added, but turning back to English he finds that the girls did better on average than the boys. He does not find this particularly surprising, given what has appeared in the press about gender differences in national GCSE results.

112. Cheered by what he has seen so far, the DoS decides to look at value added in the independent standardisation. The Mathematics is a little disappointing, with an ASR of -0.1. But he knows that the independent standardisation is a much more rigorous benchmark. The girls and the boys had similar value added on average, perhaps the boys having a very slight edge. There did not seem to be much variation in value added between the weaker and the stronger pupils, which is different from the national standardisation, but the DoS does not worry too much about this.

113. The DoS is less happy with the value added in English, whose ASR is -0.2. Inside the confidence limits, but still negative. Curious, this: the national standardisation ASR (+0.5) was not exactly borderline, and the Head is not going to like anything negative relative to the independent sector as a whole.

114. Despite, or perhaps spurred on by his consternation at this point, the DoS looks up the independent standardisation subject report for English. What he finds adds to his disquiet. The weaker pupils, who had achieved, generally, the better value added in the national standardisation, recorded less good value added than the abler pupils in the independent standardisation. The DoS could, however, see from the scatterplot of results how the negative ASR for the subject had come about.

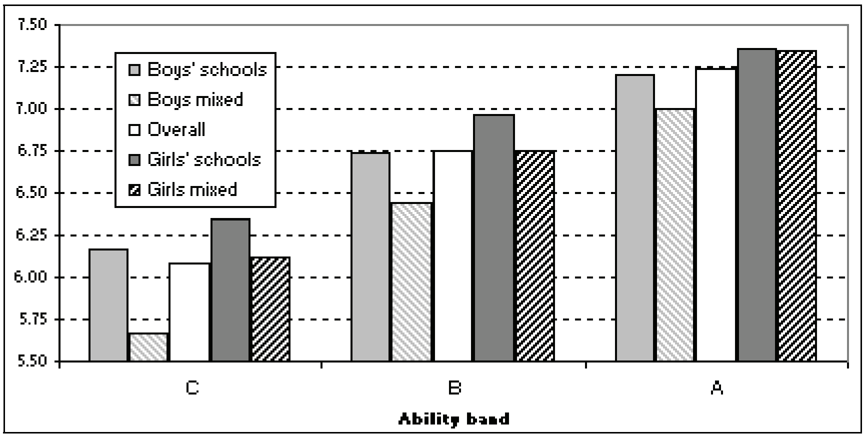
An Attempt at Clarification

115. Can we, with knowledge of the full dataset, provide our DoS with any help in his interpretation of the MidYIS feedback?

116. For a start, we can tell him that the increase of average national standardisation residuals with decreasing ability is commonplace (which is certainly a factor in this notional school, with 40% band C pupils); this does not happen in the independent standardisation, which provides a value-added measure whose average value for independent school pupils in the sample is zero at any point in the ability range. We can confirm that the smaller average residuals of his most able pupils are a manifestation of what we know as the ceiling effect.

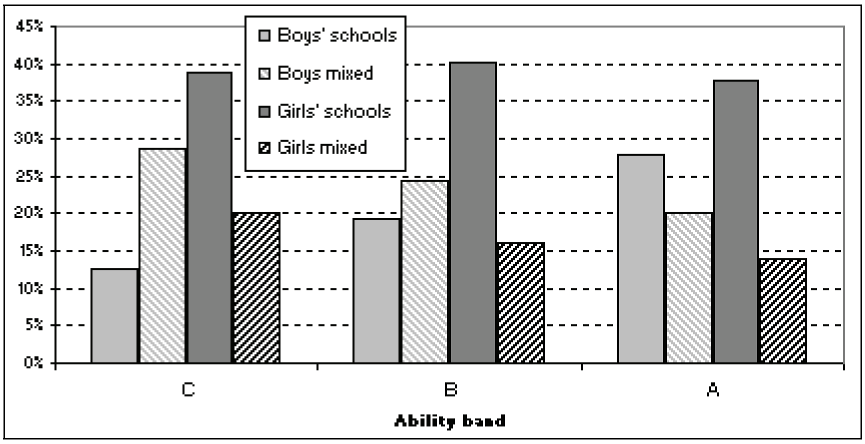
117. Next, we can show him the following diagram, representing the average independent sector results achieved in English, the breakdown being by pupil ability band and gender and school type (single-sex or mixed).

Figure 23: GCSE English results by school type and ability band



118. The unshaded columns represent the average result for all pupils in the ability band. It is easy to see that this plot reflects both the overall gender gap in the subject and the school-type differences investigated earlier, and that it shows the category convergence towards the top of the ability range that has been encountered previously. The next diagram shows the division of the English entry in each ability band by gender and school type.

Figure 24: Breakdown by gender and school type of the pupil population of each ability band



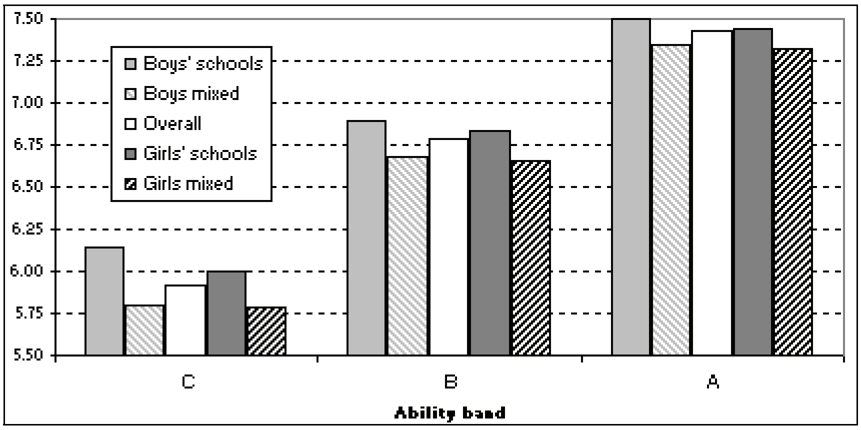
119. If there were equal numbers of boys and girls, divided equally between single-sex and mixed schools, the average grade for each band would simply be equal to the mean of the four category averages. Alternatively, if the average grades for the four pupil categories were the same or nearly so, then differential representation of pupil categories in the band would have little or no effect. Clearly, neither of these conditions applies (in English at least).

120. The pupils in the notional average mixed school performed at an average level in relation to their gender and school type. *Figure 23* shows how the boys' independent standardisation residuals came to be most negative at the lower end of the ability range, and how the girls' residuals were better. What can also be understood from this chart is that the population of band C, say, can legitimately be regarded as consisting of four sub-groups of pupils. The average grade of all band C pupils is a function of the average grades achieved by each of the sub-groups and of the numbers of pupils in each sub-group. The location of the regression curve in this ability range is a consolidation of these variables.

121. It appears probable from the above that mixed schools are on average more likely than single-sex schools to show negative independent standardisation residuals in English, with the likelihood increasing in schools with higher proportions of band C pupils. (In fact, of the 148 mixed schools with candidates for GCSE English, 104 (70%) had negative average residuals and only 25 of these schools had average MidYIS independent standardisation test scores of 100 or greater.)

122. Here is the Mathematics equivalent of *Figure 23*.

Figure 25: GCSE Mathematics results by school type and gender



123. The percentage distribution of pupils will be very similar to that in English. The school-type differences are fairly similar to the values for English, but the virtual absence of an overall gender difference in Mathematics has the effect of bringing the category averages more into line, so that

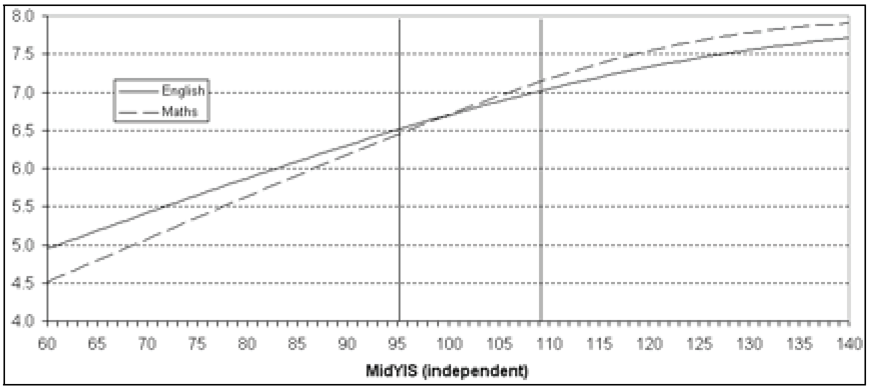
differential representation of pupil categories has less effect than in English on the equation of the regression line.

124. The diagram suggests that in our notional average mixed school the boys have done just about as well as the girls, but both genders have done a little less well than their counterparts in single-sex schools. The average school will have a negative residual in the subject from the independent standardisation, but residuals will not vary greatly across the ability range.

125. It is still likely that more mixed schools than single-sex schools will have negative average residuals, but on average the over-representation of boys and band C pupils in these schools will be slightly less of an issue than it was in English. (62% of mixed schools had negative average residuals in Mathematics: of these 91 schools 23 had average MidYIS scores of 100 or more.)

126. In real mixed schools, pupils of both genders and all three ability bands will in general have achieved average results which differ from the sector's band/school-type averages. Nevertheless, the independent standardisation value added of all mixed schools will be influenced by the above considerations to a greater or lesser extent. *Figures 23 and 25* can also indicate what the Directors of Studies in average boys' schools and girls' schools will find in their independent standardisation feedback. The diagram below plots the regression curves for the two subjects: it is interesting to note the expected correspondence with *Figures 23 and 25*. The vertical lines on the graph are drawn at the ability band cut-off scores.

Figure 26: Independent standardisation regression curves for English and Mathematics, Year 7



127. The description offered here is not, of course, an attempt to explain below-average performance. Our hypothetical Director of Studies has to accept the outcome for what it is, though he may now be in a position to understand a little more about how the measurement works. In summary, the generally favourable value added measured by the national standardisation tends to mask differences internal to the independent sector which are revealed by the independent standardisation, this being especially the case in the lower reaches of the independent sector ability range.

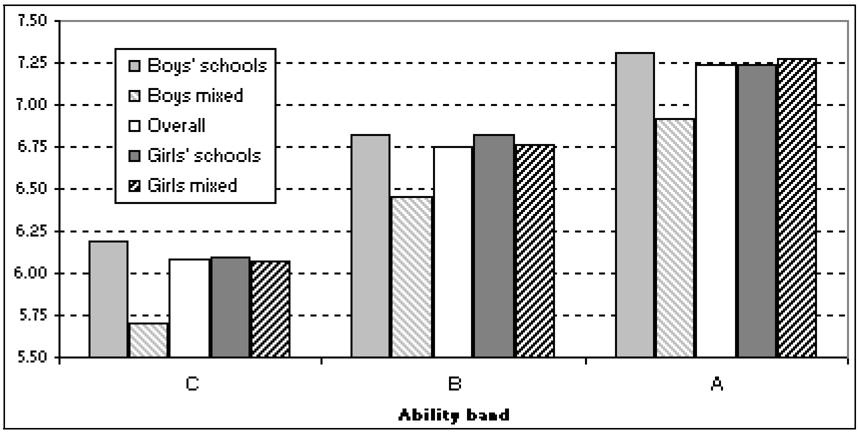
128. Every subject will have its own equivalent of *Figures 23 and 24*. In some cases, the values plotted will be influenced by variations between school types in the availability of subjects (in Science, perhaps). But in every case, the standard regression equation will be influenced to a greater or lesser extent by variations between the ability bands of the average performances of pupils in the four possible categories and of pupil representation by category.

Gender-Related Influences on the Year 9 Regression Curves

129. In general, gender and school-type difference will influence independent school Year 9 value-added measurement for similar reasons to those arising from paragraph 119. However, the effects will be unlikely to follow the same pattern as in the Year 7 project because, as already established, (a) overall gender differences are probably smaller in the Year 9 sample (and in boys' favour in some subjects), (b) the gender make-up of the Year 9 sample is very different, and (c) although a detailed investigation

has only been carried out for pupils new to MidYIS in Year 9, it appears probable that school-type differences are small for girls though still significant for boys.

Figure 27: GCSE English results (Y9 independent school sample) by school type and ability band

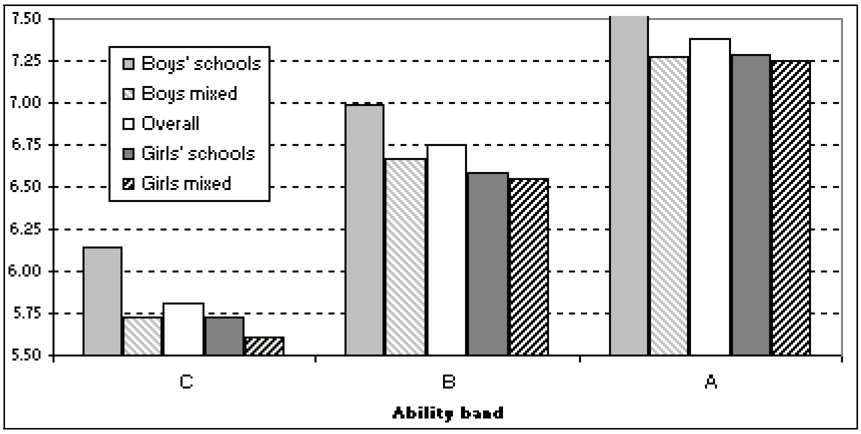


130. It appears that in this case any overall gender gap in favour of girls arises from the below-par average performance of boys in mixed schools. (For reasons explained in paragraph 91, no overall gender gap for the Year 9 sample has been included in this paper.) As expected, the school-type difference is negligible for girls, but significant for boys.

131. At school level, therefore, mixed schools are more likely than single-sex schools to have negative value added, but in the Year 9 sample this is principally determined by the average performance of boys in these schools and not, as in Year 7, because of significant school-type differences for both genders. "School level" is a less robust concept here than in the Year 7 sample, because it is known that some schools do not test complete year groups in Year 9.

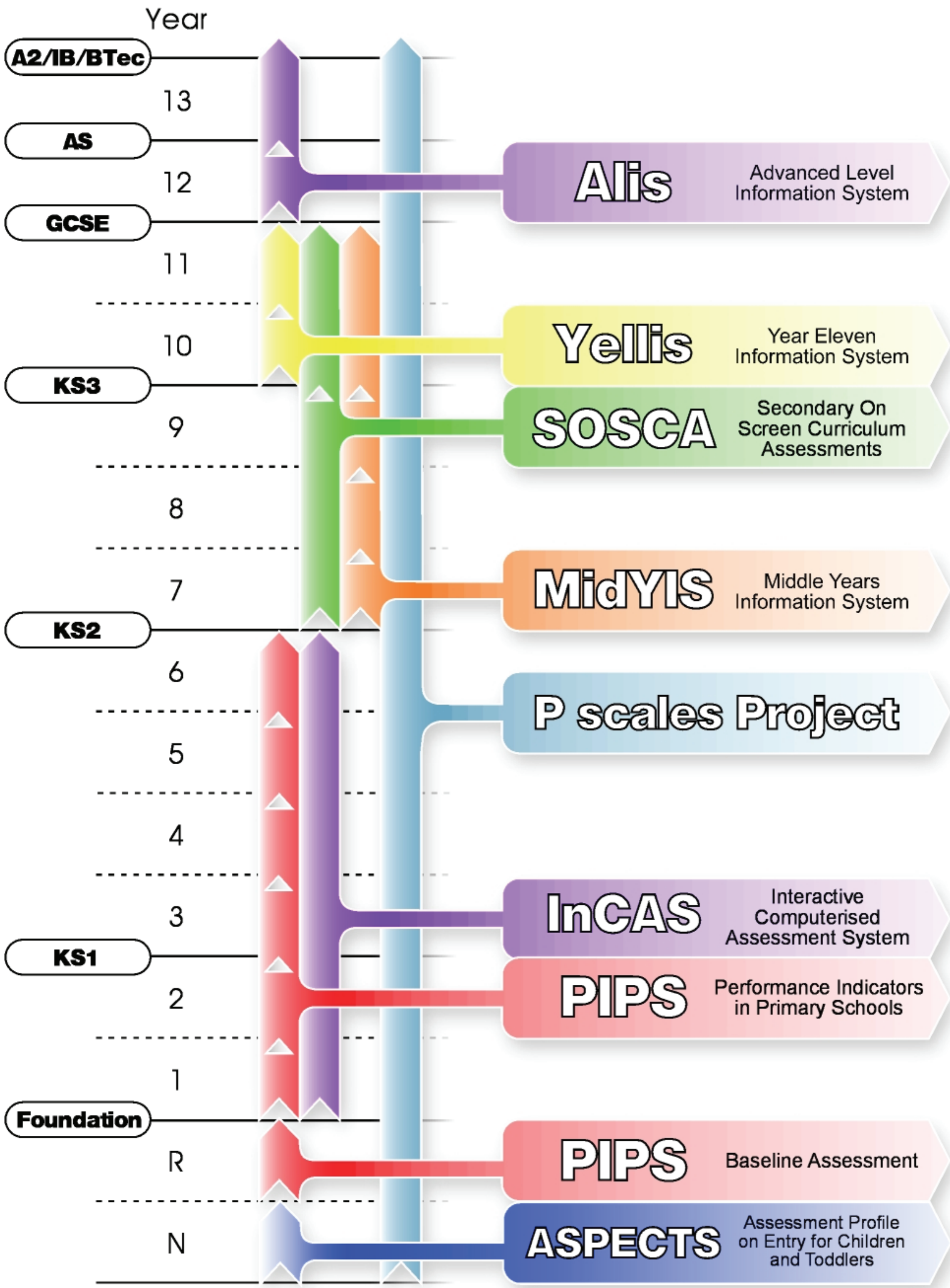
132. For information, the ability band cut-off scores for the Year 9 sample were 94.4 and 106.6, slightly lower than for the Year 7 sample.

Figure 28: GCSE Mathematics results (Y9 independent school sample) by school type and ability band



133. Despite the considerable school-type difference for boys, the overall gender gap is clearly in favour of boys in Mathematics: the boys in mixed schools achieved average grades similar to girls (the virtual absence of a school-type difference for girls being evident again here).

134. Clearly, the boys' schools are most likely to have positive residuals in this sample, but it is difficult to produce any definitive statement on the likely distribution of residuals between girls' and mixed schools.



www.cemcentre.org

CEM CENTRE

MidYIS