Changes in student entry competencies 2001 – 2017

Hodds, M., Shao, J. & Lawson, D.

Original citation & hyperlink:
Hodds, M, Shao, J & Lawson, D 2020, 'Changes in student entry competencies 2001 - 2017', International Journal of Mathematical Education in Science and Technology, vol. (In-Press), pp. (In-Press).
https://dx.doi.org/10.1080/0020739X.2020.1849836

DOI 10.1080/0020739X.2020.1849836
ISSN 0020-739X
ESSN 1464-5211

Publisher: Taylor and Francis

Changes in student entry competencies 2001 - 2017

Mark Hodds*, Jia Shao, and Duncan Lawson

sigma, Coventry University, UK

Email: ab7634@coventry.ac.uk

Word count: 6210

Changes in student entry competencies 2001 - 2017

Diagnostic testing on entry to university provides an opportunity for students to ascertain individual strengths and weaknesses and allows course teams to gain insight into the competencies of the whole cohort. Coventry University has been using the same diagnostic test since 1991 and a previous analysis showed students coming to Coventry in 2001 were performing considerably worse than their counterparts in 1991 (Lawson, 2003). Given the changes to A level Mathematics that have taken place since 2001 and the recently implemented reforms to A level Mathematics in the UK, it is timely to review performance in the diagnostic test and to create a benchmark before students enter higher education with the new (linear) A levels. This paper therefore reports on the changing entry competencies of students at Coventry University from 2001-2017 and shows that there has been a general improvement in performance across all entry grades. However, students are still not performing at the same level as they were in 1991 and the gap between students with higher entry grades (A level A* - C) and lower entry grades (A level D - E) is getting wider.

Keywords: Diagnostic Testing, Transition to University, Mathematics Education

Introduction

In the 1990s, many UK higher education institutions began administering a form of mathematics diagnostic test during the first few weeks of the academic year (Edwards, 1996; Savage, 2003; MathsTEAM, 2003). Such testing provides students with an opportunity to determine their mathematical strengths and weaknesses as they begin their courses. Coventry University introduced diagnostic testing in 1991. Initially these tests at Coventry were given to students on those courses that were considered “at risk”, i.e. courses where historically the students had seriously struggled with the mathematical content such as sub-degree Higher National Diploma engineering courses and some degree courses with low mathematical entry requirements (Lawson, 2003).

Now, the diagnostic test at Coventry University is taken by students on any course that has a reasonable level of mathematical content. In the 2017-18 academic year, students from 36 different courses took a form of the diagnostic test, with over 1600 students taking a diagnostic test. There are three different forms of the test (Higher, Intermediate and Foundation) and students take the test that is most suited to both their entry qualification in mathematics, and the mathematical content contained within their course[1]. Although there are now three forms of the diagnostic test, the questions on the Higher diagnostic test have remained the same since it was introduced in 1991. This allows us to make direct comparisons between each year's cohorts. Each test contains 50 multiple choice questions (with one correct answer, three distractors, and one option for "Don't Know") covering various topics such as algebra, arithmetic, calculus, and trigonometry. Students currently take the test in their first week and have one hour to complete the test. Calculators are not allowed and students are told not to guess but to use the "Don't Know" option if they are unsure of the answer. The students then receive their results in a letter that explains which areas they need to improve on and links these areas to a set of worksheets. The worksheets can be obtained online, or from the

---

[1] The Higher diagnostic test is taken by students on courses which have an A level Mathematics (or equivalent) entry requirement. The Intermediate diagnostic test is taken by students on courses with a GCSE Mathematics (or equivalent) entry requirement and which contain a moderate amount of mathematics. The Foundation diagnostic test is taken by students on courses with a GCSE mathematics (or equivalent) entry requirement and which contain a low level of mathematics. There are two courses, Nursing and Oil and Gas Management, which have their own specialised version of the Foundation and Intermediate tests respectively.

Mathematics and Statistics Support Centre based in the library, to work on and students can seek support with the topics if necessary.

The main intention of the diagnostic test is to provide individual students with insight into their own mathematical capabilities. However, since the test has remained the same for the past 28 years, the data from the testing provide us with the ability to make some useful comparisons between the different sets of new students. Indeed, there have been several articles regarding the analysis of such data, not just from data collected at Coventry University but also similar data from other institutions (e.g. Hunt & Lawson, 1996; Todd, 2001; Lawson, 2003; Treacy & Faulkner, 2015). This paper is an update to the paper by Lawson (2003) which described the changes in student entry competencies between 1991 and 2001 at Coventry University.

In his 2003 paper, Lawson presented some noteworthy results regarding students' entry competencies in mathematics. He found that the performance of students with grades D or E in A level Mathematics had declined in the decade starting in 1991. Furthermore, students with a grade B in A level Mathematics in 2001 were performing at the same level as those with a grade N[2] in 1991. Lawson concluded that although the data did not provide evidence that A level standards were falling, it did suggest that A level Mathematics courses were not providing students with the same competencies that they were in 1991.

Since the publication of Lawson's paper, A level Mathematics has undergone several changes. Firstly, "Curriculum 2000" was introduced, with the first cohort of students

---

[2] In the early 1990s, A levels were graded as A-E pass grades (A being the highest), then came grade N (narrow fail) and grade U (unclassified). The N grade was abandoned in the late 1990s.

reaching university in 2002. This imposed a uniform modular system on A levels in all subjects and had unfortunate consequences for mathematics. In its first year, the failure rate in AS level Mathematics was much higher than in any other subject and there was a drop of almost 20% in the numbers taking mathematics A level between 2001 and 2002 (Smith, 2004). A subsequent Government inquiry into mathematics education post-16 stated that the introduction of Curriculum 2000 was ‘a disaster for mathematics’ (Smith, 2004, p.8). Emergency measures were put in place to remedy the situation and a new approach to A level Mathematics was developed. With this revised curriculum, the popularity of A level Mathematics steadily increased and it is currently the A level subject taken by the largest number of students. In addition, A level Further Mathematics has also increased in popularity (JCQ, 2018). However, mathematics is once again at a crossroads as further significant changes to the structure of all A levels in England are currently being implemented. There is great concern that fewer students will take Mathematics (at A level), and Further Mathematics (at both AS and A level), reversing the recent trend of increasing numbers taking these subjects. Furthermore, Core Mathematics, a new qualification between A level and GCSE, has been introduced and there have been reforms to GCSE Mathematics which may also affect the number of students taking A level Mathematics and Further Mathematics. The introduction of these changes to the pre-University mathematics education framework makes now a natural time to create a baseline by reflecting on how the changes implemented since 2001 have impacted students' entry competencies at Coventry University up to 2017.

This paper, therefore, looks at the data from Coventry University’s diagnostic tests from 2001 to 2017. For the purposes of this paper, and to compare the data to that of Lawson (2003), the data reported are taken only from those students who took the Higher version of the diagnostic test and had an A level in Mathematics (in total 3152 students) to allow for fair comparisons.

Overall cohort performance 2001 - 2017

Figure 1 shows the number of students with A level mathematics taking the Higher diagnostic test each year from 2001 to 2017 and the distribution of their Mathematics A level grades. The number of students with A level Mathematics dropped to a low of 43 in 2005 and grew to a high of 439 in 2016. Overall, there is a much larger proportion of students obtaining higher A level grades now compared with in the early 2000s. In 2001, approximately 30% of students taking the diagnostic test had grades A-C whereas in 2017, this proportion was 80%. Table A1 in Appendix 1 shows the number of students with each grade in each of the years from 2001 to 2017.

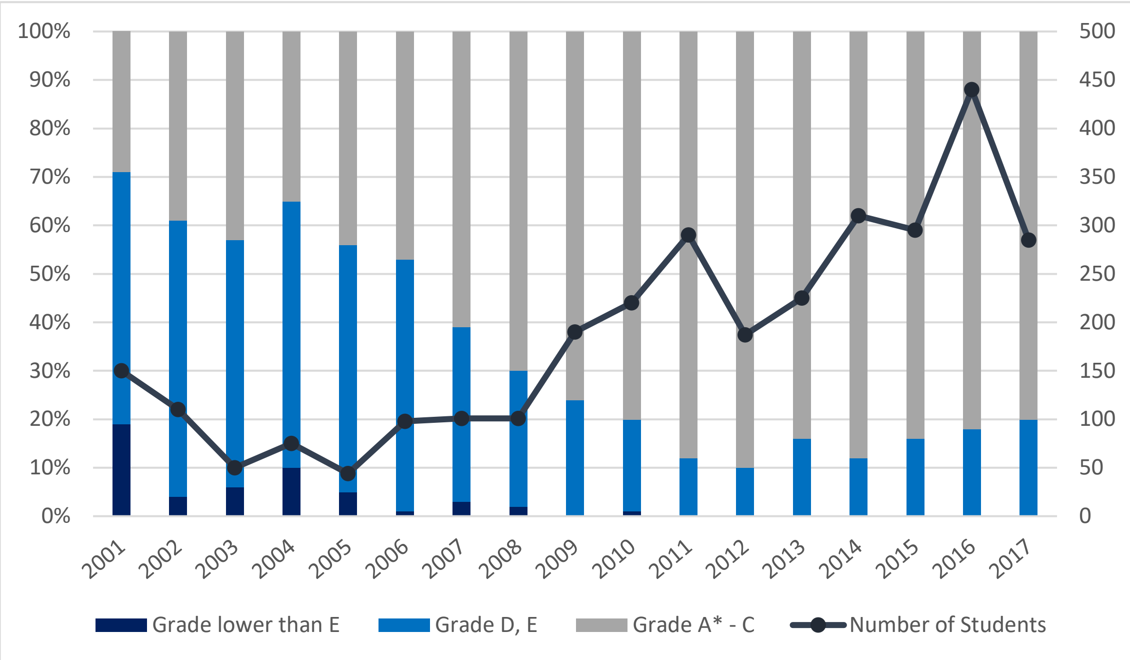

Figure 1: The number of students and proportion of students with different grades taking the higher diagnostic test, 2001 to 2017.

Figure 2 shows the average mark (out of 50) in the diagnostic test for all the students recorded in Figure 1. It shows that the average mark increased from 29.3 in 2001 to 36.9 in 2017 with a peak of 39.3 in 2014.

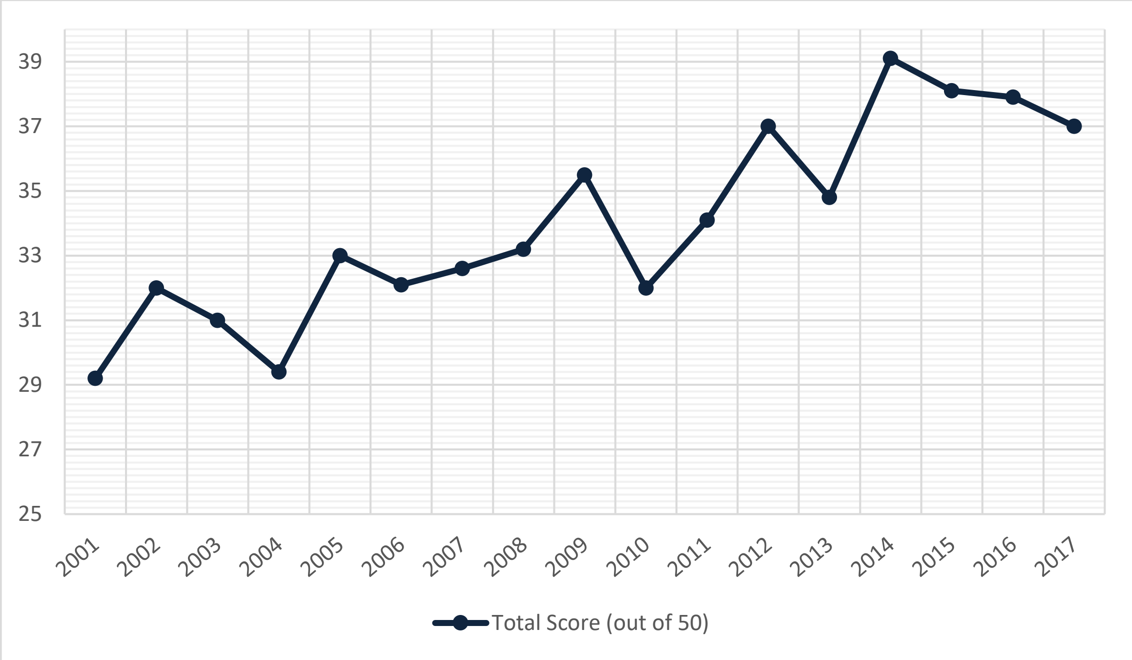

Figure 2: Average total score (out of 50) in the Higher diagnostic test of all A level Mathematics students, 2001 to 2017.

Given that there was an increase in the proportion of students achieving higher grades from 2001 to 2017, it would be expected that the average mark would increase. In order to give a visual representation of the changes in overall distribution of entry grades, we use UCAS tariff points[3] (UCAS is the body that administers university admissions in the UK) as a way of comparing different qualifications. The tariff which was in use until 2016 gave 140 points for an A level grade A*, 120 points for an A level grade A, 100 points for an A level grade B, etc.; ie reducing the points value by 20 for each grade lower. In 2017, UCAS revised its tariff so that an A level grade A* was worth 56 new tariff points, and the value reduced by 8 points per grade (ie grade A 48 points, grade B 40 points, etc.). In order to maintain consistency throughout our analysis, we have used the old tariff in all years.

[3] See https://www.ucas.com/ucas/tariff-calculator for further details.

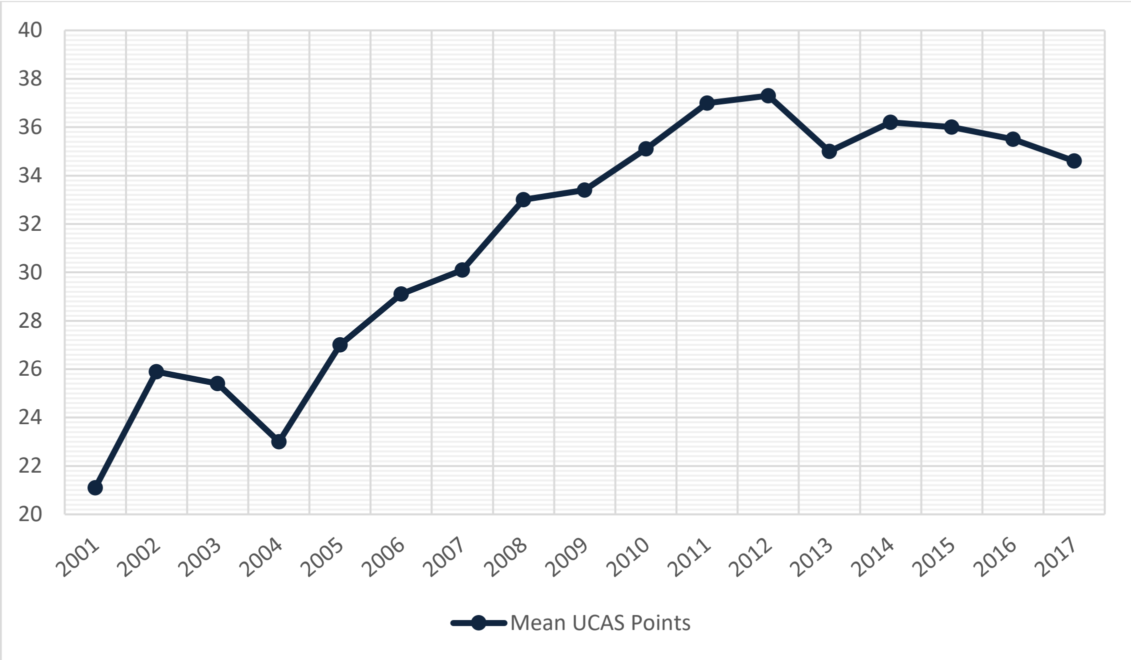

Figure 3: Average (old) UCAS tariff points in Mathematics of A level students taking the Higher diagnostic test at Coventry University, 2001 to 2017.

We can see from Figures 2 and 3 that both graphs display broadly similar trends, increasing substantially from 2001 to a maximum somewhere in the early to mid-2010s and then decreasing slightly.

In order to explore this further, we model the performance of each year's cohort using the results from 2001 as a baseline. Table 1 gives the grade cohort results (out of 50) in 2001.

A	B	C	D	E	N	U
38.42	33.82	32.69	29.25	28.38	24.84	23.08

Table 1: 2001 Grade cohort average scores in the Higher diagnostic test.

We use the 2001 grade cohort scores and the actual numbers of students with each grade at A level to predict the performance of the overall cohort average score in subsequent years. We use the following notation:

$A_1, B_1, C_1,$ etc. are the average score for the 2001 grade A cohort, grade B cohort, grade C cohort, etc. (as shown in Table 1);

NA_i, NB_i, NC_i, etc. are the number of students in the year 2000+i entry cohort with grade A, grade B, grade C, etc.;

Ni is the total number of students in the year 2000+i entry cohort i.e. the sum of NA_i through to NU_i;

P_i is the predicted whole cohort weighted average score in year 2000+i.

Then $P_i = \frac{1}{N_i}[NA_i.A_1 + NB_i.B_1 + NC_i.C_1 + ND_i.D_1 + NE_i.E_1 + NN_i.N_1 + NU_i.U_1]$

P_i is therefore what the whole cohort average score would be in year 2000+i if the performance of each grade cohort that year was the same as it had been in 2001. If the actual cohort average is greater than the predicted value then that year's cohort is performing better on the diagnostic test than the 2001 cohort; if the actual cohort average is less than the predicted value then that year's cohort is performing worse than the 2001 cohort. This is visualised by plotting the graph of "(Actual / Predicted) – 1". Years where this value is positive are performing better than 2001, years where the value is negative are performing worse.

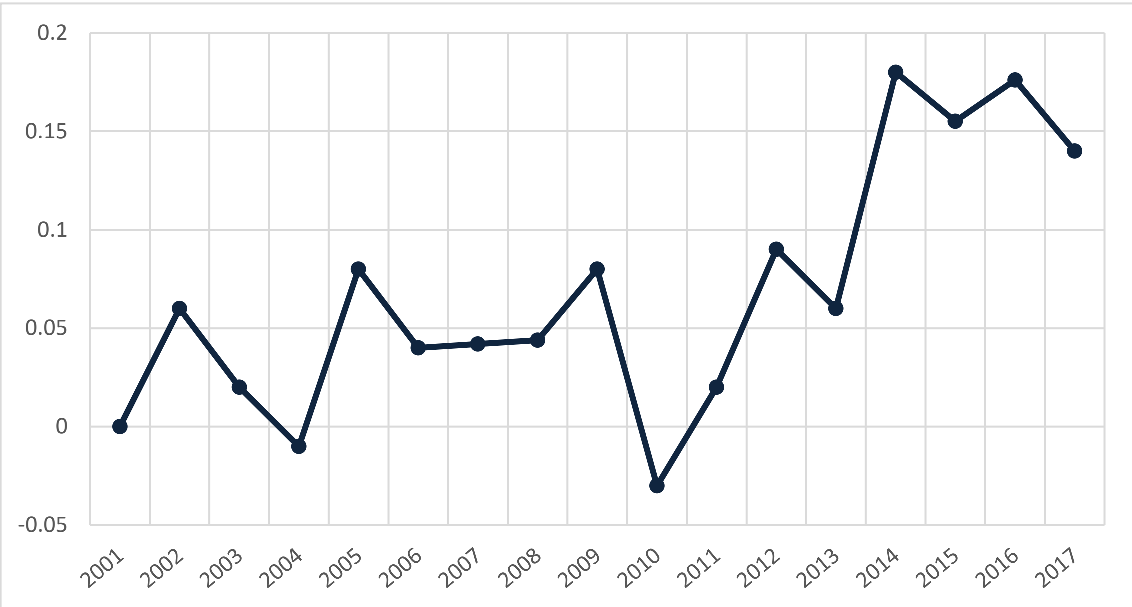

Figure 4: "(Actual/Predicted) – 1" for the year cohorts from 2001 to 2017

Clearly in 2001, the actual and predicted values are identical, so the value of "(Actual / Predicted) – 1" is equal to zero. Thereafter, we see from Figure 4 that, after allowing for the changes in the distribution of the grades of the incoming cohort, there is a general upward trend (with a major blip in 2010) from 2001 to 2014 (where the performance is approximately 18% better than in 2001) followed by a small tailing off. This suggests that the general performance of students in the diagnostic test has improved since 2001, after controlling for the increasing numbers of students with higher grades coming to the institution.

Performance of students with lower grades (grades D and E)

Figure 5 shows the average performance of the 728 students entering with A level Mathematics grades D (510) and E (218) (the traditionally 'at risk' students, according to Lawson (2003)) over the period from 2001 to 2017.

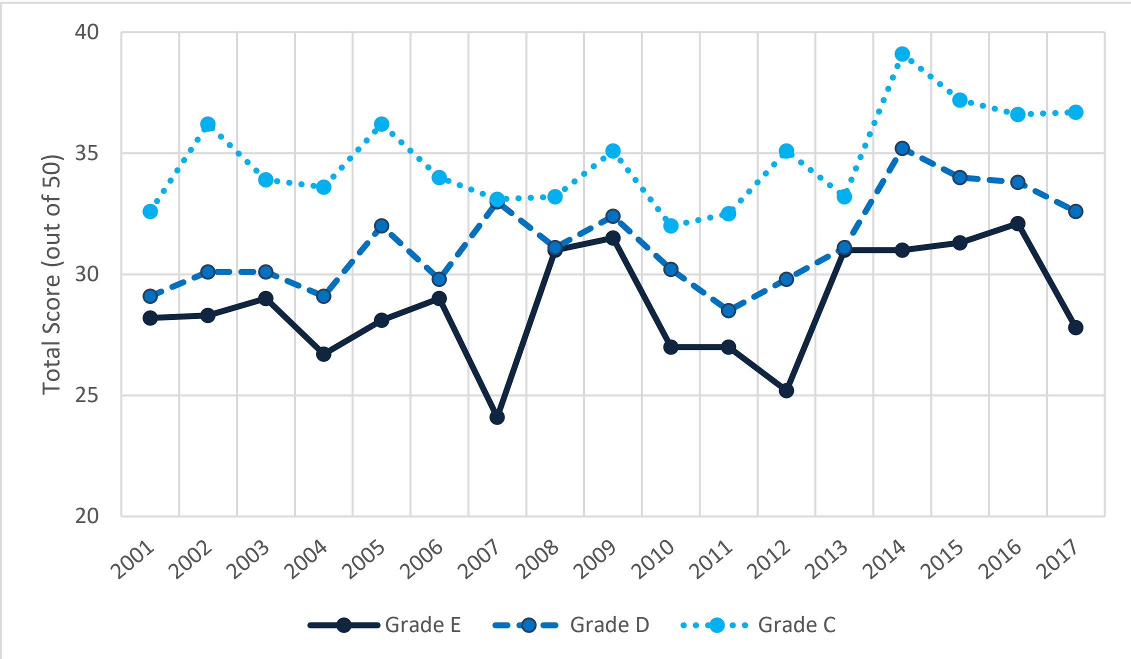

Figure 5: Average total score (out of 50) in the Higher diagnostic test of the students entering with A level Mathematics grades D and E, 2001 to 2017. [The performance of the grade C cohort is shown for reference to higher grades.]

Over these 17 years, grade E students (M = 28.67, s.d. = 7.06) perform statistically significantly worse than grade D students (M = 31.72, s.d. = 6.65), t(726) = 5.52, p < 0.001. This result is as would be expected. However, in line with the overall performance reported previously, there is evidence to suggest the performance of all students with lower grades has generally improved. Indeed, when taking these students with lower grades as a whole (the small number of grade E students in later years precludes doing this analysis for the grades separately), there is enough evidence to suggest a positive linear relationship (r = 0.23, p < 0.001) between year and these students' average performance in the diagnostic test. This shows a small increase, on average, of about 0.27 correct answers (p < 0.001) each year.

Lawson (2003) considered the homogeneity of the cohort of students with lower grades. When teaching a cohort, it is useful for the tutor to be aware of topics in the students' prior education in which the whole cohort are competent and topics in which not many students are competent. To investigate this, he looked at the number of questions where more than 90% of the students with lower grades had answered correctly (i.e. topics where almost everyone was competent) and the number of questions where less than half the cohort had answered correctly (i.e. topics where the majority lack competence). These results for the 2001 to 2017 cohorts of students with lower grades are shown in Figure 6. Also shown in Figure 6 is the homogeneity index (explained below) for the students with lower grades.

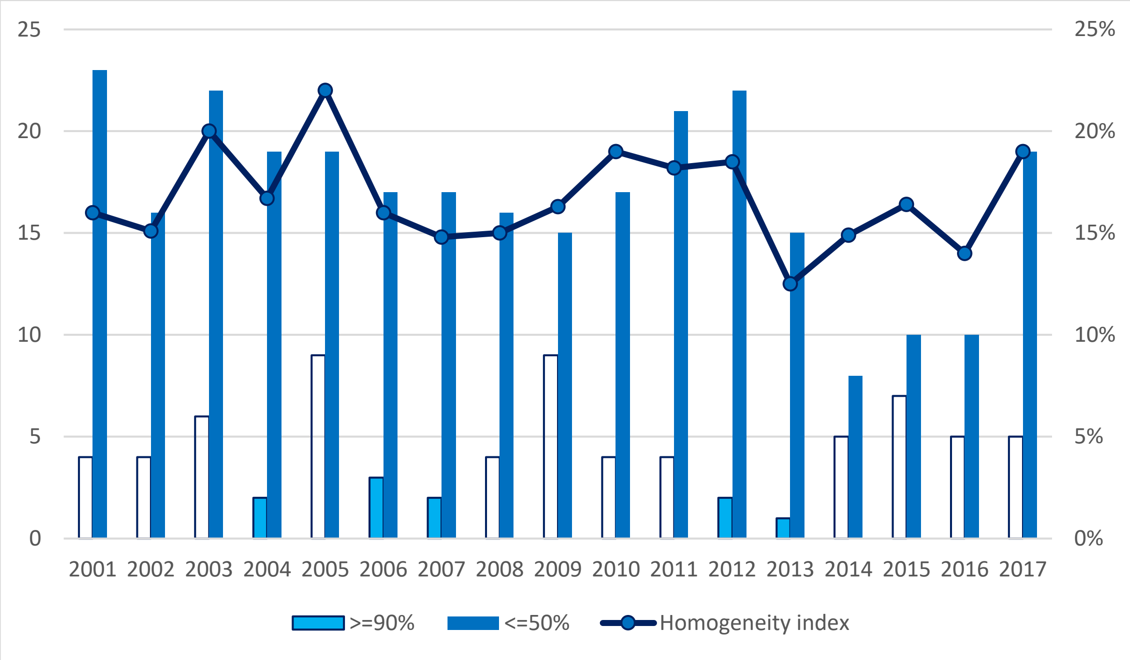

Figure 6: The number of questions answered correctly by more than 90% and by less than 50% of the cohort (left axis) and the homogeneity index (right axis) for the grades D and E cohort over the period 2001 to 2017.

There is no discernible pattern in the number of questions answered correctly by more than 90% of the students. The highest value (9) occurs twice (in 2005 and 2009) and the lowest value (1) occurs in 2013. The average number of questions answered correctly is close to 30 each year. This indicates that there is little consistency across the cohort in terms of questions that almost all students with lower grades can answer correctly. In terms of consistency between year cohorts, there are only two questions that were answered correctly by more than 90% of the students with lower grades in more than 10 of the 17 year-cohorts from 2001 to 2017. These questions are:

- to calculate the value of a 5% increase in 250;
- to evaluate the expression $\frac{3x^2-1}{4-x}$ when $x = -2$.

Likewise, the number of questions answered correctly by less than 50% of the students shows no consistent trend over the period 2001 to 2017. Since the average number of questions answered correctly each year is around 30, then having 20 or more questions answered incorrectly by over 50% of the cohort might be taken as an indication of some level of homogeneity. There are four questions that fall into this category in each of the 17 year-cohorts under consideration. These questions are:

- to simplify $\frac{2x^{-3}y}{x^2y^{-2}}$
- to determine the partial fraction form of $\frac{1}{(x+1)(x+2)}$ [see footnote [4]]
- to determine in which quadrant the angle θ lies if its tangent is positive and its sine is negative
- to identify the correct expansion of sin (*A*-*B*).

In order to measure the homogeneity of a given cohort, a homogeneity index has been developed and this is explained in Appendix 2. The homogeneity index for the year-cohorts of students with lower grades for the years 2001 to 2017 are shown in Figure 6.

We see that over the period in question, the highest values (22% in 2005 and 19.9% in 2003) of the homogeneity index occur in the early years, with the lowest values (12.1% in 2013 and 14.2% in 2016) towards the end of the time period. This suggests that this cohort is becoming less homogenous as time progresses.

Performance of students with higher grades (grades A* to C)

Figure 7 shows the average total score (out of 50) in the Higher diagnostic test of all the students entering with A level Mathematics grades A* to C from 2001 to 2017. There were 2353 students who entered Coventry University with such grades (A* (24), A (400), B (936), and C (993)) over the period from 2001 to 2017. A one-way ANOVA of their results with a Tukey post-hoc test showed that students with the highest grades

[4] Although partial fractions is no longer on the A level syllabus, since the diagnostic test is multiple choice it is possible to work back from each of the four given partial fraction expansions to determine which is equal to the given expression.

(A*: M = 46.29, s.d = 2.46 or A: M = 39.65, s.d = 7. 20) perform significantly better than those with other grades (B: M = 37.86, s.d = 6.79 or C: M = 35.31, s.d. = 6.56); with $F(3,2349) = 62.155$, $p < 0.001$. Again, this is an expected result. To explore this further, a multiple linear regression was conducted. From this analysis, we can conclude that students with these higher grades perform, on average, 0.35 questions better when compared to the previous year, $t(2350) = 9.85$, $p < 0.001$, and students in each grade achieve, on average, 2.28 more correct answers than students who have one grade lower, $t(2350) = 12.37$, $p < 0.001$.

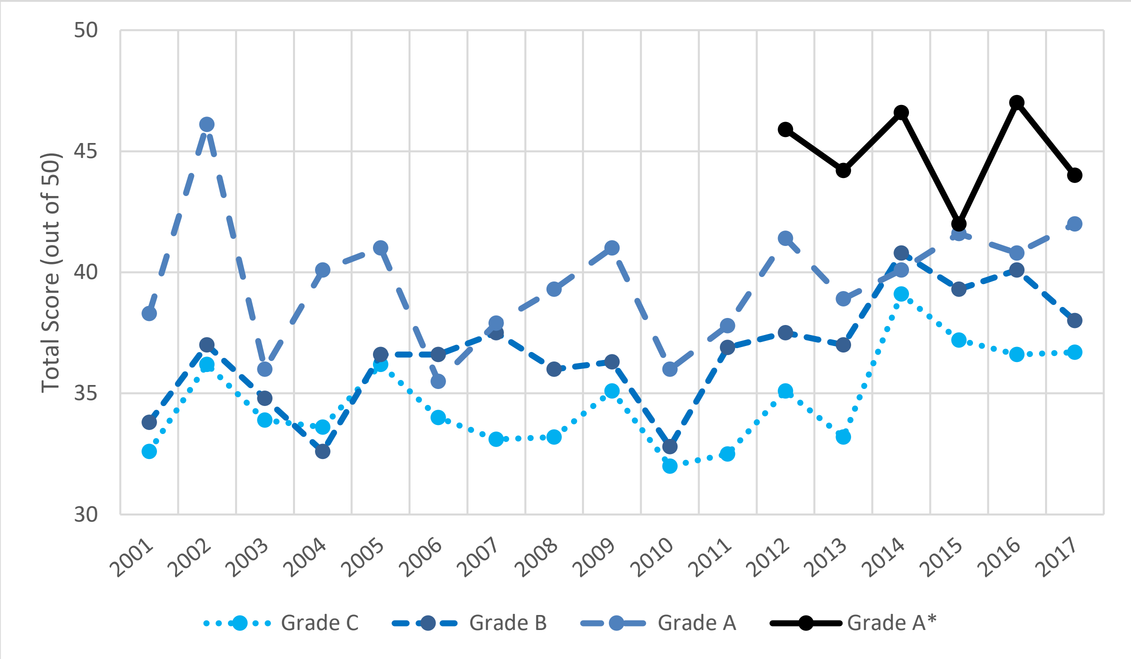

Figure 7: Average total score (out of 50) in the Higher diagnostic test of the students entering with A level Mathematics grades A to C, 2001 to 2017.*

As for the students with lower grades, cohort homogeneity was investigated for students with higher grades by looking at the number of questions answered correctly by more than 90% of the cohort or incorrectly by over half the cohort. Figure 8 shows this information and the homogeneity index for the students with higher grades.

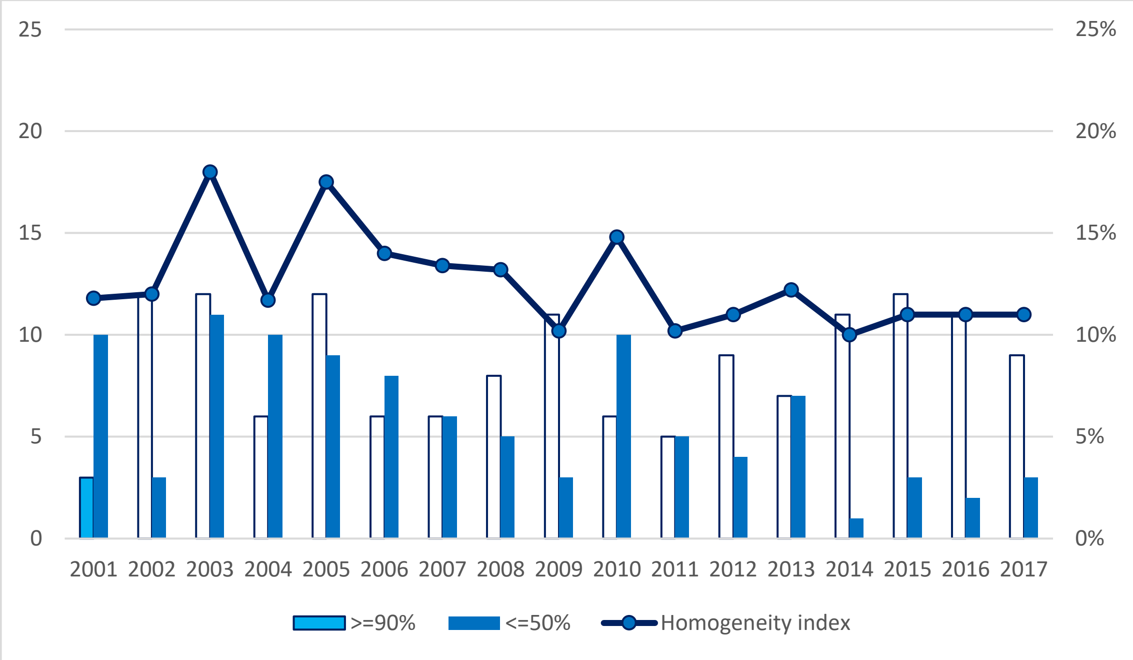

Figure 8: The number of questions answered correctly by more than 90% and by less than 50% of the cohort (left axis) and the homogeneity index (right axis) for the grades A-C cohort over the period 2001 to 2017.*

The average performance for this cohort varies from 34 out of 50 in 2001 to 40 in 2014. In nine of the 17 year-cohorts, at least nine questions (around a quarter of the average number of correct answers) were answered correctly by 90% or more of the cohort indicating some level of homogeneity. In terms of year on year consistency, one question was answered correctly by over 90% of the Grade A*- C cohort in all 17 years, one question fell into this category in 16 years, and a further question in 15 years. These questions are:

- to evaluate the expression $\frac{3x^2-1}{4-x}$ when $x = -2$ (all 17 year cohorts);
- to simplify $\frac{x}{3} + \frac{y}{2}$ (16 year cohorts);
- to determine the derivative of $y = 5x^4$ (15 year cohorts).

It is noteworthy that the percentage increase question, which was one of the two questions answered correctly by over 90% of the cohort of students with lower grades most often (ten times), is not in the list above. This question passed the 90% threshold amongst students with higher grades in 14 year cohorts, along with two other arithmetic questions.

Students in this cohort, on average, answer approximately 14 questions incorrectly. With the exception of 2010, there is a trend that the number of questions answered incorrectly by at least half the cohort, reduces as time goes on. Year on year, there is some consistency about questions that fall into this category. There are four questions that feature most often – these are:

- to determine in which quadrant the angle θ lies if its tangent is positive and its sine is negative (16 year-cohorts);
- to identify the correct expansion of sin(*A*-*B*) (13 year-cohorts)
- to determine the feasibility of constructing triangles of given dimensions (13 year-cohorts)
- to simplify $\frac{2x^{-3}y}{x^2y^{-2}}$ (10 year-cohorts).

This list includes the three questions that featured every year for the students with lower grades and one further question. This additional question showed two triangles: the first was shown with angles of 80^{o}, 70^{o} and 40^{o}; the second was shown with sides of length 2, 3, and 6. The question asked whether it was possible to construct only triangle 1, only triangle 2, neither triangle or both triangles. Interestingly, this question is answered slightly better by students from the lower grades cohort than by students from the higher grades cohort. For the lower grades cohort, this was answered correctly by less than half the cohort in nine of the 17 years under consideration and, over all of the 17 years, was answered correctly by 56% of these students compared to only 49% of the students with higher grades. It is possible that the students with lower grades outperformed higher grade students on this question because it is the only one on the whole test with a labelled diagram. Helping students create visualisations of a mathematical problem has been shown to improve problem solving performance (Carden and Cline, 2015) which is made significantly easier when the visualisation is provided. In turn, this may help students provide their own explanations on how to solve the problem, again shown to improve understanding and performance of students known to struggle with mathematics (e.g. Hodds, 2020; Wong, Lawson, and Keeves, 2002). Without the visualisations, students with lower grades may struggle to generate such understanding and explanations and therefore struggle on the more traditional type questions which contain only words and mathematical symbols.

Figure 8 also shows the homogeneity index for the higher grades. As with lower grades students, we see that the highest values of the homogeneity index (18.7% in 2003 and 18.2% in 2005) occur in the early years of the period studied. The lowest values (10.0% in 2014 and 10.7% in 2011) occur in the later years. Comparing the information in Figures 6 and 8, we see that overall the homogeneity index for students with lower grades is actually slightly higher than for students with higher grades.

Overall comparison of all students

In order to investigate the effect of the A level grade and the year in which students entered university, a multiple regression analysis was conducted. In this analysis, the following coding was used for the grades: A* and A: 5, B: 4, C: 3, D: 2, E: 1. The grades A* and A were put together because grade A* was not introduced until 2012 – the highest grade that could be achieved until then was grade A. Therefore grade A cohorts in years 2001 to 2012 may contain students who would have achieved grade A* had this been available. The estimated model from this multiple regression analysis is:

$$\text{Total Score} = 24.34 + 0.30 * (\text{Year-2001}) + 2.49 * \text{Grade}.$$

This indicates that over the whole population of students being considered from 2001 to 2017, and over all grades, the total score increases on average by 0.3 marks per year, confirming the improvement previously noted. Similarly, over all of the years, the total score increases by, on average, 2.49 marks per A level grade.

The information displayed previously in Figures 5 and 7 suggests that there is an upward trend from 2001 to 2017 for each A level grade. To investigate this further, linear regression models were developed for each grade. These are shown in Table 2 below.

Grade	Grade code	Observed Mean	Standard Deviation	N	Estimated Model
A* - A	5	40.03	7.18	424	36.47 + 0.31 * (Year – 2001)
B	4	37.86	6.69	936	32.65 + 0.45 * (Year – 2001)
C	3	35.31	6.56	993	32.15 + 0.29 * (Year – 2001)
D	2	31.72	6.65	510	29.41 + 0.25 * (Year – 2001)
E	1	28.69	7.06	218	27.89 + 0.14 * (Year – 2001)

Table 2: Descriptive statistics and estimated model for test score by A level grade.

These models indicate that, for each grade, student performance improves over the years, as suggested previously. The models also show that the improvement is faster for students with higher grades than for students with lower grades. Therefore, this suggests that the gap between students with higher A level grades and students with the lower grades is getting wider. This is particularly the case for students with grade E whose rate of improvement (0.14 marks per year) is considerably less than for students with grade D (0.25 marks per year). Furthermore, an interesting result is that the performance of students with a grade B is improving at the fastest rate (0.45 marks per year); this is discussed further in the conclusions below.

Discussion and Conclusions

This paper set out to update the work of Lawson (2003) by investigating the competencies of students entering Coventry University between 2001 and 2017. In Lawson (2003), the results showed that the entry competencies of students declined over the period 1991-2001, so much so that students with grade B in 2001 were performing at the same level as those with a grade N in 1991. The results of this paper have shown that the entry competencies of students between 2001 and 2017 have generally improved, although they have not recovered to the levels achieved in 1991, as shown in Table 3 (N.B. There is no entry for grade A*- A in 1991 since only one student with a grade A took the Higher diagnostic test that year). Indeed, the rate of improvement seen between 2001 and 2017 is slower than rate of decline that was seen between 1991 and 2001.

Year	Grade A* - A	Grade B	Grade C	Grade D	Grade E
1991	--	40.5	39.9	37.3	35.7
2001	38.4	33.8	32.7	29.3	28.4
2017	42.3	38.6	36.6	32.7	27.8

Table 3: Grade cohort scores in 1991, 2001 and 2017.

Similar analyses of changes over time in the competences of incoming undergraduates have been carried out in Ireland using results from a diagnostic test administered at the University of Limerick. The work of Faulkner, Hannigan and Gill (2010) shows that, although there is a small decline in the performance of students with the same entry grades over the period 1998 to 2008, the change is not statistically significant. However later work by Treacy and Faulkner (2015) shows a significant drop in performance by students with both higher leaving certificate and ordinary leaving certificate mathematics qualifications over the period 2003 to 2013. These trends in Ireland are at variance with the findings reported here in relation to A level mathematics in England over the period 2001-2017 where the performance of equally well qualified new undergraduates has improved.

Students with a grade A* to C in A level Mathematics are now performing considerably better on the diagnostic test than they were in 2001. In particular, students with a grade B have rapidly improved over the period and by 2016 were performing almost as well as students with an A or A* (grade B average 40.1, grade A*-A average 41.7), however the gap increased a little in 2017. Indeed, from the linear model described in section 5 it can be seen that grade A*- A students are improving by 0.31 marks each year, grade C students by 0.29 marks per year, which is a similar rate, but grade B students are improving by a much larger 0.45 marks per year. It is possible that there is a ceiling effect on the grade A*-A students – their rate of improvement may be lower than that of grade B students since they have less room for improvement.

Although those students with a grade D or E in A level Mathematics have also improved, they have improved at a slower rate than students with higher grades. There are far fewer students coming to Coventry University with these lower entry grades than there were in 2001 but those that do appear to be falling further behind their colleagues with higher grades. Indeed, the gap in entry competencies between students with higher A level Mathematics grades and students with lower A level Mathematics grades has dramatically increased since 2001.

We can only speculate as to why a general improvement in average diagnostic test score has occurred. As stated in Lawson (2003, p. 174), "...the diagnostic test is not attempting to replicate A level assessment...", rather the diagnostic test seeks to help students understand where their strengths and weaknesses are in the basic mathematical skills that are required for their courses. Having said this, we might speculate that the improvement may be because of changes to the overall A level structure. Between 2005 and 2017, the pure (or core) mathematics in the A level syllabus was taught over 2 years in 4 modules alongside 2 applied modules (usually in Mechanics or Decision and Statistics). The diagnostic test only covers topics from the pure (core) part of the syllabus. Prior to this, between 2001 and 2005, the pure mathematics was taught over 3 modules. This meant students taking A level Mathematics after 2005 were able to spend more time on each topic and perhaps obtain a deeper understanding. Furthermore, once the syllabus and examination method was settled, teachers gained a better understanding of what was required of their students and were therefore able to better prepare students for their examinations. This perhaps provides a reason why there has been an improvement for all students since 2001 but it does not explain why the gap between the higher grades and lower grades has increased.

In an ideal world, all students with grade E would be able to answer correctly the same set of questions on the diagnostic test; all students with grade D would be able to answer correctly this set of question plus a further set of questions; all students with grade C would be able to answer correctly the questions grade D students can answer plus a further set, etc. However, reality is not as clear cut as this. The homogeneity index gives a measure of how consistent the competencies of a particular cohort are. It has been shown in this paper that the cohorts of students with lower grades have higher values of the homogeneity index (by around 4 percentage points) than the cohorts of students with higher grades. This may be because students with lower grades score most of their marks on the "easier" questions (those that rely only on GCSE level mathematics) whereas the students with higher grades score highly on these questions and then score further marks across the "harder" questions (but without high levels of consistency about which "harder" questions they answer correctly). This however is informed speculation and further investigation is needed to confirm this.

A final consideration is what the future will bring. The next cohort of students entering university will have obtained their A levels through the newest change to the curriculum. In the new curriculum, assessment is no longer by module examinations taken during the two years of study but in a single, so-called, linear examination at the end of the two years. Consequently, students have to ensure they understand the whole A level syllabus when they sit their examinations at the end of their second year. It is possible that this may have a detrimental effect on weaker students since they may struggle to retain knowledge over the two years of the course and begin to feel overwhelmed. Alternatively, it may be that the weaker students will actually improve because they will need to retain the knowledge instead of only needing it for one module at a time. They will no longer be able to set aside earlier parts of the curriculum as they may have done in previous years. Indeed, retaining knowledge to use in later modules is a problem many mathematics students face whilst at university, so it is possible the new changes will have a positive impact on knowledge retention and understanding. Of course, there may also be positive benefits for stronger students since they will need to have a longer, deeper understanding of the material and be forced to make important connections between the topics. Finally, there may also be fewer students taking A level Mathematics and Further Mathematics due to the changes in the general structure of A levels. One of the changes in the new A level structure is to make AS and A levels completely separate qualifications (previously AS had been the first half of an A level). Under the former structure, students were encouraged to take four subjects at AS level with the option to drop one at the end of Year 12 and proceed with three A levels in Year 13. This made AS level mathematics popular with those students whose university study aspirations would benefit from studying some mathematics post-16 (such as geography, biology and psychology) but which did not require a full A level. Similarly, AS level Further Mathematics was popular with those aspiring to engineering and physics courses of a more theoretical nature. Some of these students, originally intending only to study Mathematics or Further Mathematics to AS level, finding an aptitude and interest in the subject continued it to A level. Under the new structure, such an approach is at best strongly discouraged and, in many instances, actually not possible because of the way the curriculum is delivered locally. Hence it is anticipated that the numbers taking AS and A level Mathematics and Further Mathematics will drop. What impact this will have on universities, university entry requirements, and first year undergraduate course content remains to be seen. If students' entry competencies in mathematics are going to change again, higher education providers will need to consider how to best support and guide them through their degrees when they arrive at our institutions. It should also be noted that as this manuscript was written before the impact of Covid-19 disruption on teaching, we plan to update our findings in a few years' time which may also provide some evidence to the suggestions above.

References

Carden, J. and Cline, T. (2015) Problem solving in mathematics: the significance of visualisation and related working memory, Educational Psychology in Practice, 31(3):235-246.

Edwards, P. (1996) Implementing diagnostic testing for non-specialist mathematics courses, OLF, London.

Faulkner, F., Hannigan, A. and Gill, O. (2010) Trends in the mathematical competency of university entrants in Ireland by leaving certificate mathematics grade. Teaching

Mathematics and Its Applications, 29(2):76-93.

Hodds, M. (2020) Supporting nursing students' mathematical understanding, MSOR Connections, 18(3): 23-29.

Hunt, D.N and Lawson, D.A. (1996) Trends in the mathematical competency of A level students on entry to university. Teaching Mathematics and Its Applications, 15(4):167-173.

JCQ (2018) A level results remain steady and STEM subjects increase in popularity, JCQ Media Release. Available at https://www.jcq.org.uk/media-centre/news-releases/gce-press-notice-uk-2018 (accessed 16 January 2020).

Lawson, D. (2003) Changes in student entry competencies 1991-2001, Teaching Mathematics and Its Applications, 22(4):171-175.

MathsTEAM (2003) Diagnostic testing for mathematics, LTSN, Birmingham.

Savage, M. (2003) Tackling the maths Problem: Is it far more extensive than we thought? In Proceedings of 4th IMA Conference on the Mathematical Education of Engineers, Hibberd, S. and Mustoe, L. (eds), April 2003, Loughborough.

Smith, A. (2004) Making mathematics count, The Stationery Office, London.

Todd, K.L. (2001) An historical study of the correlation between GCE Advanced Level grades and the subsequent academic performance of well-qualified students in a university engineering department. Mathematics Today, 37:152-156.

Treacy, P. and Faulkner, F. (2015) Trends in basic mathematical competencies of beginning undergraduates in Ireland, 2003-2013. International Journal of Mathematical

Education in Science and Technology, 46(8):1182-1196.

Wong, R., Lawson, M. and Keeves, J. (2002) The effects of self-explanation training on students' problem solving in high-school mathematics. Learning and Instruction, 12(2): 233-262.

Appendices

Appendix 1 – Numbers of students in each A level mathematics grade per year

Year	A*	A	B	C	D	E	Lower than E	Total
2001	0	7	11	26	36	45	31	156
2002	0	4	14	24	34	27	6	109
2003	0	1	9	11	10	12	4	47
2004	0	3	5	16	19	21	9	73
2005	0	2	5	13	15	5	3	43
2006	0	3	21	25	26	15	1	91
2007	0	14	15	34	23	13	4	103
2008	0	14	25	32	24	5	3	103
2009	0	27	39	68	39	12	0	185
2010	0	40	74	68	31	8	5	226
2011	0	57	104	86	27	8	0	282
2012	2	31	71	57	15	3	0	179
2013	2	23	78	97	36	1	1	238
2014	3	48	117	112	28	9	2	319

2015	1	39	119	83	38	6	1	287
2016	13	59	141	143	66	17	0	439
2017	3	28	88	98	43	11	1	272
Total	24	400	936	993	510	218	71	3152

Appendix 2 – Development of the homogeneity index

When teaching a cohort of students, the ideal is that they all have the same set of skills. In terms of the diagnostic test (assuming that they are not all going to answer all the questions correctly), it is preferable if they all answer the same questions correctly (and by implication all answer the same questions incorrectly). This would be a perfectly homogeneous cohort – the average score of the cohort on every question is either 1 or 0.

The opposite situation is one where some students answer some questions correctly and other students answer a different set of questions correctly. Then there is no prior knowledge shared across the cohort. For a cohort where the overall average mark is M, then a perfectly inhomogeneous cohort is one where the average score on each question is $M/50$, indicating that different students have competency in different areas.

We use these ideas to formalise the idea of a homogeneity index.

Suppose a cohort of students has completed the 50 question diagnostic test and the overall cohort average is M, with the proportion of students answering question i correctly being x_i.

Let $\bar{x}$ be the mean of the x_i values (i = 1, …, 50)

Then $M = 50\bar{x}$.

A perfectly homogeneous cohort would have $x_i = 1$ for M questions and $x_i = 0$ for the other 50-M questions. If we order the questions so that those answered correctly come first, we can represent this cohort by a 50-vector

$\boldsymbol{x_H}$ = (1, 1, …, 1, 0, 0, …, 0)

Likewise, a perfectly inhomogeneous cohort would have $x_i = M/50 = \bar{x}$ for $i = 1, \ldots, 50$ and we represent this by a 50-vector

$$\boldsymbol{x_I} = (\bar{x}, \ldots, \bar{x})$$

By using the standard Euclidean norm, we can calculate the distance-squared between the perfectly homogenous cohort ($\boldsymbol{x_H}$) and the perfectly inhomogeneous cohort ($\boldsymbol{x_I}$) which we call the maximum homogeneity.

Maximum homogeneity = $M(1-\bar{x})^2 + (50-M)(0-\bar{x})^2$

After simplifying (using $M = 50\bar{x}$) we find

Maximum homogeneity = $M(1-\bar{x}) = 50\bar{x}(1-\bar{x})$

For a general diagnostic test vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_{50})$ we can measure the distance-squared of this vector from the perfectly inhomogeneous cohort $\boldsymbol{x_I}$ and then express this as a percentage of the maximum homogeneity. We call this the homogeneity index.

Distance-squared from $\boldsymbol{x_I}$ = $\sum_{i=1}^{50}(x_i - \bar{x})^2$

Homogeneity index = $100\,\frac{\sum_{i=1}^{50}(x_i-\bar{x})^2}{50\bar{x}(1-\bar{x})}$

The larger the homogeneity index, the more homogeneous the cohort. A cohort with $50\bar{x}$ values of 1 and $50(1-\bar{x})$ values of 0 would have a homogeneity index of 100%.

In essence, we are measuring how far away from the perfectly inhomogeneous cohort our general cohort vector is and then comparing this to the furthest away it is possible to be (ie the perfectly homogeneous cohort). Defining the homogeneity index in this way allows fair comparison of the homogeneity of cohorts with different average total marks since it is a measure of how close a given cohort is to the maximum possible homogeneity for a cohort scoring that average total mark.

It is important to note that the maximum homogeneity depends on the average score of a cohort and this is why it is necessary to construct a measure such as the homogeneity index, which is independent of the average score of the cohort.

It should also be noted, that for most cohorts $50\bar{x}$ will not be an integer and so the most homogeneous cohort cannot have $50\bar{x}$ values of 1 and $50(1-\bar{x})$ values of 0, but this does not invalidate the final formula for the homogeneity index given above.