TARGETING QUESTIONS TO GRADES IN GCSE MATHEMATICS B (MARCH 2007)

Ashleigh Seldon

Background

For GCSE Mathematics B, the Qualifications and Curriculum Authority (QCA) provide grade descriptors indicating the level of attainment characteristic of grades A, C and F. These are designed to provide a general indication of the anticipated learning outcomes at each grade. All awarding bodies are issued with the same grade descriptors, the most recent version of which is given in Appendix 1. Awarders are asked to refer to these descriptors when making judgemental decisions. In addition to grade descriptors, QCA specifies that for GCSE Mathematics 50 *per cent* of the marks on any question paper should be targeted at the bottom two grades. The remaining marks can be allocated as individual awarding bodies deem appropriate. In AQA's new two-tier Mathematics B specification (4302), for example, question papers are written with the mark allocation 15 *per cent*, 15 *per cent*, 20 *per cent*, 20 *per cent* and 30 *per cent* in mind. Information regarding what constitutes a suitable question targeted at each of the available grades is shared within the forum of an inter-awarding body committee. This committee also decides upon common coursework boundaries, which all awarding bodies must adhere to.

With the transition from a three-tier (3302) to a two-tier specification (4302) in Mathematics B, there is evidence that the orthodoxy of targeting questions to grades has hindered the awarding process, with awarders having pre-defined expectations of the difficulty of questions targeted at particular grades. In December 2006, for example, the Director General stated in his report for the award that:

> *"It is time for a sustained challenge to be made on the orthodoxy which has grown up in Mathematics about the wisdom of targeting particular questions at particular grades. All the evidence suggests that this is a hindrance to the maintenance of comparable standards for reasons well-documented in the assessment literature."*
> (Director General, December 2006)

The purpose of this paper is to assess how targeted questions perform, and whether it is advisable for awarders to draw on information regarding targeted questions when making judgemental decisions.

Method

Item level data for the three-tier Module 3 Intermediate paper (33003/I) and both two-tier Module 3 papers were obtained (43003/F, 43003/H) for March 2007. Module 3 formed the focus of the investigation as the two-tier version of the module proved particularly difficult to award in both December 2006 and March 2007. In both series, recommended grade changes were made with a view to making the module outcome less severe. Only the Intermediate paper on the three-tier specification has been considered, as this tier has historically had the largest entry. In March 2007, for example, 27,343 candidates entered for the Module 3 Intermediate tier, compared to 3,822 on the Foundation tier, and 3164 on the Higher tier.

The analysis conducted on the item level data for each paper was threefold:

1. Candidates' performance (in terms of raw marks) *across all questions targeted at a particular grade* by final module grade was displayed graphically (see Appendix 2).
2. Candidates' performance (in terms of raw marks) on *each item* by final module grade was also graphed, in order to ensure that compensatory influences had not been masked by only examining candidates' scores across a number of items targeted at a certain grade[1].
3. Candidates' performance (in terms of mean mark and mean mark as a percentage of the maximum mark/facility index) on the questions targeted at their final module grade was assessed. So, for example, A* candidates' mean mark on each of the questions targeted at A* was calculated (Tables 1-3)

Results

Three-tier specification

Intermediate paper (33003/I)

Figures 1-4 (see appendix 2) show the total mean mark that candidates' achieved across questions targeted at each grade. For each group of targeted grade items a linear relationship was evident between final module grade and total mean mark. Analysis of each individual item revealed a similar pattern. Since there were no instances where C grade candidates, for example, scored more marks on an item targeted at grade B (or across a number of items targeted at grade B) than B grade candidates did, one may argue that targeted questions worked well in this instance.

However, when candidates' performance on items targeted at the final grade they achieved on Module 3 in March 2007 was considered, a different pattern emerged. Table 1 shows the mean mark and facility index (mean mark as a percentage of the maximum mark) that candidates awarded each grade at module level obtained on the questions targeted at that grade. So, for example, B grade candidates' performance on each of the B grade questions is given. One may expect that candidates that achieved a particular grade at module level may have displayed a similar level of ability on each of the questions targeted at that grade. However, as Table 1 demonstrates, this was clearly not the case. Facility indices for the B grade candidates, for example, varied from 31.7 *per cent* to 90.0 *per cent.* Grade E candidates showed the greatest degree of variability across questions targeted at their final grade (1.5 – 83.0%) and grade C the least (24.7 – 74.7%). Facility indexes across all questions targeted at each grade ranged from 40.0 *per cent* for grade E candidates to 54.7 *per cent* for grade D candidates. This suggests that on average candidates obtained around half of the marks targeted at the final grade they received.

[1] For reasons of prudent practice, individual graphs for each item have not been included in an appendix.

Two-tier specification

Foundation paper (43003/F)

As with the three-tier Module 3 Intermediate paper, candidates' scores on items on the two-tier Module 3 Foundation paper increased with module grade awarded (see Appendix 2, Figures 5-9). However, candidates' scores on the items targeted at their final grade varied substantially (Table 2). In terms of mean mark as a percentage of the maximum mark, grade F candidates showed the greatest degree of variability across targeted questions (2.0 – 82.0%) and grade D candidates the least (32.3 – 79.0 %). Facility indices across all questions targeted at each grade indicated that grade E candidates only obtained on average a quarter (25.5 %) of the marks targeted at the grade. The most successful candidates were grade G candidates, who obtained nearly two-thirds (65.9%) of the marks targeted at this grade.

Higher paper (43003/H)

Candidates' scores on items on the two-tier Module 3 Higher paper showed a positive linear relationship with module grade awarded. As expected, this was also true for groups of targeted items (Appendix 2, Figures 10 – 14). Table 3 gives the mean mark and facility indices that candidates achieved for items targeted at their final module grade. Candidates' performance varied across targeted items. Grade D candidates scored between 12.0 *per cent* and 93.0 *per cent* of the marks available on targeted grade D items (a range of 81.0%). Grade A* candidates exhibited the least amount of variability with regards to scores on targeted items (43.7 – 83.5%). Across all targeted questions, A* candidates obtained over two-thirds (68.6 %) of the marks targeted at this grade, whilst grade A candidates only obtained over a third of targeted marks (37.2%).

Table 1: Candidates' performance on each of the items targeted at the final grade they achieved on the three-tier module three intermediate paper

Component: 33003/I			
Item	**Max. mark**	**Mean**	**% of max.**
Grade B			
A8	3	2.70	90.0
A9a	1	0.73	73.0
A9b	2	0.79	39.5
B18	3	1.28	42.6
B19a	1	0.57	57.0
B19b	3	0.95	31.7
Total	**13**	**7.02**	**54.0**
Grade C			
A6b	3	1.60	53.3
A6c	2	0.57	28.5
A7a	1	0.46	46.0
A7b	1	0.35	35.0
B16	3	2.24	74.7
B17	3	0.74	24.7
Total	**13**	**5.96**	**45.9**
Grade D			
A5a	2	1.43	71.5
A5b	4	1.60	40.0
A6a	2	1.16	58.0
B13b	1	0.62	62.0
B14a	2	1.56	78.0
B14b	2	1.08	54.0
B15a	1	0.86	86.0
B15b	1	0.06	6.0
B15c	1	0.39	39.0
Total	**16**	**8.75**	**54.7**
Grade E			
A1	2	0.97	48.5
A2	4	1.56	39.0
A3	2	1.66	83.0
A4	3	1.42	47.3
B10a	2	0.03	1.5
B10b	1	0.45	45.0
B11	2	0.53	26.5
B12a	1	0.33	33.0
B12b	1	0.28	28.0
B12c	2	0.96	48.0
B12d	1	0.33	33.0
B13a	1	0.30	30.0
Total	**22**	**8.81**	**40.0**

Table 2: Candidates' performance on each of the items targeted at the final grade they achieved on the two-tier module three foundation paper

Component: 43003/F			
Item	**Max. mark**	**Mean**	**% of max.**
Grade C			
A9	3	1.57	52.3
B16	3	2.01	67.0
B17	3	0.57	19.0
Total	**9**	**4.10**	**45.6**
Grade D			
A8a	2	1.32	66.0
A8b	4	1.29	32.3
B15a	2	1.58	79.0
B15b	2	1.04	52.0
Total	**10**	**5.20**	**52.0**
Grade E			
A5	2	0.56	28.0
A6	4	1.40	35.0
A7a	1	0.39	39.0
A7b	1	0.19	19.0
B13a	2	0.01	0.5
B13b	1	0.38	38.0
B14	2	0.38	19.0
Total	**13**	**3.31**	**25.5**
Grade F			
A2e	1	0.72	72.0
A2f	1	0.22	22.0
A4a	2	1.64	82.0
A4b	1	0.02	2.0
B11c	3	0.83	27.7
B12b	3	2.37	79.0
B12c	2	0.60	30.0
Total	**13**	**6.40**	**49.2**
Grade G			
A1a	1	0.94	94.0
A1b	3	2.03	67.7
A2a	1	0.73	73.0
A2b	1	0.67	67.0
A2c	1	0.28	28.0
A2d	1	0.71	71.0
B10a	1	0.81	81.0
B10b	1	0.89	89.0
B10c	1	0.61	61.0
B10D	1	0.68	68.0
B10e	1	0.87	87.0
B11a	1	0.37	37.0
B11b	2	0.34	34.0
B12a	1	0.95	95.0
Total	**17**	**11.21**	**65.9**

Table 3: Candidates' performance on each of the items targeted at the final grade they achieved on the two-tier module three higher paper

Component: 43003/H			
Item	**Max. mark**	**Mean**	**% of max.**
Grade A*			
A7	4	3.34	83.5
B16a	3	2.21	73.7
B16b	3	1.31	43.7
Total	**10**	**6.86**	**68.6**
Grade A			
A6a	3	1.26	42.0
A6b	3	0.84	28.0
B15a	1	0.36	36.0
B15b	2	0.90	45.0
Total	**9**	**3.35**	**37.2**
Grade B			
A4	3	2.61	87.0
A5a	1	0.65	65.0
A5b	2	0.54	27.0
B13	3	0.97	32.0
B14a	1	0.57	57.0
B14b	3	0.78	26.0
Total	**13**	**6.11**	**47.0**
Grade C			
A2b	3	1.74	58.0
A2c	2	0.92	46.0
A3a	1	0.38	38.0
A3c	1	0.30	30.0
B11	3	2.31	77.0
B12	3	0.87	29.0
Total	**13**	**6.5**	**50.0**
Grade D			
A1a	2	1.41	70.5
A1b	4	1.85	46.3
A2a	2	1.19	59.5
A3b	1	0.27	27.0
B8a	2	1.68	84.0
B8b	2	1.28	64.0
B9	3	1.63	54.3
B10a	1	0.93	93.0
B10b	1	0.12	12.0
B10c	1	0.52	52.0
Total	**19**	**10.89**	**57.3**

Conclusion

Overall, a positive linear relationship was evident between candidates' scores on individual items (and across targeted items) for three Module 3 papers in March 2007. However, for each of the papers investigated candidates' scores on items targeted at their final module grade varied substantially. At best candidates achieved 68.6 *per cent* of targeted marks (Table 3), and at worst 25.5 *per cent* (Table 1). This has implications for how awarders use information regarding targeted questions when making judgemental decisions. When deciding upon the cut-scores for papers it would have been inappropriate to have based judgements upon whether candidates demonstrated the ability to answer items targeted at the grade in question.

Lucy Billington & Ashleigh Seldon, 2 May 2007
W:\Lucy\Misc\Ash - Targeting grades\Targeting questions to grades.doc

Appendices

Appendix 1: Grade Descriptions[2]

Mathematics B (4302)

The following grade descriptors indicate the level of attainment characteristic of the given grade at GCSE. They give a general indication of the required learning outcomes at each specific grade. The descriptors should be interpreted in relation to the content outlined in the specification; they are not designed to define that content.

The grade awarded will depend in practice upon the extent to which the candidate has met the assessment objectives (as in Section 6) overall. Shortcomings in some aspects of the examination may be balanced by better performances in others.

Grade A

Candidates give reasons for the choices they make when investigating within mathematics itself or when using mathematics to analyse tasks: these reasons explain why particular lines of enquiry or procedures are followed and others rejected. Candidates apply the mathematics they know in familiar and unfamiliar contexts. Candidates use mathematical language and symbols effectively in presenting a convincing reasoned argument. Their reports include mathematical justifications, explaining their solutions to problems involving a number of features or variables.

Candidates manipulate simple surds. They determine the bounds of Intervals. Candidates understand and use direct and inverse proportion. They manipulate algebraic formulae, equations and expressions, finding common factors and multiplying two linear expressions. In simplifying algebraic expressions, they use rules of indices for negative and fractional values. They solve problems using intersections and gradients of graphs.

[2] Grade descriptors can be accessed at **http://www.aqa.org.uk/qual/pdf/AQA-4306-W-SP-09.PDF**

Candidates sketch the graphs of sine, cosine and tangent functions for any angle and generate and interpret graphs based on these functions. Candidates use sine, cosine and tangent of angles of any size, and Pythagoras Theorem when solving problems in two and three dimensions. They use the conditions for congruent triangles in formal geometric proofs. They calculate lengths of circular arcs and areas of sectors, and calculate the surface area of cylinders and volumes of cones and spheres. They understand and use the effect of enlargement on areas and volumes of shapes and solids.

Candidates interpret and construct histograms. They understand how different methods of sampling and different sample sizes may affect the reliability of conclusions drawn; they select and justify a sample and method, to investigate a population. They recognise when and how to work with probabilities associated with independent and mutually exclusive events.

Grade C

Starting from problems or contexts that have been presented to them, candidates refine or extend the mathematics used to generate fuller solutions. They give a reason for their choice of mathematical presentation, explaining features they have selected. Candidates justify their generalisations, arguments or solutions, showing some insight into the mathematical structure of the problem. They appreciate the difference between mathematical explanation and experimental evidence.

In making estimates candidates use appropriate techniques and multiply and divide mentally. They solve numerical problems involving multiplication and division with numbers of any size using a calculator efficiently and appropriately. They understand the effects of multiplying and dividing by numbers between 0 and 1. They use ratios in appropriate situations. They understand and use proportional changes. Candidates find and describe in symbols the next term or the nth term of a sequence, where the rule is linear.

Candidates calculate one quantity as a percentage of another. They multiply two expressions of the form $(x + n)$; they simplify the corresponding quadratic expressions. They solve simple polynomial equations by trial and improvement and represent inequalities using a number line. They formulate and solve linear equations with whole number coefficients. They manipulate simple algebraic formulae, equations and expressions. Candidates draw and use graphs of quadratic functions.

Candidates solve problems using angle and symmetry properties of polygons and properties of intersecting and parallel lines. They understand and apply Pythagoras theorem when solving problems in two-dimensions. Candidates solve problems involving areas and circumferences of circles. They calculate lengths, areas and volumes in plane shapes and right prisms. Candidates enlarge shapes by a positive whole number or fractional scale factor. They appreciate the imprecision of measurement and recognise that a measurement given to the nearest whole number may be inaccurate by up to one half in either direction. They understand and use compound measures such as speed. Candidates use mathematical instruments to carry out accurate constructions of loci.

Candidates construct and interpret frequency diagrams with grouped data. They specify hypotheses and test them. They determine the modal class and estimate the mean, median and range of a set of grouped data, selecting the statistic most appropriate to their line of enquiry. They use measures of average and range with associated frequency polygons, as appropriate, to compare distributions and make inferences. Candidates understand relative frequency as an estimate of probability and use this to compare outcomes of experiments.

Grade F

In order to carry through tasks and solve mathematical problems, candidates identify and obtain necessary information; they check their results, considering whether these are sensible. Candidates show understanding of situations by describing them mathematically using symbols, words and diagrams. They draw simple conclusions of their own and give an explanation of their reasoning.

Candidates use their understanding of place value to multiply and divide whole numbers and decimals by 10, 100 and 1000. They order, add and subtract negative numbers in context. They use all four operations with decimals to two places. They reduce a fraction to its simplest form by canceling common factors and solve simple problems involving ratio and direct proportion. They calculate fractional or percentage parts of quantities and measurements, using a calculator where necessary. Candidates understand and use an appropriate non-calculator method for solving problems involving multiplying and dividing any three-digit by any two-digit number.

In solving problems with or without a calculator, candidates check the reasonableness of their results by reference to their knowledge of the context or to the size of the numbers, by applying inverse operations or by estimating using approximations. Candidates explore and describe number patterns and relationships including multiple, factor and square. They construct, express in symbolic form, and use simple formulae involving one or two operations.

When constructing models and when drawing, or using shapes, candidates measure and draw angles as accurately as practicable, and use language associated with angle. They know the angle sum of a triangle and that of angles at a point. They identify all the symmetries of 2-D shapes. They know the rough metric equivalents of imperial units still in daily use and convert one metric unit to another. They make sensible estimates of a range of measures in relation to everyday situations. Candidates calculate areas of rectangles. Candidates use coordinates in all four quadrants to locate and specify points.

Candidates understand and use the mean of discrete data. They compare two simple distributions using the range and one of the mode, median or mean. They interpret graphs and diagrams, including pie charts, and draw conclusions. They understand and use the probability scale from 0 to 1. Candidates make and justify estimates of probability by selecting and using a method based on equally likely outcomes or on experimental evidence as appropriate. They understand that different outcomes may result from repeating an experiment.

Ashleigh Seldon
May 2007

Appendix 2: Candidates' performance across all questions targeted at each grade by final awarded grade

Three tier specification

3303/I

Figure 1: Candidates' mean mark across grade B questions by final module grade

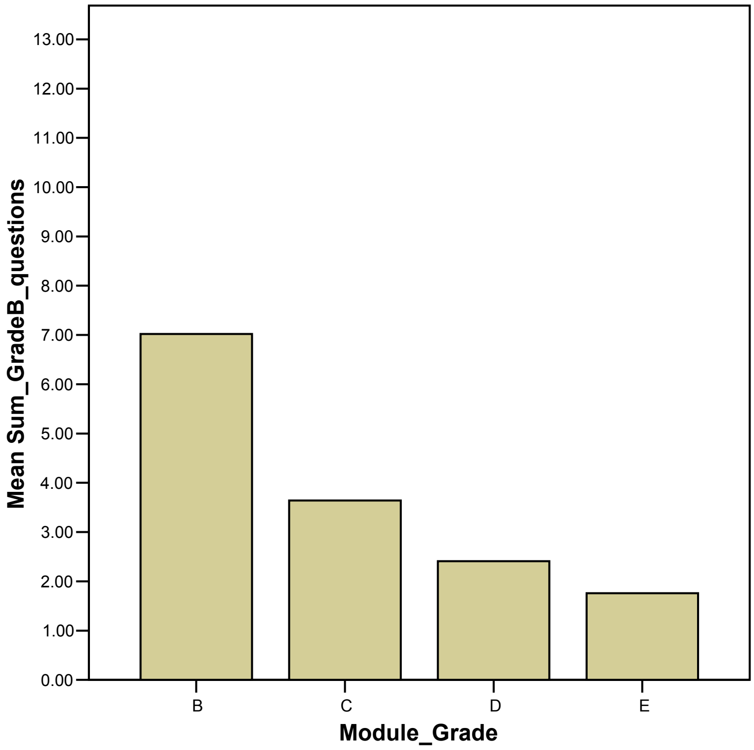


Figure 2: Candidates' mean mark across grade C questions by final module grade

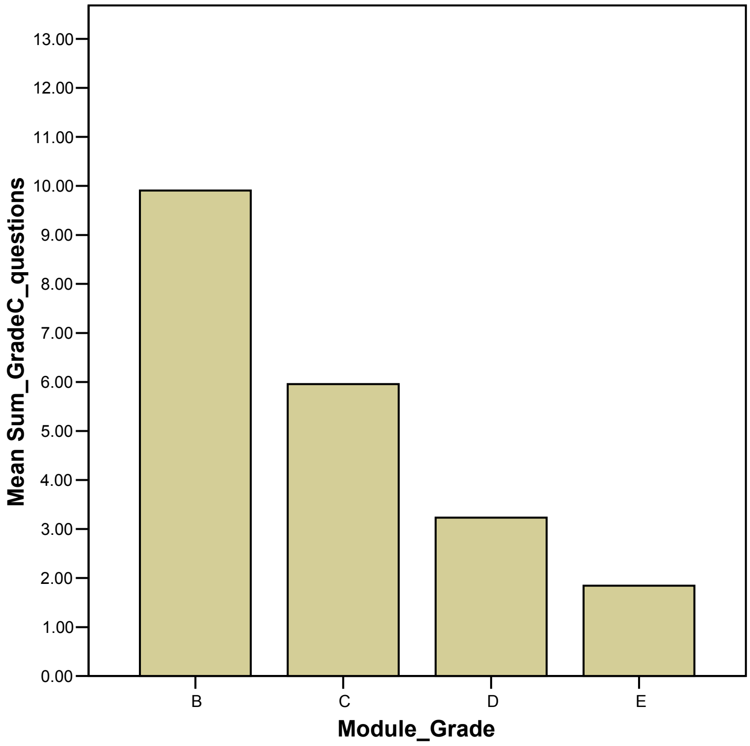


Figure 3: Candidates' mean mark across grade D questions by final module grade

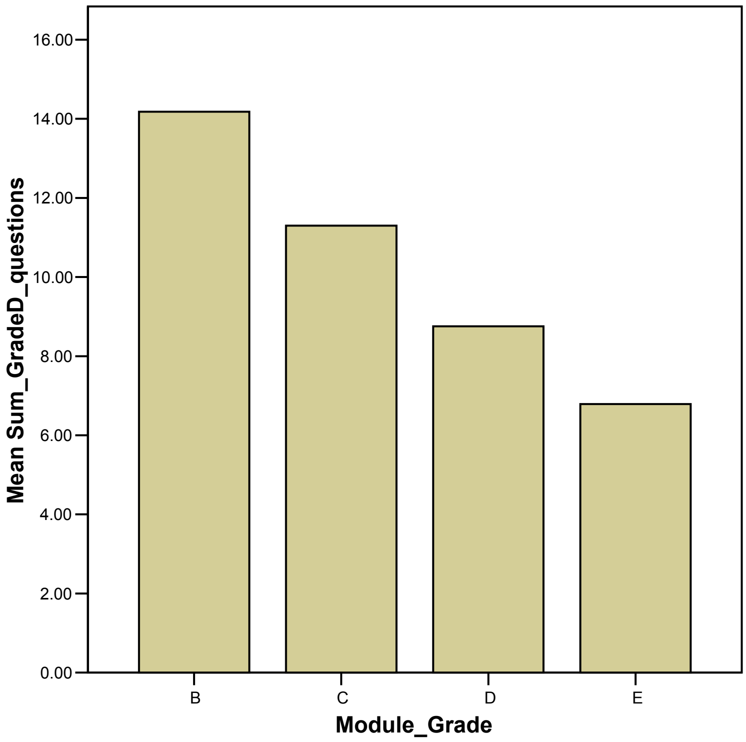
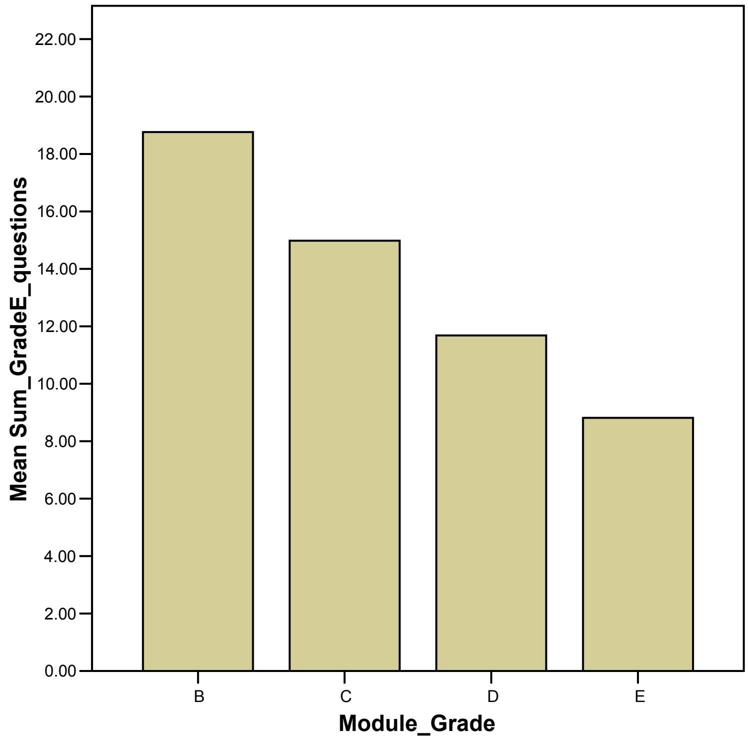


Figure 4: Candidates' mean mark across grade E questions by final module grade



43003/F

Figure 5: Candidates' mean mark across grade C questions by final module grade

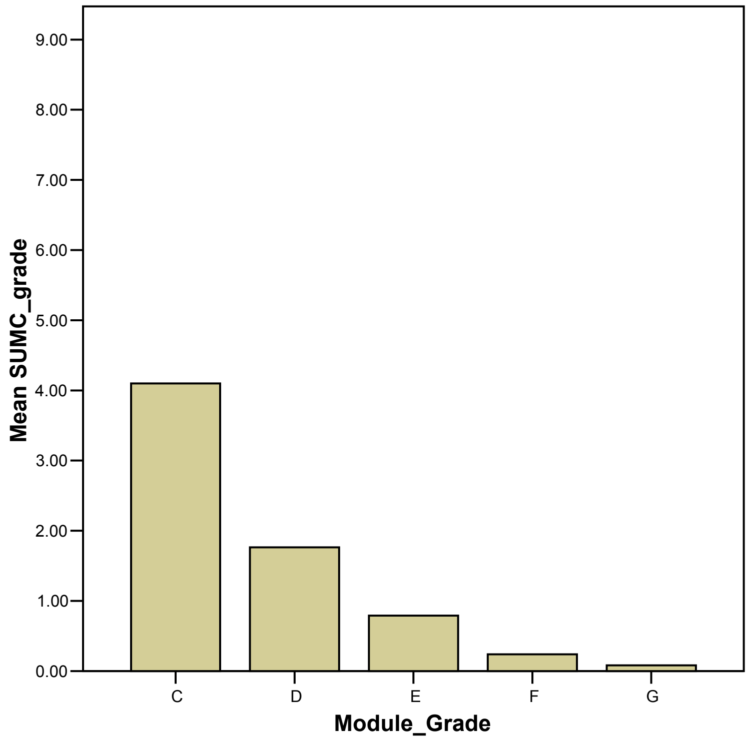


Figure 6: Candidates' mean mark across grade D questions by final module grade

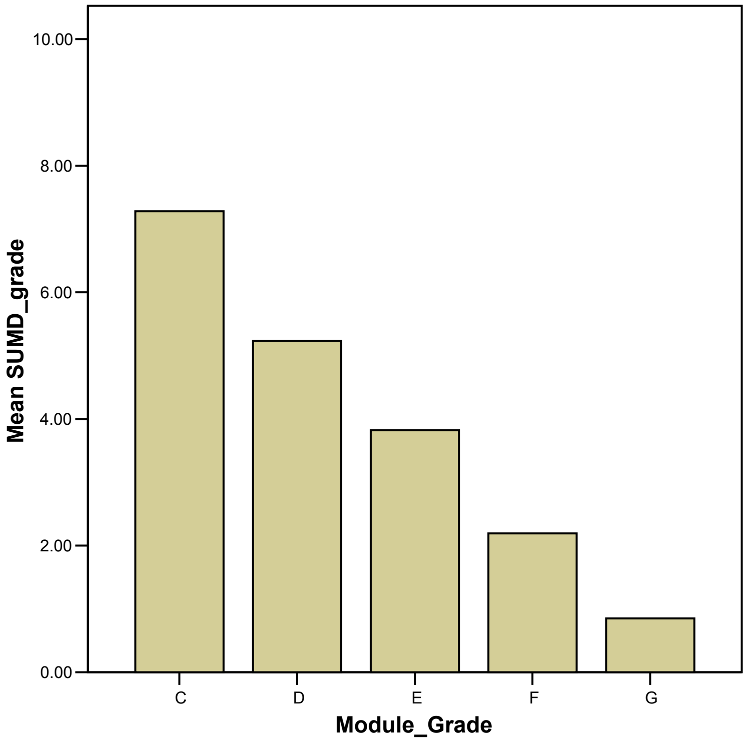


Figure 7: Candidates' mean mark across grade E questions by final module grade

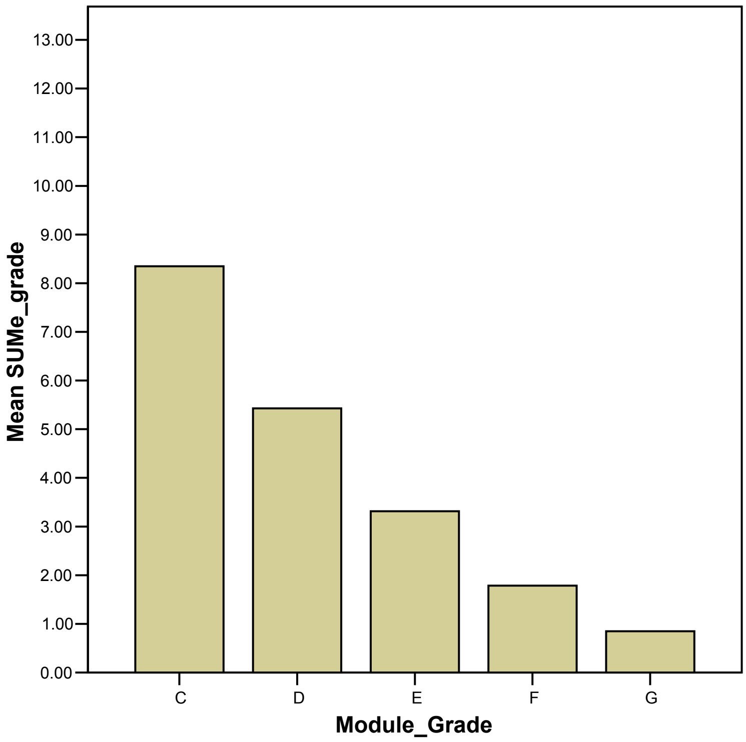


Figure 8: Candidates' mean mark across grade F questions by final module grade

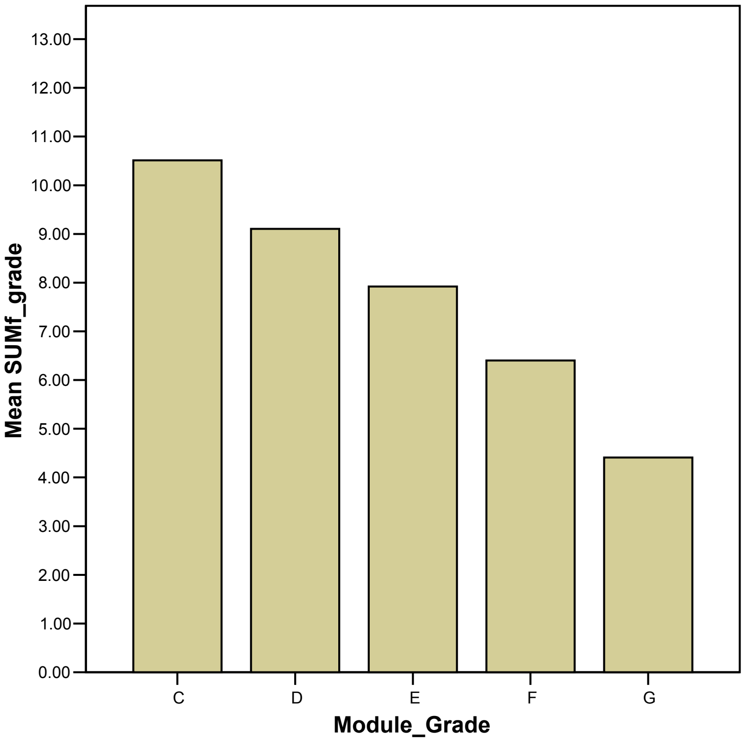
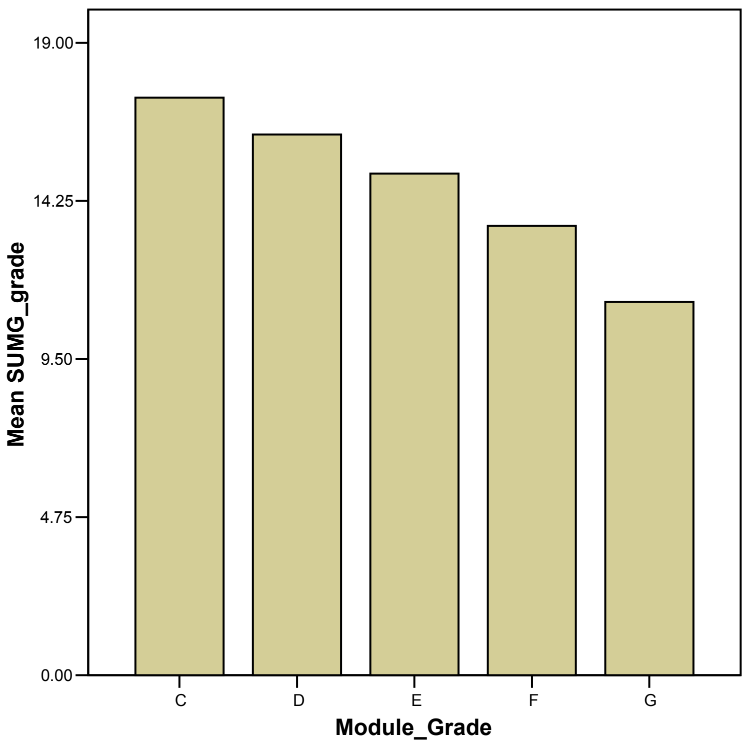


Figure 9: Candidates' mean mark across grade G questions by final module grade



-

43003/H

Figure 10: Candidates' mean mark across grade A* questions by final module grade

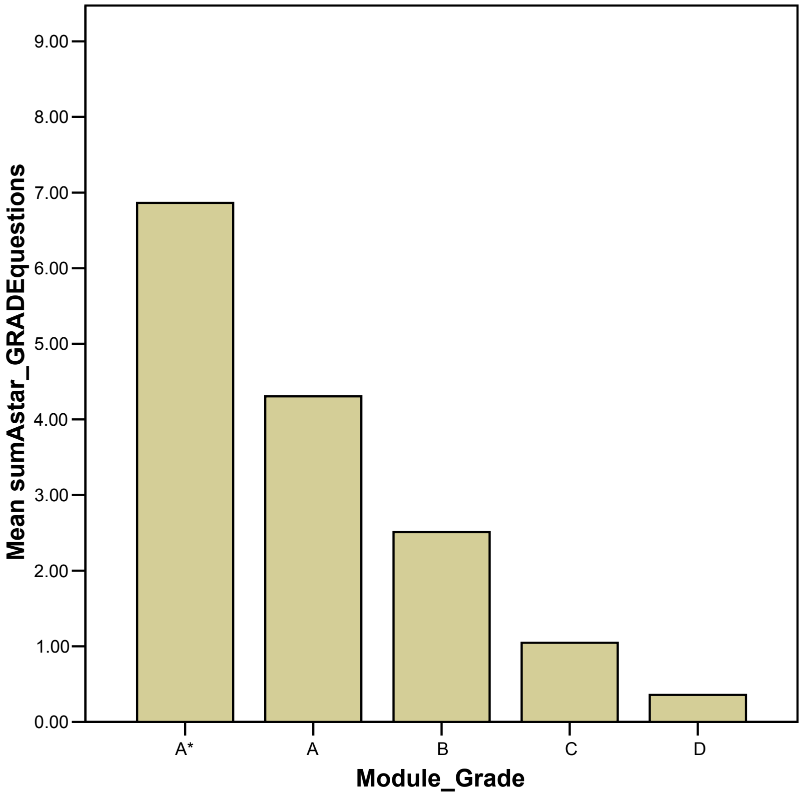


Figure 11: Candidates' mean mark across grade A questions by final module grade

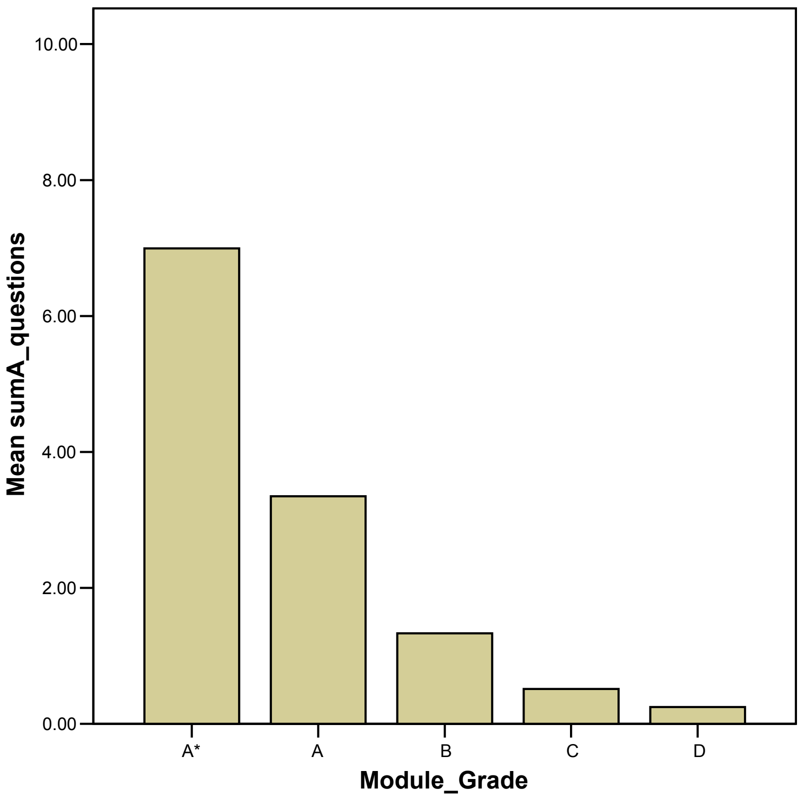


Figure 12: Candidates' mean mark across grade B questions by final module grade

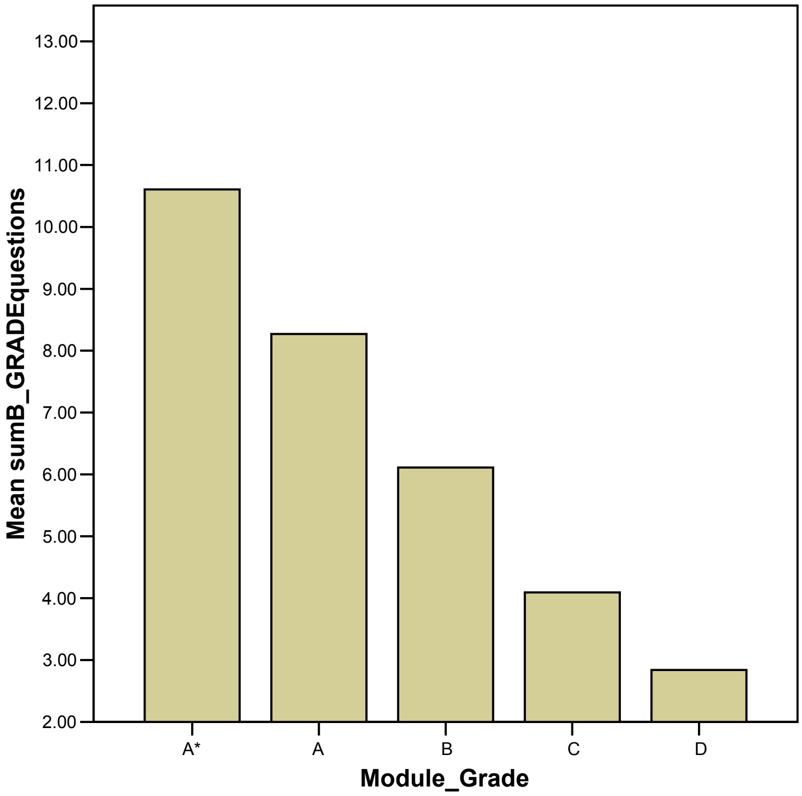


Figure 13: Candidates' mean mark across grade C questions by final module grade

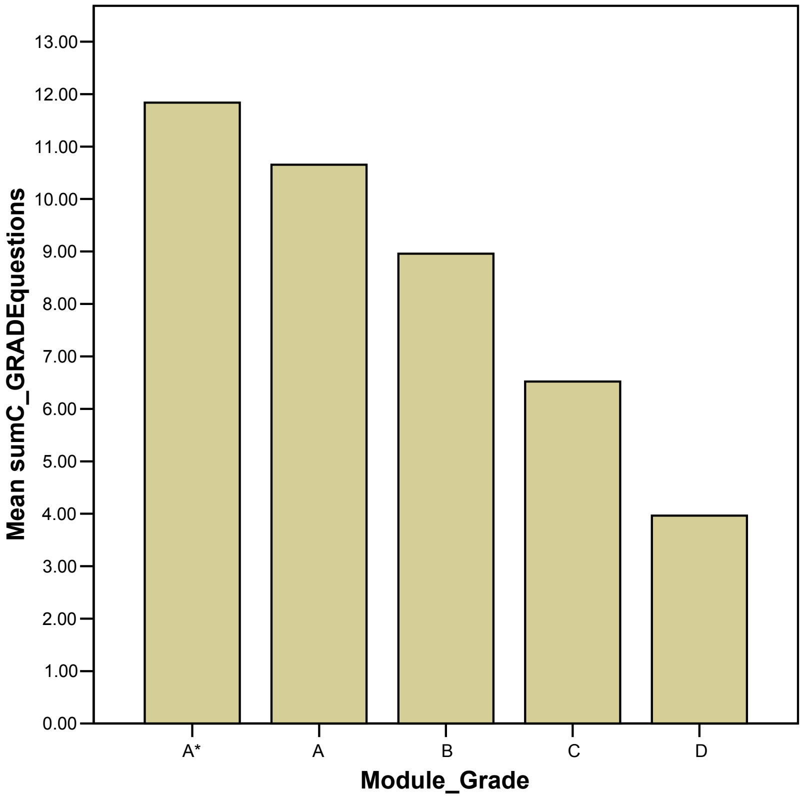


Figure 14: Candidates' mean mark across grade D questions by final module grade