ROUTINE QUESTIONS AND A-LEVEL MATHEMATICS GRADES

John Berry & Wendy Maull — University of Plymouth
Peter Johnson & John Monaghan — University of Leeds

Abstract

This study concerns student performance in A-level examination questions. In particular whether lower attaining students in mathematics examinations generally gain their marks on routine parts of questions? Students' scripts in a recent mathematics examination were examined in an attempt to evaluate this question. The question is an important one because routine questions could be awarded fewer marks if algebraic calculators are allowed in examinations. The results are not conclusive but indicate that a problem of this type does exist.

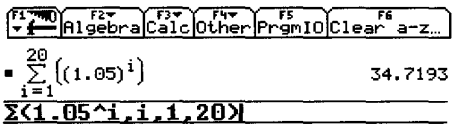
Introduction

Our starting point is: do A-level Mathematics students who attain lower pass grades (D and E) generally obtain these grades by answering 'routine' parts of A-level Mathematics questions? Routine questions may be viewed as those for which students may be expected to execute a rehearsed procedure consisting of a limited number of steps. Problems in characterising routine questions are considered later. The next three paragraphs explain the rationale for and import of the study.

During the period 1994-1996 the then Schools Curriculum and Assessment Authority (SCAA) set up a number of working groups investigating possible consequences of student use of a new generation of algebraic calculators on A-level Mathematics papers. One debate was whether such use would accentuate the difference between higher and lower attaining students, e.g. between those attaining grades A & B and those attaining grades D & E. An example should clarify matters.

A typical question on geometric series, for example, may start by a request to evaluate $\sum_{i=1}^{20} 1.05^i$ and then proceed to a question on compound interest, e.g. "If I invest £550 at a rate of 5% per annum, how many years must I wait until I have more than £1 000 in this account?" It should be noted that the new generation of algebraic calculators can perform the first part of this question,

e.g. the TI-92 screendump on the right. The discussion amongst working group members was interesting. Initially everyone assumed, as a generality with exceptions, that students attaining lower A-level grades learnt how to do the first routine part but would have difficulty with the second non-routine part. At the next meeting, however, the discussion continued



with several people saying they were not sure that this really was the case. There appears to be no literature of direct relevance in this area

Now if, as a generality, lower attaining students obtain the majority of their marks on routine questions and if, as seems likely, such questions are allocated a relatively smaller share of the total mark scheme when algebraic calculators are permitted in examinations (see Monaghan (in press) for a discussion of this issue), then these students will find it more difficult to pass these examinations. Mathematics is already considered a difficult subject at senior school level (see Fitz-Gibbon & Vincent (1994, p.23) for UK data) and we, along with many mathematics educators, would be extremely concerned if mathematics examinations became more difficult to pass.

Methodology

To address the question we analysed the performance of students with A-level grades A, B, C, D & E in questions which have routine and non-routine parts. An Examination Board provided us with the scripts from a recently marked Pure Mathematics paper. Pure Mathematics was chosen, rather than Mechanics, Statistics or Decision Mathematics paper, since Pure Mathematics is the core for all options and because it is the area most likely to be affected by algebraic calculators (see Monaghan (in press) for a discussion of this). Over 300 scripts from equal numbers of male and female students who obtained scores at the boundaries of the A, B, C, D & E grades were provided.

Each question part was coded as routine or non-routine (further details of the paper are noted below) and the marks for the parts of the questions were adjusted in accordance with our expections of what future marks, where routine questions were allocated a relatively smaller share of the total mark scheme, would be like. Each student's script was then remarked.

The A-level **paper and mark** scheme(s)

The paper used was the first of six papers. It was one of three papers that all students following a popular modular A-level scheme had to take. There were four questions each worth 15 marks. Questions 1,2 and 4 had four parts. Question 3 had five parts. We reproduce question 2 below as an example of the kind of question asked. The original marks are given in square brackets.

The gradient of a curve is given by $\frac{dy}{dx} = 3x^2 - 8x + 5$. The curve passes through the point (0, 3).	
(i)	Find the equation of the curve. [4]
(ii)	Find the coordinates of the two stationary points on the curve. State, with a reason, the nature of each stationary point. [6]

\2,

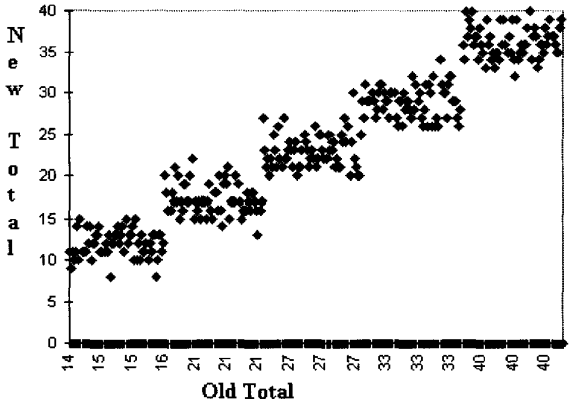
(iii)	State the range of values of k for which the curve has three distinct intersections with the line $y=k$.	[2]
(iv)	State the range of values of x for which the curve has a negative gradient. Find the x-coordinate of the point within this range where the curve is steepest.	[3]

The other questions concerned: Ql, trigonometry in context; Q3, coordinate geometry (lines, circles and ellipses); Q4, integration in context (comparison of exact and numeric methods). We classified each question part as routine (R) or non-routine (N) and obtained: Ql (R, N, N, N); Q2 (R, R, N, N); Q3 (R, R, N, N, N); Q4 (R, R, R, N). The division of marks for routine and non-routine parts was 30 marks each. Various alternative mark schemes were developed, all adjusting the mark ratio so that routine questions scored fewer marks. The agreed version left each question with 15 marks, left the mark allocation of Q 1 unchanged but adjusted the others so that routine parts totalled 23 marks and non-routine parts totalled 37 marks. The parts of Q2, for example, were allocated 3, 4, 4 and 4 marks respectively. The grade boundaries for this paper were: A, 40; B, 33; C, 27; D, 21; and E, 15. A, B and E grade boundaries are, by convention, determined by examiners' judgement while C and D boundaries were fixed at equal intervals between B and E grade boundaries.

Results

311 scripts (63, 63, 62, 62 and 61 at grades A, B, C, D and E respectively) were re-marked to the new mark scheme. The use of statistics in this study must be carefully examined for much of the data is far from independent (consider, for example, the relationship between the total on the original mark scheme and the total on the revised mark scheme). The statistics which follow are intended to give the reader a feel for the general patterns in the data. Three aspects are examined: the overall scores; the proportion of marks obtained in routine and non-routine parts of questions; and factor analytic results suggesting that students follow through whole questions.

The scatter diagram on the right illustrates the general pattern. The ranges overlap but the ranges from the original grades retain their heirarchical structure. The new totals are generally lower than the original totals. In fact of the 311 student scripts examined 297 obtained lower scores from the new mark scheme, **11** scores remained the same (3, 3, 2 and 3 from A, C, D and E grade



students respectively) and three obtained higher marks (a B and a D grade student obtaining one more mark and a C grade student gaining three more marks). This indicates that increased emphasis on non-routine questions leads to a general lowering of the overall marks obtained.

The table on the right displays the mean marks obtained by the groups of students at each grade level for:

- Old routine (O-R) parts (out of 30)
- Old non-routine (O-N) parts(out of30)
- New (revised) routine (N-R) parts (out of23)
- New non-routine (N-N) parts (out of37)

	O-R	O-N	N-R	N-N
A	26.3	13.7	19.9	16.4
B	23.2	9.3	17.4	11.2
C	19.9	7.1	14.8	8.1
D	15.8	5.2	11.1	6.2
E	11.8	3.2	8.3	3.6

Note that each column decreases with decreasing grades (hardly surprising) and that mean marks obtained for routine parts of questions are consistently greater than mean marks obtained for nonroutine parts of questions even though there were more marks for non-routine parts of questions in the revised mark scheme. This may be interpreted as evidence that students attaining at all pass grades gain more marks on routine parts of questions. An examination of the ratios O-R:O-N and N-R:N-N is interesting. For grades A-E we get, respectively: 1.9,2.5, 2.8, 3.0, 3.7 and 1.2, 1.6, 1.8, 1.8, 2.3. The decrease in the second list, relative to the first list, clearly mirrors the higher weighting given to non-routine parts of questions in the revised mark scheme. The increase in both lists, however, may be interpreted as evidence that higher (respectively lower) attaining students gain proportionally (to their overall mark) more (resp. less) marks on non-routine parts of questions.

Principal component analysis of both the old and the new scores yielded 6 factors with eigenvalues greater than one. In both old and new scores the question parts which loaded significantly on the factors were as follows: all parts of Q2; all parts of Q4; Q1 parts i, ii, and iii; Q3 parts i, ii, iii, and iv; Q3 parts ii, iv and v; Q1iv, Question Ii (negatively) and Q2i. The first five factors suggest an interpretation that the correlations of the scores of the parts within a question dominate the analysis, i.e. if you do well on one part ofQ2, you tend to do well on all of it.

Discussion

So, do lower attining students gain a substantial proportion of their marks on routine parts of questions, compared to higher attaining students? The results are not conclusive but they are not without interest. Before considering them we address their surface validity. The results arise from at least two semi-arbitrary decisions: the categorization of parts of questions as routine or not (and the

dichotomy implicit in this categorization); the weightings gIven III the revised mark scheme. Against these criticisms it should be noted that both decisions were made after considerable debate by a group of people with considerable experience of the type of examination paper in question.

The overall lower scores obtained from the revised mark scheme and the distribution of mean marks over routine and non-routine question in both mark schemes clearly inidcate that all students score substantially more marks on what we have designated routine parts of questions. The increasing ratios of routine to non-routine mean marks over grades in both mark schemes, however, does provide evidence that lower attaining students do obtain proportionally more of their marks on routine parts of questions. Looking at the scatter diagram, however, it would not appear that this would make a substantial difference to the overall grades (if they were still determined by the same judgement/equal interval rubric) given the small 'new total' overlap over old grades.

The principal component analysis results alerted us to a possibility we had not, but in retrospect should have, anticipated: that many students at all levels of attainment exhibit a propensity to follow a question through. This may have many bi-causal connections with other influences, including a familiarity with a specific content area - a dialectic may exist between attainment and familiarity with a range of content areas. Again we must view these results with caution, due to regularity conditions implicit in the analysis and a lack of prior hypotheses, but five out of six located factors indicating such a propensity in students certainly deserves consideration.

What these results suggest is the need for further investigations. Two such investigations are a refinement of what is meant by the term 'routine question' and a more realistic interpretation of curriculum development vis a vis assessment development. We turn to a considerations of these now but note that space does not allow as full a discussion as we would like.

This study implies that routineness is located in a question rather than being a socio-psychological construct ofthe relation between an individual/group and a question. The latter appears more valid. Routineness as located in a question was, however, a natural approach for the study given its origins in the SCAA working group. This study should be viewed as exploring one avenue of routineness.

There are few references to routine questions in the literature. Seldon et al. (1989) address a similar level/type of mathematics to this study. They distinguish between problems and exercises and view a problem as having two componenets: task and solver(s). They view 'cognitively non-trivial'

problems as those where "the solver does not begin knowing a method of solution" and note that traditional calculus courses contain few cognitively nontrivial problems. Boaler (1997) explored a range of issues from two schools with contrasting ethoses and teaching methods. In her analyses of GCSE students' performance on conceptual/procedural questions she defines procedural questions as "those questions that could be answered by a simplistic rehearsal of a rule, method or formula." whereas conceptual questions require "the use of some thought and rules or methods committed to memory in lessons would not be of great help". She claims that conceptual questions are more difficult, for students, but shows that similar overall results in examination performance may be obtained in different ways (ratios of success in the two types of questions). This raises the obvious question of the relationship between routineness and teaching methods. Nagy et al. (1991) examine the relationship between test content and instructional content in High School calculus. Their analysis of assessment, ranging from skills to situational problems, showed wide variation in teachers' emphasis, especially at the skill level. This calls attention to the importance of further studies on instruction. The Question Difficulty Project (Fisher-Hock et aI., 1997) examined mathematics and other subject GCSE examinations focusing on a model of question answering based on reading, application and communication. The findings noted the difficulty of both social and mathematical language, to the presentation of answers and the concomitant recording of steps. Trials revealed 22 sources of difficulty, from command words to irrelevant information and alert us to the sheer number of factors impinging on what might make a question routine or not.

In the medium and long term, when algebraic calculators become commonplace in classrooms, it is possible that the curriculum will develop to incorporate the potential offered by these calculators. Ideally examination questions, if not the form of examinations, will change in line with curriculum developments. In this projected setting the current study is partially misplaced as an attempt to second guess the future without taking account of curriculum development.

References

Boaler, J. (1997) *Experiencing School Mathematics.* Buckingham: Open University Press.

Fisher-Hoch, H. & Hughes, S. (1997) What Makes Mathematics Exam Questions Difficult? http://gps.1eeds.ac.uk/cgi-bin/ucslbrs_engine_vl

Monaghan, J. (in press) Some issues surrounding the use of algebraic calculators in traditional examinations. *International Journal for Mathematics Education in Science and Technology.*

Nagy, P., Traub, R., MacRury, K. & Roslyn, K. (1991) High School Calculus: Comparing the Content of Assignments and Tests, *Journalfor Research in Math Education,* 22(1), 69-75.

Selden, J., Mason, A. & Selden, A. (1989) Can Average Calculus Students Solve Nonroutine Problems? *Journal of Mathematical Behaviour,* **8,** 45-50.