

# Analysis of the HMC and Girls' Schools Association (GSA) report on the 2015 Cambridge International Examinations (CIE) IGCSE® First Language English (entry number 0500) results in GSA and HMC schools



The nine analyses reported have been conducted on 4,903 candidates which studied at "...GSA & HMC schools that felt they had concerns about their CIE 0500 results" (pg5). This is an unrepresentative sample from a population of 17,619 students taking the qualification and this makes it difficult, if not impossible, to interpret the findings of the research. Data for all the candidates entered and for more than one year, and different, more sophisticated analyses would be necessary to draw conclusions about the fairness of the award for the 0500 entry. This will be illustrated below by discussing each analysis in turn. We have analysed available data from the considerably larger 0522 entry to test some of the assertions made in the report.

**Analysis 1** simply reports the structure of the specification and the sizes of each entry. The report focuses on the extended paper 2 which can be combined with either a writing exam (unit 3) or coursework (unit 4).

**Analysis 2** compares the grading outcomes for these two options. Students taking the 2/3 option do better than students taking the 2/4 option. The report claims that the difference in outcomes shows that "*There is clearly a problem*" (pg7). The difference in outcomes cannot be interpreted in this way. It is, of course, possible that the two options are differentially motivating or may be taught to different standards. Moreover, it is likely that the two options attract different abilities. While the report goes on to compare the paper 2 grades of the two options and shows that on this measure the two groups are similar, it seems probable that the coursework option attracts a lower ability entry than the exam option. To test this hypothesis we compared the Key Stage 2 (KS2) prior attainment of the two routes for 0522 entry candidates. The proportion of candidates that could be matched was relatively low (35.6% for unit 3 and 52.6% for unit 4). However, the analysis showed that the KS2 scores of the unit 3 candidates were better than those of the unit 4 candidates. For example, 10% of unit 3 candidates were in the most able octile of KS2 results compared to 7% of unit 4 candidates. Further, 15% of unit 3 candidates were in the least able octile compared to 18% of unit

4 candidates. It may be the case that this difference in ability profile does not generalise to the candidates of the 0500 entry but the analysis illustrates that it is unwise to interpret differences in raw grading outcomes without controlling for other factors such as the general academic ability of the cohorts.

It is also worth noting that the analyses reveal that the sample is unrepresentative in ways beyond their dissatisfaction with results. A CIE report (Benton, October 2015) gives a breakdown of the entries for the 2/3 and 2/4 routes. For 0522, 91% of candidates take the 2/4 coursework option, compared to 36% for 0500. In the sample analysed in the report, 75% of candidates take the 2/4 coursework option.

**Analysis 3** compares the unit grade distributions and comments that the higher grade distribution of unit 3 is more characteristic of the overall grades expected from candidates at the schools. This may be true but no supporting evidence for the statement is given. The report goes on to comment that the grades on the coursework unit 4 are “*much lower than would be expected*” (pg8). Again no supporting evidence is given for this statement and this unit is marked by teachers. The report states that this is due to the high boundaries. An A\* on the maximum raw mark is unsatisfactory. However, this will have been the case in previous years and it is not clear why it would have caused problems in this year only. The report states that there have not been grading concerns in the past.

**Analysis 4** uses the correlation between marks on units 2 and 3 to explore whether there is evidence of poor marking. A weaker than usual correlation between scores on two assessments might be caused by poor marking reliability. However, there are many other potential explanations, for example, the units assess different constructs.

We calculated the correlation between the units for the 0522 entry. Note, these are the same papers, marked in the same way as those of the 0500 entry. The correlation is higher than that reported (0.435 on page 8 of the report) but more importantly it is stable and slightly higher in 2015 than 2014 (see Table 1 below).

Table 1. Correlations between unit marks

Unit Combination	2015		2014	
	<i>r</i>	N	<i>r</i>	N
1 and 3	.614	8,215	.593	8,018
1 and 4	.437	47,489	.445	33,947
2 and 3	.590	9,246	.563	9,814
2 and 4	.499	125,617	.478	52,073

The report interprets the ‘low’ correlation as evidence that there were issues with the marking in 2015. These correlations do not support that hypothesis. If we were to interpret the correlation as telling us something about marking quality, then the increase suggests very slightly better marking in 2015 than 2014.

The report then goes on to correlate marks on unit 2 and unit 4 and concludes that “*no reliable correlation can be derived*” (pg9). We calculated the correlation for the 0522 entry candidates and it is moderate (.499) and stable compared to 2014.

The final part of this analysis compares the grades achieved on units 2 and 3, and on units 2 and 4. The report notes that over 200 students were ungraded in paper 2 and claims that “*many of these students would never have achieved less than grade A in any examination*”. Paper 2 is an extended paper (higher tier) and research shows that small numbers of students entered for the higher tier papers of GCSEs are often ungraded (see Wheadon and Béguin, 2010). It may be unusual for candidates from these schools to be ungraded in this unit but no evidence (e.g. comparisons over time) is given to support the assertion made.

**Analysis 5** compares the achieved grades with those forecast. It is usual for forecast grades to be higher than achieved grades. This is the case here. It is impossible to draw any conclusions from this analysis without the comparison for a number of years. Has the gap this year widened?

**Analysis 6** compares candidates’ grades with those in English literature, history and all other subjects. This shows that candidates’ grades in English were lower than in other GCSEs. However, this sample of schools are those with concerns about their *English* results and not their results in other subjects. It is therefore impossible to know whether the results of this analysis are worrying or not.

Indeed, using data from the 0522 entry we examined the results of candidates who took both CIE's English and English literature (N = 8,215) and CIE's English and history (N = 877). Taking English literature first, 31.2% of candidates were awarded the same grade in both subjects, 37.1% did better in English and 31.7% did better in English literature. For history, 34.9% of candidates were awarded the same grade in both subjects, 35.7% did better in English and 29.4% did better in history. These findings do not suggest a systematically severe award in English.

**Analysis 7** shows that the grade distribution for English nationally is lower than those in English literature, history and all other subjects. But the report claims that the gap at the top grades is less than that found in Analysis 6. However, this is not unexpected given the flawed basis of the latter analysis.

**Analysis 8** repeats Analysis 6 for a sample of 25 selective grammar schools. It is not clear how these 25 schools were selected. Lower outcomes in English compared to English literature, history and all other subjects remain at A and A\*. Again the report claims that the gap at the top grades is less than that found in Analysis 6. Again, this is not unexpected given the flawed basis of the latter analysis being based on a sample of concerned schools.

**Analysis 9** repeats Analysis 6 but disaggregates the grades of those taking unit 3 from those taking unit 4. It finds that the gap in grade outcomes is greater for those taking unit 4. This analysis is subject to the same issues as outlined for Analyses 6, 7 and 8. Further, as stated earlier, it is possible that the two options are differentially motivating or may be taught to different standards. It is also possible that the two entries attracted candidates of different abilities. Indeed, we compared the KS2 prior attainment for those taking the two routes within both the 0500 and 0522 entries using data for the 0522 cohort. The analysis showed that the KS2 scores of the candidates taking the unit 2/3 route were higher than those taking the unit 2/4 route. For example, 15% of unit 2/3 route candidates were in the most able octile of KS2 results compared to 8% of unit 2/4 candidates. Further, 7% of unit 2/3 candidates were in the least able octile compared to 11% of unit 2/4 candidates. Data for the 0500 entry were not available but it seems likely that there will be systematic differences in the abilities of the two groups of candidates.

In summary, none of the analyses presented provide compelling evidence that the 0500 grade boundaries were inappropriately set or otherwise.