

ANOVA on A*s at A-level and Tripos performance

Introduction

This paper provides an assessment of the relationship between the attainment of A* grades in GCE A-levels and Tripos performance. It builds on previous work¹, which demonstrated a clear positive relationship between the total number of A* grades attained at A-level, and performance in first year Tripos examinations.

This paper replicates the 2011 analysis, using additional data years so as to draw more robust statistical conclusions, for Arts, Sciences and Social Sciences grouped Triposes in the University. This analysis is then extended to individual Triposes. The relationship between A*s attained at A-level and Tripos performance by subject group is further examined by school type.

The paper concludes that the relationship between A* attainment at A-level and first year Tripos performance remains consistently positive for all subjects. In Arts and Social Sciences, there is a linear relationship between the number of A* grades achieved and the first year Tripos percentage obtained. In the Sciences, although there is a significant difference in the Tripos performance of all attainment groups, those attaining three and four A* grades appear to benefit most.

Methodology

The basis of the analysis undertaken in this paper is the Analysis of Variance, which is a statistical technique used to examine significant differences between groups in the mean of a dependent variable. The dependent variable is first year Tripos percentage, which is taken as a proxy for overall Tripos examination performance.² We will assess how far group means vary for:

- Students with one, two, three and four or more A* grades at A-level, by the subject group of their Tripos
- Students with the above level of grouped A* attainment and school type, by the subject group of their Tripos
- Students with the above level of grouped A* attainment, by their Tripos

ANOVA by subject group

Previous work has suggested that, at a University-wide level, there appeared to be a statistically significant positive linear relationship between the number of A* grades attained, and the performance (measured by Tripos percentage) in first year examinations. This relationship was sustained at the subject group level, although with less of a clear linear relationship. Most notably, Sciences subjects showed little difference in Tripos percentage between those students who achieved one and two A* grades at A-level, but a significantly higher performance for those with three and four A* grades.

This analysis focuses on subject groups only, rather than the University-wide population, because sufficient data are now available to ensure that the statistics calculated in the ANOVA model are

¹ P. Chetwynd, 2011, *Tripos performance and A*s at A level*, ARWP October 2011.

² With suitable caution about the problems involved in translating percentage marks into classmarks, or an overall measure of performance, across all the Triposes at Cambridge.

robust. We begin by replicating the results of Chetwynd's (2011) analysis, with additional data years, from the 2010 to 2013 cycles inclusive, examining the relationship between grouped A* attainment and Tripos percentage, by subject group. We would expect his tentative results to be confirmed by further analysis: that there will be a positive relationship between the number of A* grades attained and Tripos performance.

Attainment of A*s uses four groups for comparison: one A*, two A*s, three A*s and four or more A*s. Although there are students who gain higher numbers of A* grades at A-level than four, they are relatively few in number, and so they have been included in the four A* grouping. Those very few students who do not attain any A*s at A-level are excluded from this analysis.

Tables 1 below shows the results of the ANOVA undertaken on the dependent variable (Tripos percentage) and the grouped A* variable, for the subject groups Arts y1,³ Sciences y1 and Social Sciences y1 respectively.

Table 1. Descriptive statistics and ANOVA on Tripos percentage by A group.*

ARTS year one				SCIENCES year one				SOCSCI year one			
A* Grp	Mean Tripos %	StDev	N	A* Grp	Mean Tripos %	St Dev Tripos %	N	A* Grp	Mean Tripos %	St Dev Tripos %	N
1	0.618	0.062	458	1	0.555	0.099	313	1	0.603	0.051	276
2	0.640	0.057	597	2	0.576	0.084	808	2	0.618	0.052	416
3	0.656	0.056	390	3	0.622	0.086	1605	3	0.632	0.056	387
4	0.676	0.052	119	4	0.660	0.091	2286	4	0.655	0.042	269
overall	0.640	0.060	1564	overall	0.628	0.096	5012	overall	0.627	0.054	1348

Subject Group	F	p	Welch's F
Arts	47.0	0.00	n/a
Sciences	n/a	0.00	259.8
Social Sciences	n/a	0.00	64.9

We see significant F/Welch's F⁴ statistics for all subject groups, and therefore we can conclude that there is a statistically significant difference in the mean Tripos percentage. For all three subject groups, it seems that there is an association between our dependent variable, Tripos percentage, and the attainment group of students at A-level.

In order to work out what kind of relationship there is between the different levels of A* attainment and Tripos percentage, we need more detail than the F statistic alone provides. Consequently, post-

³ Inevitably this excludes ASNAC, English, History and MML from the Arts group results.

⁴ Welch's F is used in Sciences and Social Sciences where Levene's test is significant and a more robust F statistic is required.

hoc tests, which look at the differences observed between each level of the variable A*, can be used, where the original model yields a significant F statistic.⁵ In this analysis, two kinds of post-hoc tests are used. The first, Hochberg's GT2, is used for the Arts y1 sample, because it is especially well-calibrated to deal with the differences in the sample size of each group (students with two A*s outnumber students with only one A*), when Levene's test is non-significant. In this case, the differences between each level, and every other level are included in the far left column, and significant differences in the mean percentage mark between that and the comparator groups are flagged at the 0.05 significance level. The results of Hochberg's GT2 are presented in table 2 below, and all of the post-hoc comparisons are significant, meaning that there is a statistically significant difference in Tripos percentage between each A* group.

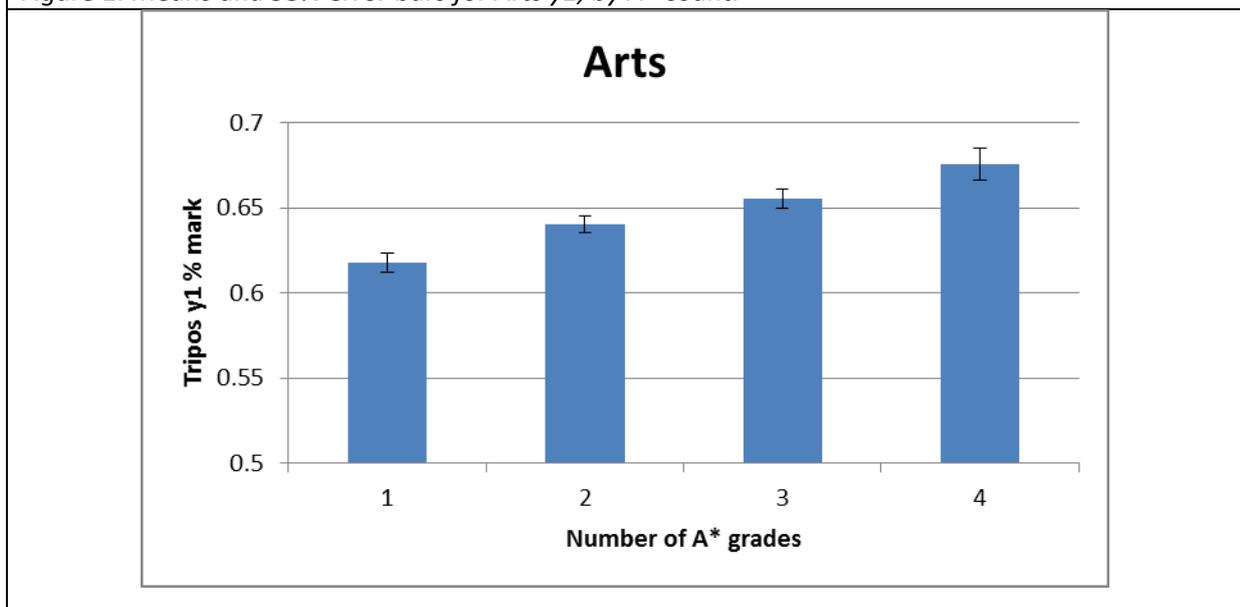
⁵ These tests essentially conduct t-tests, looking for significant differences between the mean Tripos percentage for two groups (for example those with one A* and those with two A*s), but, crucially, they must control for Type 1 errors, and so they are harder to satisfy. It is theoretically possible for the overall ANOVA to be significant, but for differences between the groups in post-hoc tests to be insignificant.

Table 2. Hochberg's GT2 post-hoc test: Arts y1

Profile	Comparison	Mean difference	Std. Err.	p	Significant
A*	A*A*	-0.022	0.00	0.00	y
	A*A*A*	-0.038	0.00	0.00	y
	A*A*A*A*	-0.058	0.01	0.00	y
A*A*	A*	0.022	0.00	0.00	y
	A*A*A*	-0.015	0.00	0.00	y
	A*A*A*A*	-0.035	0.01	0.00	y
A*A*A*	A*	0.038	0.00	0.00	y
	A*A*	0.015	0.00	0.00	y
	A*A*A*A*	-0.020	0.01	0.01	y
A*A*A*A*	A*	0.058	0.01	0.00	y
	A*A*	0.035	0.01	0.00	y
	A*A*A*	0.020	0.01	0.01	y

The results in table 2 are more easily visualised in figure 1 below, which shows the mean Tripos percentages, and 95% confidence intervals for the mean, for each group of A* attainment. From the graph below, we can see that the confidence intervals for the means do not overlap. There is a positive, linear relationship between the total number of A*s achieved at A-level, and Tripos attainment in the first year of Arts Triposes, with each additional A* making a statistically significant and positive difference to attainment.

Figure 1. Means and 95% error bars for Arts y1, by A count.*



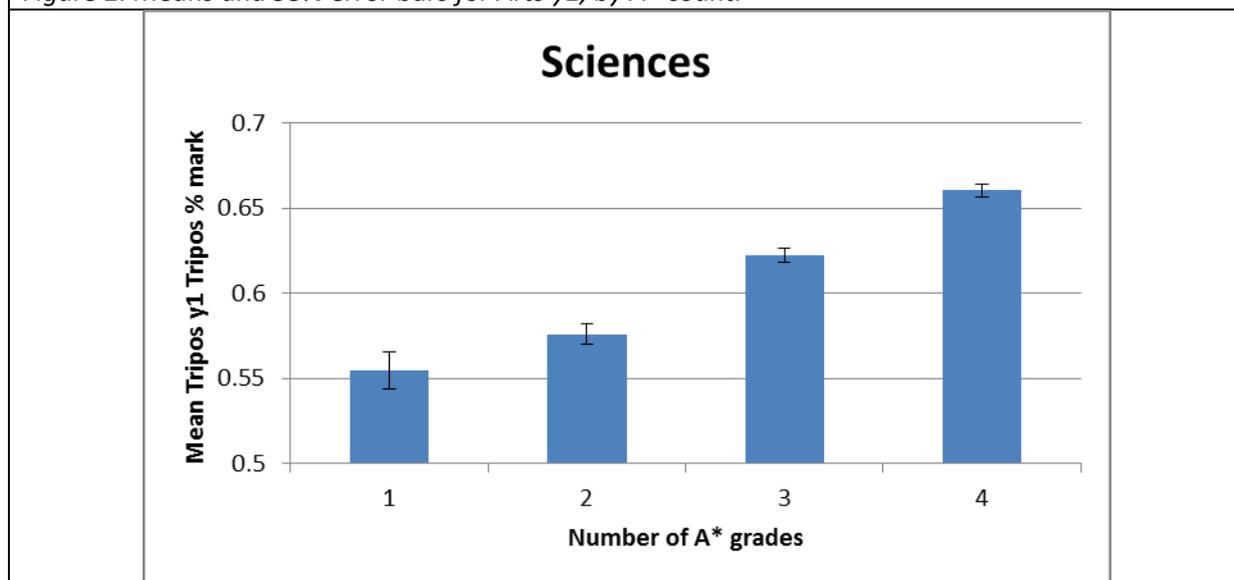
When we turn to the results for Sciences and Social Sciences year 1, although we are also interested in the statistical significance of differences between A* attainment groups, we change the post-hoc tests we employ to test for statistical significance. In both subject groups, the Levene's test results, presented in tables 1b and 1c above, were significant. Consequently, we cannot assume that there is homogeneity of variance in our subject group samples: we therefore use Games-Howell post-hoc procedures.

Table 3 below presents the results for Games-Howell post-hoc tests on Sciences y1. Figure 2 follows it, and presents the means and 95% confidence intervals for the mean in graphical form.

Table 3. Games-Howell post-hoc tests: Sciences y1.

Profile	Comparison	Mean difference	Std. Err.	p	Significant
A*	A*A*	-0.021	0.01	0.00	y
	A*A*A*	-0.067	0.01	0.00	y
	A*A*A*A*	-0.106	0.01	0.00	y
A*A*	A*	0.021	0.01	0.00	y
	A*A*A*	-0.046	0.00	0.00	y
	A*A*A*A*	-0.084	0.00	0.00	y
A*A*A*	A*	0.067	0.01	0.00	y
	A*A*	0.046	0.00	0.00	y
	A*A*A*A*	-0.038	0.00	0.00	y
A*A*A*A*	A*	0.106	0.01	0.00	y
	A*A*	0.084	0.00	0.00	y
	A*A*A*	0.038	0.00	0.00	y

Figure 2. Means and 95% error bars for Arts y1, by A count.*



All the comparisons undertaken in the Games-Howell post-hoc tests yielded significant results, which means that we can conclude that there is a statistically significant difference in means of the groups.

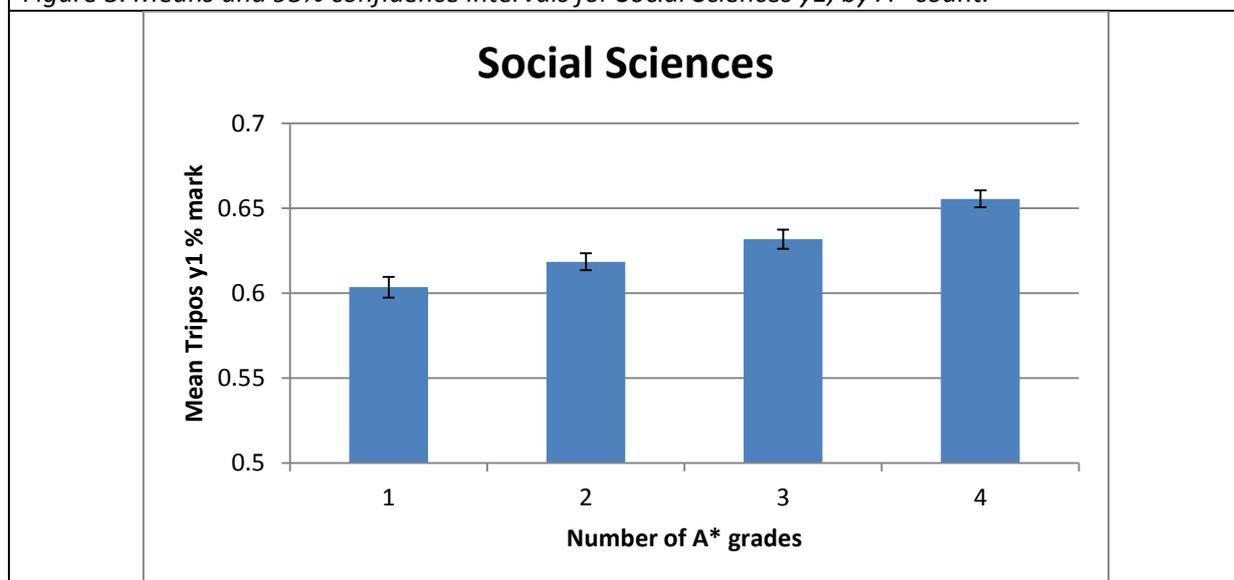
Figure 2 in particular emphasises the magnitude of the difference between the means for those studying Sciences y1 with two A*s and fewer, and three A*s and more.

The Games-Howell results for Social Sciences are presented in Table 4 below, and Figure 3 graphs the differences in group means and 95% confidence intervals. As with Arts and Sciences, we can see that there is a statistically significant difference between all group means, and that the greatest magnitude of difference occurs in the four or more A* category.

Table 4. Games-Howell post-hoc tests for Social Sciences y1.

Profile	Comparison	Mean difference	Std. Err.	p	Significant
A*	A*A*	-0.015	0.00	0.00	y
	A*A*A*	-0.028	0.00	0.00	y
	A*A*A*A*	-0.052	0.00	0.00	y
A*A*	A*	0.015	0.00	0.00	y
	A*A*A*	-0.013	0.00	0.00	y
	A*A*A*A*	-0.037	0.00	0.00	y
A*A*A*	A*	0.028	0.00	0.00	y
	A*A*	0.013	0.00	0.00	y
	A*A*A*A*	-0.024	0.00	0.00	y
A*A*A*A*	A*	0.052	0.00	0.00	y
	A*A*	0.037	0.00	0.00	y
	A*A*A*	0.024	0.00	0.00	y

Figure 3. Means and 95% confidence intervals for Social Sciences y1, by A count.*



ANOVA by Tripos subject

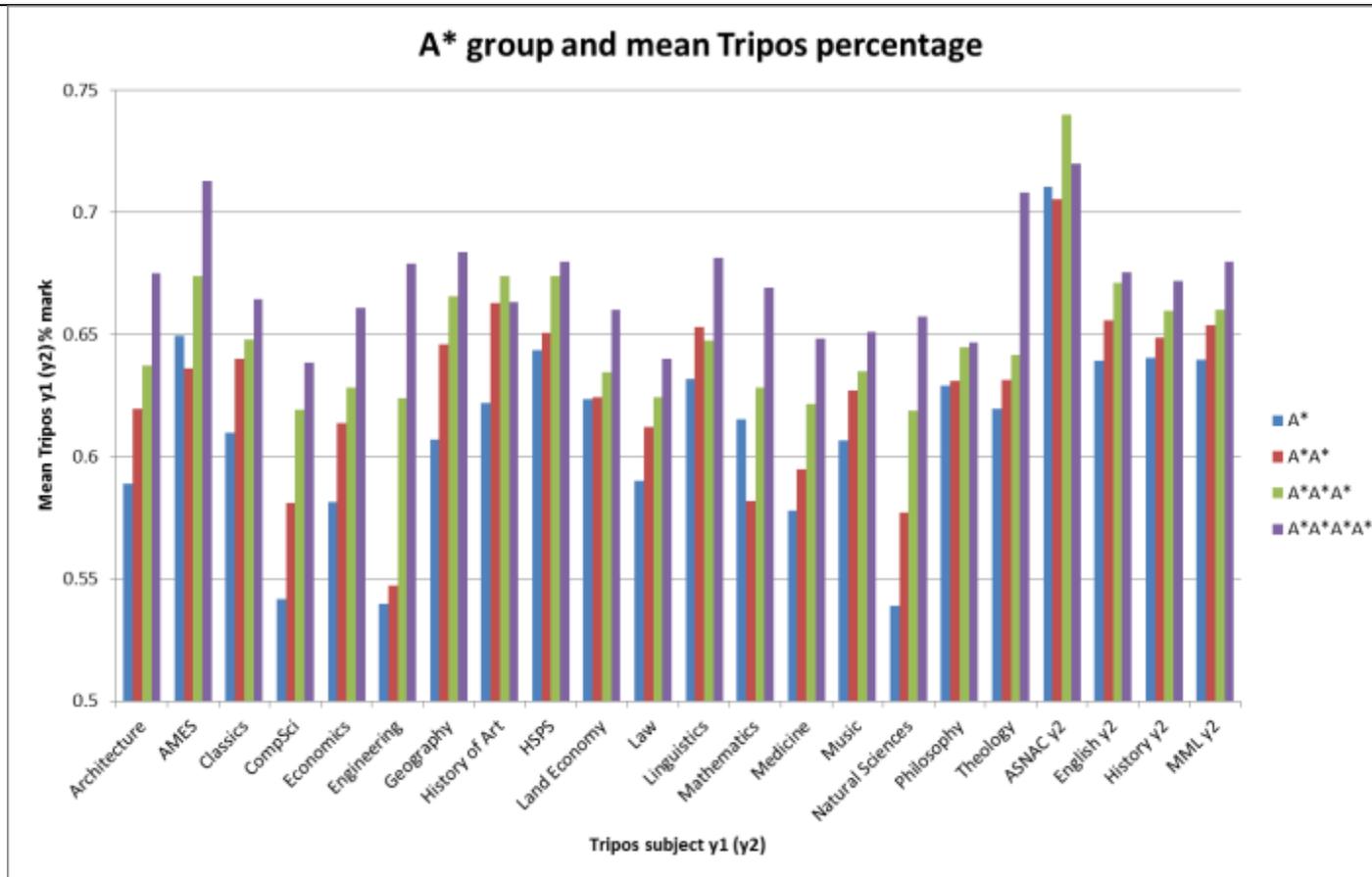
The patterns obtained at the subject group level are both convincing and consistent with previous results. While it is the next logical step to conduct similar ANOVA procedures on individual Tripos

subjects, in some of the smaller Triposes, a less statistically consistent and significant pattern can emerge.

On the whole, however, the individual Triposes conform to the general pattern of the subject groups, in that they show a significant F statistic in ANOVA, indicating a difference in the mean Tripos percentage of A* attainment groups. Where Tripos subjects can differ from their parent subject group is in the variation between levels of the A* groupings, identified through post-hoc tests. In some subjects there is considerable difference between those who attain one and two A* grades, while in others, significant differences only emerge at the point at which three A* grades are attained. A graph of the mean Tripos percentage for the A* groupings can be seen in figure 4 below.

PREDICTION OF TRIPOS PERFORMANCE

Figure 4. Mean Tripos percentage by A* group for individual subjects.



The graph shows that there are some subjects where there is a clear differential between the lower two groups of A* attainment and the higher two groups in terms of Tripos performance. Engineering, Natural Sciences and Computer Sciences are three examples of where the mean mark is under 60% for those with two or fewer A* grades. For those subjects where we use second year results, there is substantially reduced variability in Tripos outcome, but, in the case of the larger subjects of English, History and MML, there is a discernible relationship between A* attainment and Tripos percentage.

By contrast, subjects such as Philosophy and Land Economy show very limited between-group variability in Tripos percentage, at all attainment levels. It is likely that the relatively few observations in each group of attainment for these Triposes has substantially reduced our ability to detect any meaningful relationship between A* attainment and Tripos. With more data years, a different pattern might emerge.

Table 5 below shows a summary of the ANOVA models and appropriate post-hoc tests for each Tripos subject. In some smaller subjects no significant F statistic is returned in the ANOVA model. This is because there is either no clear direction in the relationship between A* attainment and Tripos percentage, or because observed between-group differences are too small to obtain a significant value of F. Such subjects are marked as 'n' in the ANOVA F sig column.

Table 5. Summary of ANOVA models for all subjects

Subject	ANOVA F significant?	A* and A** sig diff	Effect size est. A* and A**	Probability	A** and A**A** sig diff	Effect size est. A** and A**A**	Probability
Architecture	y	n	0.4	0.61	n	0.2	0.56
AMES	y	n	-0.2	-0.56	n	0.6	0.66
Classics	y	y	0.6	0.66	n	0.2	0.56
CompSci	y	n	0.5	0.64	n	0.5	0.64
Economics	y	y	0.6	0.66	n	0.2	0.56
Engineering	y	n	0.1	0.53	y	0.8	0.69
Geography	y	y	0.5	0.64	n	0.3	0.58
History of Art	y	y	1.5	0.85	n	0.4	0.61
HSPS	y	n	0.2	0.56	y	0.7	0.71
Land Economy	n	n	0	0.5	n	0.2	0.56
Law	y	y	0.4	0.61	n	0.2	0.56
Linguistics	n	n	0.4	0.61	n	-0.1	-0.53
Mathematics	y	n	-0.3	-0.58	y	0.4	0.61
Medicine	y	n	0.3	0.58	y	0.5	0.64
Music	y	n	0.4	0.61	n	0.2	0.56
Natural Sciences	y	y	0.5	0.64	y	0.5	0.64
Philosophy	n	n	0	0.5	n	0.3	0.58
Theology	y	n	0.2	0.56	n	0.2	0.56
ASNAC y2	n	n	-0.1	-0.53	n	0.4	0.61
English y2	y	y	0.4	0.61	y	0.3	0.58
History y2	y	n	0.3	0.58	y	0.4	0.61
MML y2	y	n	0.4	0.61	n	0.1	0.53

NOTE that Probability columns show the likelihood that, when a person is selected at random from the control (lower A) and experimental (higher A*) group, that higher Tripos percentage will be obtained by the person from the experimental (higher A*) group.

Models where the post-hoc tests return a significant difference between specified levels of A* attainment are shown in bold, and marked with a 'y' in the column appropriate for each level. For all subjects, a simple estimate of the effect size is given, alongside an associated crude probability that a random person from the higher A* group would obtain a Tripos percentage higher than a random person from the lower A* group.

When we combine this table with the graphs of mean values in Figure 4, we might speculate that:

- Engineering, Medicine and Natural Sciences show a significant difference in Tripos percentage between two A* and three A* grades, a probability greater than 60% that a person with three A*s would outperform one with two A*s, and a mean percentage above 60 for those with three or more A* grades.
- Mathematics returns an anomalous result for one A*, but this is based on only 21 observations. It is therefore unsurprising that we see no significant difference only between two A* and three A*s, along with a probability of higher marks over 60%, and a mean mark of over 60% for those with three or more A*s.
- Computer Science does not return a significant difference at any group level, but the overall F statistic is significant, and there is a clear trend for Tripos percentage to increase positively with the total number of A* grades attained.
- Classics, Geography, History of Art and Law show a significant difference in Tripos performance for the one A* and two A* groups, a probability greater than 60% that the higher attaining student will perform better in Tripos, and have a higher mean Tripos percentage for those attaining two A*s or better.

ANOVA by subject group and school type

The last part of this paper focuses on the incorporation of school type into the ANOVA models, and post-hoc tests, outlined above. We will test the effect on subject group Tripos performance of A* attainment and school type, and any interaction effect between school type and attainment. School type is operationalised as a binary measure, where schools can be classified either as independent or maintained. Any schools which cannot be classified in this way are excluded from the analysis.

As a result of adding the independent/maintained school variable, our ANOVA becomes more complex, and the number of groups investigated increases: for example, the original A*AA only group is changed, to A*AA independent and A*AA maintained. The ANOVA, and its associated post-hoc tests, continues to look for significant differences in the mean Tripos percentage obtained across this matrix of different groups.

Table 5 below shows the descriptive statistics, Levene's test results and main ANOVA table for each of the subject groups. Significant F statistics are those with p-values less than 0.05 in the appropriate column.

Table 5. ANOVA by subject group, school type, and A achievement.*

ARTS year one	Mean Tripos %		St Dev Tripos %		N	
A* Group	independent	maintained	independent	maintained	independent	maintained
1	0.619	0.617	0.060	0.060	196	250
2	0.642	0.639	0.060	0.060	302	279
3	0.655	0.656	0.050	0.060	196	187
4	0.673	0.681	0.050	0.060	62	54

SCIENCES year one	Mean Tripos %		St Dev Tripos %		N	
A* Group	independent	maintained	independent	maintained	independent	maintained
1	0.553	0.549	0.108	0.085	79	209
2	0.575	0.578	0.084	0.082	234	525
3	0.616	0.626	0.090	0.083	575	955
4	0.662	0.660	0.094	0.087	949	1217

SOCSCI year one	Mean Tripos %		St Dev Tripos %		N	
A* Group	independent	maintained	independent	maintained	independent	maintained
1	0.611	0.598	0.050	0.050	69	192
2	0.615	0.619	0.047	0.055	140	260
3	0.632	0.631	0.066	0.048	171	186
4	0.653	0.659	0.042	0.043	114	124

Subject Group		F	p
Arts	A* group	46.3	0.00
	school_type	0.1	0.74
	school + A*	0.3	0.82
Sciences	A* group	238.7	0.00
	school_type	0.3	0.58
	school + A*	1.4	0.24
Social Sciences	A* group	42.7	0.00
	school_type	0.1	0.77
	school + A*	1.6	0.20

In all three of the subject groups, there is no statistically significant difference by school type of the mean Tripos percentage obtained in first year examinations. There is also no significant interaction effect between school type and A* attainment which impacts on first year Tripos percentage.

This result echoes Chetwynd's earlier findings, which also showed no significant difference between the means of Tripos percentage for different school types. It amplifies too the long-standing claim that school type has a negligible impact on educational outcomes at this University, once account has been taken of prior academic attainment.

Conclusion

Overall, then, a conclusive and powerful picture of the relationship between A* attainment at A-level and first year Tripos performance emerges from the new ANOVA using data from 2010 to 2013 inclusive. There is a statistically significant difference between A*AA candidates and those with two or more A*s in all subject groups. In the Sciences, the pattern towards higher Tripos attainment for those with more A*s is even more pronounced, but is also present in the Arts and Social Sciences.

Individual Triposes show a more complex and diverse set of relationships, but the broad trend toward higher attainment for those entering with higher grades is not challenged. Many Sciences subjects already show a performance bonus for those with three A* grades, while in some Arts subjects, a similar pattern is seen for those with two A* grades.

Attendance at an independent or maintained sector school has no statistically significant effect on the ANOVA model for any subject group, nor does it have a significant interaction effect with A* attainment level.

The indicative results presented in the 2011 study are therefore substantiated by this updated analysis, with the exception that where his data enabled him only to suggest trends, we now have sufficient information available to demonstrate the statistical significance of prior attainment for performance in the first year of Tripos for all subject groups in the University.

Catherine Sumnall

Research Officer, CAO

July 2015