

Contrasts in mathematical challenges in A-level Mathematics and Further Mathematics, and undergraduate mathematics examinations

ELLIE DARLINGTON*

Department of Education, University of Oxford, 15 Norham Gardens, Oxford, OX2 6PY, UK

**Email: ellie.darlington@education.ox.ac.uk*

[Submitted January 2014; accepted June 2014]

This article describes part of a study which investigated the role of questions in students' approaches to learning mathematics at the secondary–tertiary interface, focussing on the enculturation of students at the University of Oxford. Use of the Mathematical Assessment Task Hierarchy taxonomy revealed A-level Mathematics and Further Mathematics questions in England and Wales to focus on requiring students to demonstrate a routine use of procedures, whereas those in first-year undergraduate mathematics primarily required students to be able to draw implications, conclusions and to justify their answers and make conjectures. While these findings confirm the need for reforms of examinations at this level, questions must also be raised over the nature of undergraduate mathematics assessment, since it is sometimes possible for students to be awarded a first-class examination mark solely through stating known facts or reproducing something verbatim from lecture notes.

I. Introduction

Schooling in England and Wales is compulsory only till the age of 16 years, after which students may choose to continue in education (be it academic or vocational) or to go into work. The most-commonly used academic qualification at this age is the A-level, for which the majority of students study three subjects of their choice and are awarded separate grades in each. An A-level is typically completed in two years, although students are currently¹ able to stop after one year and leave with an Advanced Subsidiary (AS) level. An AS-level is composed of three equally weighted modules, and an A-level consists of these plus three additional equally weighted modules, with AS results contributing towards the overall A-level result. In A-level Mathematics, students have the option of studying a variety of modules which include pure mathematics, mechanics, statistics and discrete mathematics and there are six possible routes through an A-level Mathematics qualification based on students' choices of

¹ From September 2015, AS-levels will become standalone qualifications and will cease contributing towards an A-level grade.

modules (Qualifications & Curriculum Authority, 2007). These optional modules form one-third of the A-level Mathematics curriculum. There are currently two examination sittings each year for each module² (one in January and one in the summer) and there are four major examination boards offering them (AQA, OCR, Edexcel, WJEC).

English A-level examinations are often accused of becoming easier and allegations of ‘grade inflation’ abound which, in the case of mathematics, is compounded by comments made by observers, educational researchers and universities that A-level Mathematics no longer acts as adequate preparation for undergraduate mathematics courses. While much comment has been made regarding the content of A-level Mathematics, and the styles of questions compared over time (e.g. Bassett *et al.*, 2009), legitimate use of a taxonomy has not been made to compare the demands of questions at the secondary and tertiary level. The Mathematical Assessment Task Hierarchy (MATH) taxonomy (Smith *et al.*, 1996) was devised as a means of classifying advanced mathematics questions into one of three groups (A, B and C) based on the skills that the question required of the student. While research by Etchells & Monaghan (1994) found that the majority of A-level Mathematics questions were of a ‘Group A’ type (see Section 2.3.1), the taxonomy has not been applied to undergraduate level examination questions or, indeed, A-level Further Mathematics questions.

2. Analysing questions in assessment

2.1 Routine and non-routine questions

At different levels of schooling, students are asked questions of varying and increasing levels of complexity and difficulty. In this sense, one could distinguish between what Pólya (1945) defined as ‘non-routine’ and ‘routine’ questions:

a problem is a ‘routine problem’ if it can be solved either by substituting special data into a formally solved general problem, or by following step by step, without any trace of originality, some well-worn conspicuous example.

(p. 171)

Routine questions have been found by Berry *et al.* (1999) to form the basis for the majority of marks awarded in A-level Mathematics examinations. Furthermore, when they redistributed the marks on pure mathematics papers to give more reward for solutions to non-routine problems, they found that 297/311 of the scripts analysed would have had reduced marks.

For example, a routine A-level Mathematics question might have been:

Express $\frac{1}{x^2(x-1)}$ in the form $\frac{A}{x} + \frac{B}{x^2} + \frac{C}{x-1}$ where A , B and C are constants.

(UCLES³ Syllabus C Paper 1 June 1994, question 9)

Alternatively, one might consider the following to be a routine undergraduate mathematics question:

Consider the 2×2 real square matrix $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$

² This will cease to be the case from 2014, when the January examination sessions will be discontinued.

³ The University of Cambridge Local Examinations Syndicate has since merged and evolved to become the OCR examination board.

Show that A has at least one real eigenvalue; and that if $a \neq d$ or $b \neq 0$, then it has two distinct real eigenvalues.

Deduce that A is diagonalisable.

(University of Oxford Algebra I 2010, question 3)

As such questions do not require original thought or application of knowledge in new situations, it is perhaps understandable to find that such questions are those which lower-attaining students are better at (Berry *et al.*, 1999). This is because they do not require the respondent to have a conceptual understanding of the material employed, as it is possible for them to practise, drill-style, similar questions in advance of formal assessment.

Conversely, Elia *et al.* (2009) describe a non-routine question as requiring ‘creative thinking and the application of a certain heuristic strategy to understand the problem situation and find a way to solve the problem’ (pp. 606–607). In order to solve such a problem, a thorough understanding of all of the component concepts would be necessary.

However, it is important to caution that a question may be considered routine by one person but not by another. Indeed, ‘routineness has to do with what the solver is used to’ (Hughes *et al.*, 2006, p. 91). Furthermore, one question may be considered routine in one particular instance, and yet non-routine in another. For example, Hughes *et al.* (2006) found that the ‘routineness’ of a question changed over time in that an absence from study changed the nature of the question for a student. For example, a question asked in a weekly assignment and, later, repeated in an examination may be tackled differently by the student each time, i.e.

A student who succeeds in proving an unseen theorem is demonstrating an ability to apply knowledge to new situations, but may only be demonstrating factual recall when proving it for a second time.

(Smith *et al.*, 1996, p. 68)

Rote learning is not of benefit to the student in answering further, non-routine questions, as this knowledge cannot be adapted or manipulated for use in problem-solving contexts (Novak, 1978). Should students resort to memorizing for reproduction, they will lack the conceptual understanding necessary to use such a definition or theorem in an attempt to answer more involved, advanced questions which require their application. These questions would be considered non-routine as they cannot be answered through stating what has already been learnt or by carrying out a series of tried-and-tested steps in search of the answer.

This may be further compounded by new undergraduates’ perceptions of mathematics. For example, Kathryn Crawford and her colleagues found that many mathematics students start their first year viewing the subject as a rote-learning task (Crawford *et al.*, 1994; 1998a,b), a phenomenon which has been found to persist as they progress through their degree (Anderson *et al.*, 1998; Maguire *et al.*, 2001).

2.2 Taxonomies

In response to criticisms of current assessment formats, specific means by which questions may be classified have been developed. Such classifications are referred to as ‘taxonomies’, and may be used to help create questions which assess particular skills and concepts according to the guidelines set forth by governing bodies or suggested and encouraged by research.

Kadijević (2002) strongly encourages the operationalization of taxonomies when designing assessment, describing it as ‘a useful framework’, as opposed to ‘a dogmatic recipe’ (p. 101), which can be used to ‘guide and foster an adequate mathematics learning as well as achieve a comprehensive evaluation of its outcome’ (p. 97). Various different taxonomies have been proposed; some designed for general assessment, some for mathematics assessment, and some for tertiary-level mathematics assessment. However, as with most frameworks in education, it is necessary to exercise an amount of caution.

Resnick (1987) cautions against taxonomies which either explicitly state, or suggest, a hierarchy of skills. That is, taxonomies which claim that factual and procedural knowledge is necessary before one can answer a question requiring deeper understanding. Kreitzer & Madaus (1994) describe many taxonomies as cumulative, in that ‘each class of behaviours was presumed to include all the behaviours of the less complex classes’ (p. 66). Resnick (1987) claims that this type of practice and the ‘relative ease of assessing people’s knowledge, as opposed to their thought processes, further feeds this tendency in educational practice’ (pp. 48–49).

In mathematics, where there have been moves towards diversifying types of assessment made available to students—for example through coursework or questions which utilize computers or graphic calculators—Huntley *et al.* (2009) have claimed that general assessment taxonomies are not applicable. In requiring such a different skillset to the humanities and sciences, he claims that general taxonomies ‘are not pertinent to mathematics’ (p. 3) and its idiosyncratic demands and topics. Indeed, Duval (2006) asks ‘Is the way of thinking the same in mathematics as in the other areas of knowledge?’ (p. 105). Hence, questions can be asked regarding whether such general taxonomies of learning and assessment can be reliably applied to mathematics.

This criticism has been made of Bloom’s taxonomy of educational objectives (Bloom *et al.*, 1956), which has been described as being ill-fitting to mathematics⁴ (Ormell, 1974; Romberg *et al.*, 1990; Kilpatrick, 1993). Furthermore, its hierarchical nature has been described as being both poor and inappropriate for the study of question types (Pring, 1971; Anderson & Soxniak, 1994). Criticism has been made of the structure of the SOLO taxonomy (Biggs & Collis, 1982), a taxonomy which is concerned with learning outcomes and the student, as opposed to the mathematical demands of a task. Chan *et al.* (2001) claim it is overly ambiguous and that the divisions between categories are too ‘blurred’. The utility of Galbraith and Haines’ taxonomy (1995) has its limitations to a study such as this, given its description of three types of task (mechanical, interpretive and constructive) appear to resonate and correspond with the MATH taxonomy’s groups (see below), though without the benefits of sub-categorization.

2.3 MATH taxonomy

A modification of Bloom’s (1956) taxonomy for the context of undergraduate mathematics was made by Smith *et al.* (1996), whose focus was on using a taxonomy to classify the skills required to complete a particular mathematical task. They designed the MATH taxonomy to assist the development and construction of advanced mathematics assessments in order to ensure that students are assessed on a range of knowledge and skills. The taxonomy ensures that students have the opportunity to demonstrate their understanding of mathematical concepts at different levels.

⁴ While the Assessment Component Taxonomy (Huntley *et al.*, 2009) appears to be a strong means of classifying advanced mathematics questions, it was designed for use when classifying tasks given by alternative assessment formats such as multiple choice questions or CAS.

The mathematical skills associated with Group C—‘those that we associate with a practising mathematician and problem solver’ (Pountney *et al.*, 2002, p. 15)—are those which, unfortunately, have been found to be most lacking among undergraduate mathematicians.

2.3.1 Categories The categories in the MATH taxonomy, which are designed in order to describe the ‘nature of the activity . . . not the degree of difficulty’ (Smith *et al.*, 1996, p. 68), are given in Table 1. That is, a Group A task might be considered more difficult than a Group C task by a particular student, depending on their perception of difficulty, as well as the particular challenges associated with the task.

One of the reasons why Group C tasks have very little attention paid to them may be the amount of time it would take to introduce such skills and nurture their development in students (Leinbach *et al.*, 2002). Time pressures, deadlines and targets are frequently blamed for tendencies to teach for assessment, which can be to the detriment of students’ learning. This is even more troubling should students go from one environment where their Group C skills are not fostered into one where they are required and are presumed to be well-engrained in their view of mathematics. Indeed, Leinbach *et al.* (2002) argue that ‘all students deserve the opportunity to attempt projects that develop the Group C-level skills’ (pp. 4–5). Such tasks are also likely to require more time to develop and mark, a commodity often in short supply for mathematics educators.

2.3.2 Uses The MATH taxonomy was designed with assessment construction in mind—that is, it was devised in order to provide a framework for ensuring that assessment is varied and tests a variety of skills. Smith *et al.* (1996) claim that assessment tasks ‘show students what we value and how we expect them to direct their time. Good questions are those which help to build concepts, alert students to misconceptions and introduce applications and theoretical ideas’ (p. 66).

Leinbach *et al.* (2002) complement the taxonomy, describing it as ‘a useful tool when determining the role of problems posed to students in the development of their mathematical skills’ (p. 13). This is particularly the case in areas such as algebra, where computer algebra systems (CAS) can be used to perform the more routine procedures and algorithms that otherwise would engulf the majority of students’ study time on particular concepts. The use of such categories as those in the MATH taxonomy, therefore, is helpful in order to ensure that students may progress onto studying and developing skills in solving higher level learning tasks (Wood, 2011).

Examples of questions which might fit into each of the MATH taxonomy’s groups and subcategories as described in Table 1 are given in Table 2.

Ball *et al.* (1998) and Smith *et al.* (1996) applied the MATH taxonomy to a study of existing tertiary-level examination papers and found that the majority that they analysed were ‘heavily biased towards Group A tasks’ (p. 828). Similarly, Etchells and Monaghan (1994; cited by Pountney *et al.*, 2002) found that A-level Mathematics examinations awarded marks mainly for Group A tasks. Indeed, Crawford (1983, 1986) and Crawford *et al.* (1993) found that most new entrants to higher education in Australia were most familiar with Group A tasks, with virtually no experience of Group C tasks.

While the MATH taxonomy was not designed to imply that associated questions were of an increasing level of difficulty, research suggests that students do perceive tasks in Group C to be more difficult than those in Group B, and those in Group B more difficult than Group A (Wood & Smith, 2002). Perceived difficulty appeared to be associated with conceptual difficulty (see also d’Souza & Wood, 2003). Furthermore, familiarity is not necessarily related to perceptions of difficulty, as the students participating in the Wood & Smith (2002) study were able to use their familiarity with certain question types in order to be aware of inherent difficulties associated with them.

TABLE 1. *The MATH taxonomy*

Routine procedures		
Group A	Factual Knowledge and Fact Systems [FKFS]	Bring to mind previously learnt information in the form that it was given.
	Comprehension [COMP]	Decide whether conditions of a simple definition are satisfied, understand the significance of symbols in a formula and substitute into that, recognize examples and counterexamples.
	Routine Use of Procedures [RUOP]	Using a procedure/algorithm in a familiar context. When performed properly, all people solve the problem correctly and in the same way. Students will have been previously exposed to these in drill exercises.
Using existing mathematical knowledge in new ways		
Group B	Information Transfer[†] [IT]	Transferring information from verbal to numerical or vice versa, deciding whether conditions of a conceptual definition are satisfied, recognizing applicability of a generic formula in particular contexts, summarizing in non-technical terms, framing a mathematical argument from a verbal outline, explaining relationships between component parts, explaining processes, resembling given components of an argument in their logical order.
	Application in New Situations [AINS]	Choose and apply appropriate methods/information in new situations.
Application of conceptual knowledge to construct mathematical arguments		
Group C	Justifying and Interpreting [JI]	Proving a theorem in order to justify a result/method/model, finding errors in reasoning, recognizing limitations in a model, ascertaining appropriateness of a model, discussing significance of given examples, recognition of unstated assumptions.
	Implications, Conjectures and Comparisons [ICC]	Given or having found a result/situation, draw implications and make conjectures and the ability to justify/prove these. The student also has the ability to make comparisons, with justification, in various mathematical contexts.
	Evaluation [EVAL]	Judge the value of material for a given purpose based on define criteria—the students may be given the criteria or may have to determine them.

Adapted from Smith *et al.* (1996).

[†]This category covers a type of understanding consistent with Duval's (2006) description of mathematical activity as consisting 'in the transformation of [semiotic] representations' (p. 111).

It appears to be the general consensus in the literature that Group C skills should be a consequence of higher education, as these higher-order skills are synonymous with a "higher education" [which] resides in the higher order states of mind' (Barnett, 1990, p. 202). It has been debated whether assessment which allows students who only have Group A and B skills to pass serves its purpose because 'high marks should be reserved for those who have demonstrated that they have acquired Group C level skills' (Pountney *et al.*, 2002, p. 16). Leinbach *et al.* (2002) suggest that Group B and C skills should be gradually introduced and utilized in such a fashion that 'they become *for that individual* student Group A... tasks because they have developed the insight into the problem-solving process that makes the solution of the problem straightforward' (p. 13).

TABLE 2. *MATH taxonomy applied to secondary and tertiary mathematics questions (sources available upon request)*

Category	A-level standard [†] example	Undergraduate example
Group A		
FKFS	State the cosine rule.	Let $f : X \rightarrow Y$ be a function and $B \subseteq Y$. Define what is meant by $f^{-1}(B)$, the inverse image of B under f .
COMP	Given that $zz^* = 9$, describe the locus of z .	If the function f is continuous on the interval (a, b) but not bounded then $\int_a^b f(x) dx$ does not make sense as a proper Riemann integral. Briefly explain why not.
RUOP	The equation of a curve is $y = f(x)$, where $f(x) = \frac{3x+1}{(x+2)(x-3)}$. Express $f(x)$ in partial fractions.	Use L'Hôpital's Rule to find the limit of the sequence $\left(\frac{\log(n^3 + 1)}{\log(2n^5 - 1)} \right)_{n \in \mathbb{N}}$
Group B		
IT	Describe a sequence of two geometrical transformations that maps the graph of $y = x^2$ onto the graph of $y = 50 - x^2$.	Describe, in about 10 lines, the ideas of the Mean Value Theorem. Imagine that you are describing the theorem to a student about to start university.
AINS	Show, with the aid of a sketch, that $y > \tanh\left(\frac{y}{2}\right)$ for $y > 0$ and deduce that $\operatorname{arcosh} x > \frac{x-1}{\sqrt{x^2-1}}$ for $x > 1$.	Prove that $\max(a, b) = \frac{1}{2}(a+b) + \frac{1}{2} a-b $
Group C		
JI	The matrix A is given by $A = \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix}$. Prove by induction that, for $n \geq 1$, $A^n = \begin{pmatrix} 3^n & \frac{1}{2}(3^n - 1) \\ 0 & 1 \end{pmatrix}$	Explain why the Mean Value Theorem does not apply to the function $f(x) = x+1 $ on the interval $[-3, 1]$.
ICC	The fact that $6 \times 7 = 42$ is a counter-example of which of the following statements? (a) The product of any two integers is odd; (b) If the product of two integers is not a multiple of 4 then the integers are not consecutive; (c) If the product of two integers is a multiple of 4 then the integers are not consecutive; (d) Any even integer can be written as the product of two even integers.	Prove that if f and g are continuous at x_0 , then $\max\{f, g\}$ is continuous at x_0 . What about $\min\{f, g\}$?
EVAL	Given a particular function, discuss the accuracy of the trapezium rule in finding the area under the curve.	The Mean Value Theorem is a powerful tool in calculus. List 3 consequences of the Mean Value Theorem and show how the theorem is used in the proofs of these consequences.

[†]Where examples could be found, these are from either A-level Mathematics, Smith *et al.* (1996), A-level Further Mathematics, the University of Oxford mathematics admissions test (which relies on only AS-level Mathematics material) or Sixth Term Extension Papers (which only rely on A-level Mathematics or Further Mathematics material).

2.3.3 Limitations The main limitations associated with the MATH taxonomy are also common to most other taxonomies. In particular, some tasks may involve the use of more than one type of knowledge or activity—‘even in the higher-level skills there are some mechanical parts’ (Leinbach *et al.*, 2002, p. 6). For example, before coming to a greater conclusion, it may be necessary for a student to perform a routine use of procedures or demonstrate comprehension in order to proceed.

Ball *et al.* (1998) give the following example:

In your own words, describe each of the following and give an example of each:

- (i) A vector space
- (ii) A subspace of a vector space
- (iii) A spanning set
- (iv) A linearly independent set
- (v) A basis of a vector space V

Here, asking students to use their own words requires students to demonstrate a *comprehension* of the terms, and then to use *information transfer* to give their own explanation. The second part of the question which asks for an example then requires the student to use *implications, conjectures and comparisons*.⁵

Furthermore, there may be occasions when answering one particular question may call on different skills from different students. This is particularly the case when constructing proofs, as students may answer this through rote learning (FKFS), whereas others may construct the proof based on an appreciation of definitions and following through an argument (JI; ICC). So this could involve Group A skills from the former and Group C skills from the latter. At the more novice level, a well-rehearsed student who is familiar with many past questions may not view a particular question as requiring ‘application in new situations’ (AINS) because they have seen similar examples of ‘novel’ questions. For example, in A-level examinations, patterns and trends can often be picked up in the types of questions posed. Therefore, a student familiar enough with sufficient past papers might begin to view questions which require AINS as being rather routine as they identify the common procedures required in order to answer such questions correctly. Consequently, such a question may be classified as RUOP for one student and AINS for another.

As individual students may view the same mathematical task differently in terms of perceptions of difficulty and approach, perhaps due to their own mathematical experience, classification can also be problematic (Berry *et al.*, 1999). However, since the MATH taxonomy focuses on mathematical demands rather than difficulty, the impact of student perceptions of difficulty on classification is minimized with this particular taxonomy. Students who are well-prepared for questions of a more demanding type may not find them as difficult as a student who is not.

It is possible that one person will think that a question belongs in one category, whereas another may disagree. In fact, a task may not fit easily into one particular category. Smith *et al.* (1996) address this issue, explaining that they do not aim to permit the classification of every task, but to provide examiners with a means by which to design assessments which call upon a variety of skills and knowledge from students. They say, ‘the examiner’s judgement, objectives and experience . . . determine the final evaluation of an assessment task’ (p. 68). However, this will be the same of any taxonomy as there are bound to be subjective cases in all instances where two individuals may not agree on what category the question belongs in, save for the most basic distinctions.

⁵ This was not a concern for this particular project. Use of mark schemes meant that part marks awarded for such questions could be distributed by group.

3. Data

In order to ascertain whether there are any differences in the challenges presented by mathematics assessment at A-level and undergraduate level, the MATH taxonomy (Smith *et al.*, 1996) was applied to a selection of papers from each of the two.

Each individual question was analysed using the MATH taxonomy by the researcher. Samples of the coding were validated by a recent undergraduate mathematics graduate, a current mathematics undergraduate, a practising secondary mathematics teacher and a professor of mathematics education. The mark schemes were consulted in order to ascertain the methods by which examiners expected students to answer the questions, which served to provide an insight into the skills that were anticipated and required in the examinations. The proportions of marks available in each paper were calculated for each of Groups A, B and C.

3.1 A-level examinations

In order to gain an insight into the challenges posed by the questions in A-level Mathematics and Further Mathematics, I decided to analyse only the modules Core Mathematics 1 (C1) and Further Pure Mathematics 3 (FP3). These were chosen because C1 is, for most students, the first A-level Mathematics module they study, and FP3 is the most advanced module in pure mathematics for students of A-level Further Mathematics.⁶ This means that concentrating on these two modules was an opportunity to analyse questions at both novice and advanced levels in this qualification. One paper from each examination board for each of FP3 and C1 were analysed, with these being chosen at random. Only papers from 2006 or later were considered, because this is when the A-level Mathematics and Further Mathematics syllabi and examination structures underwent their most recent change.⁷

Exact papers and examination boards have been anonymized in order to prevent comparisons between examination boards, given such a small sample.

Table 3 shows that the majority of marks in both C1 and FP3 papers were awarded for responses to Group A questions, with this being particularly high in C1. Very few marks were awarded for responses to Group C questions in both modules; however, more marks were awarded for Group C questions than Group B questions in paper G. This appears to be an anomaly, although applying the MATH taxonomy to more papers would be the only way of establishing whether that is the case.

The main difference which is apparent between C1 and FP3 papers is that FP3 papers award proportionally fewer marks for Group A questions than C1, with many more Group B questions. As FP3 is the most advanced pure mathematics module in A-level Further Mathematics, it was expected that it would involve more advanced questions which test students on their ability to make justifications, comparisons, interpretations and conjectures. However, despite predictions that FP3 would contain more Group C questions than C1 due to the fact that the further pure modules introduce proof and group theory and because of claims by examining boards that Further Mathematics ‘students are encouraged to think logically, practically and analytically’ (AQA, 2013, online), this transpired not to be the case other than in paper G. Mathematical proof is not generally something which can be done by use of Group A or Group B skills; however, it is possible that some can be reduced to a ‘routine use of procedures’ or reproduced as ‘factual knowledge and fact systems’ should the syllabus and question type facilitate it.

⁶ AQA also offers FP4.

⁷ Changes will come into place in January 2014 when the January examination sessions will stop, and A-levels will move to a linear system of examinations, with modular A-levels scrapped from September 2015.

TABLE 3. Total marks available for Group A, B and C questions in each A-level paper analysed (%)

Module	Paper	Year	MATH		
			Group A	Group B	Group C
C1	A	2006	90.7	4.0	5.3
C1	B	2006	68.0	29.3	2.7
C1	C	2007	88.9	11.1	0.0
C1	D	2010	88.0	5.3	6.7
Mean C1			83.9	12.4	3.7
FP3	E	2007	76.0	24.0	0.0
FP3	F	2006	66.7	33.3	0.0
FP3	G	2007	48.6	22.2	29.2
FP3	H	2010	53.3	40.0	6.7
Mean FP3			61.2	29.9	9.0
Mean C1 and FP3			72.6	21.2	6.4

Amongst A-level papers, the majority of Group A marks consisted of ‘routine use of procedures’ (88.7%), the majority of Group B marks from ‘application in new situations’ (93.9%) and all of the Group C marks were for ‘justifying and interpreting’ (Table 4). This supports claims and empirical studies concerning the procedural nature of A-level Mathematics (Crawford *et al.*, 1994; Taverner, 1997; Crawford *et al.*, 1998a,b; Alton, 2008; Bassett *et al.*, 2009). However, it should be noted that more able students of A-level Mathematics (such as future undergraduates) are better able to turn rehearsed procedures akin to Group A questions into conceptual knowledge and incorporate them into their schema of mathematical concepts (Tall & Razali, 1993). Therefore, the reliance on procedures at A-level does not necessarily mean that students will have no conceptual understanding of the mathematics that they have studied.

These findings support the work of Berry *et al.* (1999) and Monaghan (1998), whose categorization of A-level Mathematics questions as either ‘routine’ or ‘non-routine’ concluded that the majority of available marks in the examinations were awarded for answers to routine questions. Furthermore, it corroborates with Etchells & Monaghan’s (1994) research using the MATH taxonomy which found that most marks awarded in A-level Mathematics were for Group A tasks.

3.2 Undergraduate examinations

Based on the definitions of each group in the MATH taxonomy, each part question of each examination between 2006 and 2012 was classified accordingly, and the average number of marks available for each examination in each group calculated.

First-year students do a number of examinations in the summer, but only questions from Pure Mathematics I and Pure Mathematics II were analysed because the A-level examinations analysed were also pure mathematics. Both these topics are core to undergraduate courses nationwide and the examination format used is the traditional means of assessment in the UK (Iannone & Simpson, 2011). Pure Mathematics I focused on topics in algebra and Pure Mathematics II on topics in real analysis.

TABLE 4. *Proportions of available A-level marks by subgroup*

Group	A			B		C		
Subgroup	FKFS	COMP	RUOP	IT	AINS	JI	ICC	EVAL
Marks within group (%)	0.9	10.7	88.7	6.1	93.9	100.0	0.0	0.0

TABLE 5. *Proportions of Group A, B and C questions in Pure I and Pure II examinations*

Paper	Group		
	A (%)	B (%)	C (%)
Pure Mathematics I (Algebra)	29.1	21.2	49.9
Pure Mathematics II (Analysis)	34.0	7.6	58.3
Mean	31.6	14.4	54.1

TABLE 6. *Proportion of available undergraduate marks by subgroup*

Group	A			B		C		
Subgroup	FKFS	COMP	RUOP	IT	AINS	JI	ICC	EVAL
Marks within group (undergraduate) (%)	90	6	4	17	83	87	13	0
Marks within group (A-level) (%)	1	11	89	6	94	100	0	0

Table 5 shows that the majority of marks awarded in first-year pure mathematics examinations (54.1%) were for questions requiring students to demonstrate skills from Group C of the MATH taxonomy—that is, ‘justifying and interpreting’ and ‘implications, conjectures and comparisons’. The minority of marks were awarded for questions requiring Group B skills, and approximately one-third of marks available for the demonstration of Group A skills.

The spread of marks by group was slightly different between Pure Mathematics I and Pure Mathematics II (Table 4). Pure Mathematics II is much more abstract than Pure Mathematics I as it focuses on analysis, whereas topics in algebra have the potential to be assessed through asking students to demonstrate their understanding of certain procedures, or applying them, in this topic area. This means that the possibility of being awarded marks for Group A and B questions might be greater in algebra. This is reflected in the proportion of marks available for correct responses to Group B questions being higher in Pure Mathematics I than II.

Unlike in A-level Mathematics where the majority of Group A marks were for a ‘routine use of procedures’ (88.7%), in undergraduate mathematics examinations, the majority of Group A marks were for ‘factual knowledge and fact systems’ (90.0%), with ‘routine use of procedures’ forming a

very small percentage of these marks (4.0%) (Table 6). The majority of Group B marks were for ‘application in new situations’ (83.0%) and the majority of Group C marks for ‘justifying and interpreting’ (87.1%).

4. Observations

The analysis of these three different examination types has revealed substantial differences in terms of the types of questions posed to students at each level.

The vast majority of A-level questions were Group A, in particular most of these questions asking students to perform a routine use of procedures in answering the questions. This is common to both C1 and FP3. The most extreme instance of this is in paper C, where 88.9% of the available marks were awarded for answers to Group A questions. All but a few of these questions were ‘routine use of procedures’; that is, students are able to answer the majority of questions through the use of well-practised procedures which they will be familiar with from doing work during lessons, and which are themselves similar in nature to previous examinations. Both C1 and FP3 have the majority of their marks in Group A which suggests that, despite FP3 being a more advanced module, the examinations do not challenge students in a different way to earlier A-level Mathematics and Further Mathematics modules.

Conversely, undergraduate mathematics examinations focus on students’ responses to Group C questions, with 54% of available marks awarded for answers to Group C questions. Notwithstanding this, a large number of marks were also available for responses to Group A questions (32%). These questions either take the form of statements of definitions and theorems or of proofs of statements which students will have seen in their lecture notes. In such instances, for the purpose of this analysis, the students have been assumed to have answered those questions as a consequence of memorization (FKFS) than through using proof techniques to do it themselves. This leaves a fairly large proportion of the marks available for reproduction of knowledge—something which could be achieved by someone who took the time to memorize it, not necessarily an undergraduate mathematician. In all of the examinations analysed, being able to answer these questions alone would easily be sufficient to earn a candidate at least a pass (30% or higher), with most examinations analysed having the potential to reward students with only Group A skills with a third-class mark (40% or higher). In two cases, both in Pure Mathematics II, a candidate only demonstrating Group A skills could have earned a first-class mark (70% or higher).

In advanced mathematics, knowledge of precise definitions and theorems is key, hence their being tested in undergraduate examinations; however, it seems that there are instances when too great a proportion of questions asked are from Group A and not enough in Group C. The difference in proportions of Groups A and C in A-level Mathematics and undergraduate mathematics is great and it appears that the gap between the two is not bridged by Further Mathematics. At undergraduate level, there is little opportunity for students to demonstrate routine use of procedures, as the mathematics being studied is very abstract and any computational processes which could be assessed do not form a large part of what is studied. However, this is more likely to take place in algebra than analysis as algebra directly lends itself to applied processes more so than analysis.

The nature of advanced mathematics is reflected in the higher concentration of Group C questions at the undergraduate level compared to A-level. At the tertiary level, mathematics is formal, rigorous and deductive, unlike secondary mathematics which inculcates ‘a purely knee-jerk response in students’ (Bibby, 1991, p. 43). Furthermore, the type of mathematics studied at school and university ‘changes from “What is the result?” to “Is it true that . . .?”’ (Dreyfus, 1999, p. 106).

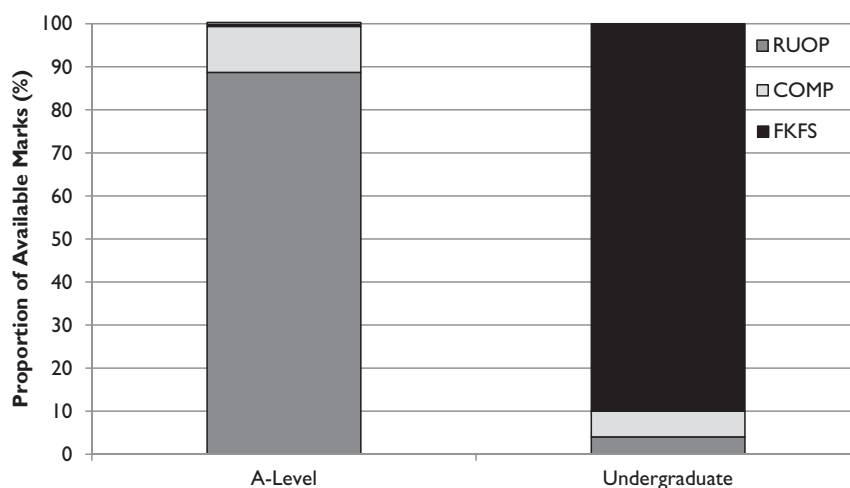


FIG. 1. Make-up of Group A questions.

An obvious difference lies in the types of questions asked between each of these stages of assessment comes from the grouping of questions within each of Groups A, B and C. For example, at A-level, the majority of Group A questions were ‘routine use of procedures’, whereas at undergraduate level these were ‘factual knowledge and fact systems’ (Fig. 1). Unlike at A-level, where routine use of procedures form the majority of in the demands of the examination in general, this is not the case at undergraduate level, where Group C questions tend to constitute the majority of the marks on offer. This difference in the make-up of the Group A questions stems from the fact that A-level Mathematics appears to be very procedural, and that undergraduate mathematics is not, but the questions which require Group A skills tend to just be those which involve factual recall of definitions or, to a lesser extent, the construction of proofs which could be memorized from lecture notes. The ‘definitions questions’ can give students who do not believe themselves capable of answering other questions which require application of these definitions the opportunity to gain some marks, while the questions which require students to reproduce proofs do not necessarily require students to memorize them parrot fashion without understanding what they are learning.

5. Conclusion and implications

While the structure and content of A-levels are currently going under change, this article nonetheless highlights the mismatches between secondary and tertiary mathematics. The gaps identified go towards explaining undergraduate mathematicians’ difficulties in the first year, and the data could inform policy and development of the new qualification in mathematics.

The difference between secondary and tertiary mathematics in terms of the types of questions that are posed appears to be great (Fig. 2), aside from the difficulty of both the questions themselves and the material being examined. The large increase in the proportion of Group C questions between school and university is indicative of the changing nature of mathematics between these two points, and relates to previous work which found that the majority of marks awarded in the University of Oxford’s mathematics admissions test were for Group C questions (Darlington, 2013). Furthermore, the differences in the nature of Group A questions at each level is indicative of the contrasting nature of the mathematics studied and examined at school and university.

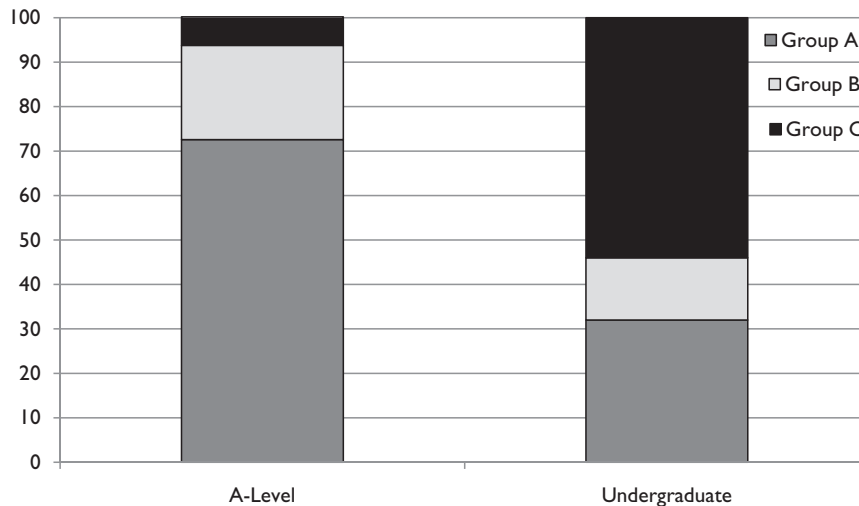


FIG. 2. Proportion of questions by group.

While the A-level is not intended purely for those students wishing to study undergraduate mathematics, it arguably has a responsibility of preparing them for tertiary mathematics study. Further Mathematics perhaps has a greater responsibility in terms of exposing students to advanced mathematics; however, the analysis here suggests that it does not provide students with a broader conception of mathematics than the standard Mathematics course. It is important that A-level students are exposed to the mathematics and mathematics questions which can give them an indication of how secondary and tertiary mathematics might differ, as well as having them develop sufficient technical skills in calculus, algebra and trigonometry which they require as the basis for undergraduate study. This has become a matter of concern at a number of UK universities, as increasing numbers are conducting diagnostic tests on new undergraduates in these topics in order to ensure that they are sufficiently proficient in their use of basic concepts (Edwards, 1996; Lawson, 2003; MathsTEAM, 2003; Gillard *et al.*, 2010; Williams *et al.*, 2010).

The findings here suggest that A-level Mathematics and Further Mathematics' main focus is on assessing students' abilities to repeat procedures, rather than to develop mathematical skills. For students to be successful at undergraduate level it is necessary that they are able to solve problems, to interpret information that they are given and choose methods for a purpose, to justify why things may be true, to deduce the implications of results and to make judgements about the value of material that they are given for a purpose. All of these skills are Group C skills, and very few of them are tested in A-level students. Any reform of examinations at this level and at undergraduate level therefore would benefit from taking these Group C skills into account as it is important that school-level advanced mathematics examinations serve to develop the skills that are necessary for a successful transition into university study.

Funding

Economic and Social Research Council (ESRC) studentship at the University of Oxford's Department of Education.

REFERENCES

- ALTON, A. (2008) When is predictability a problem in an examination? *Paper Presented at the 34th International Association for Educational Assessment (IAEA) Annual Conference*. Cambridge, UK.
- ANDERSON, J., AUSTIN, K., BARNARD, T. & JAGGER, J. (1998) Do third-year mathematics undergraduates know what they are supposed to know? *Internat. J. Math. Ed. Sci. Tech.*, **29**, 401–420.
- ANDERSON, L. & SOSNIAK, L. (eds) (1994) *Bloom's Taxonomy: A Forty-Year Retrospective*. Chicago, IL: National Society for the Study of Education.
- AQA (2013) *A-Level Further Mathematics: Why Choose this Specification*. Retrieved 5 November 2013, from <http://www.aqa.org.uk/subjects/mathematics/a-level/further-mathematics-6360/why-choose>.
- BALL, B., SMITH, G., WOOD, L., COUPLAND, M. & CRAWFORD, K. (1998) Creating a Diversity of Mathematical Experiences for Tertiary Students. *Internat. J. Math. Ed. Sci. Tech.*, **29**, 827–841.
- BARNETT, R. (1990) *The Idea of Higher Education*. Buckingham: Open University Press.
- BASSETT, D., CAWSTON, T., THRIVES, L. & TRUSS, E. (2009) *A New Level*. London: Reform.
- BERRY, J., MAULL, W., JOHNSON, P. & MONAGHAN, J. (1999) Routine questions & examination performance. *Proceedings of the 23rd International Conference of the International Group for the Psychology of Mathematics Education*, 2 (O. Zaslavsky ed.). Haifa, Israel, pp. 105–112.
- BIBBY, N. (1991) Wherefore 'Plug-and-Chug'? Computer algebra vs. A-level mathematics. *Math. Gazette*, **75**, 40–48.
- BIGGS, J. B. & COLLIS, K. F. (1982) *Evaluating the Quality of Learning: The SOLO Taxonomy*. New York, NY: Academic Press.
- BLOOM, B., ENGELHART, M., FURST, E., HILL, W. & KRATHWOHL, D. (eds) (1956) *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I: Cognitive Domain*. New York, NY: David McKay.
- CHAN, C. C., HONG, J. H. & CHAN, M. Y. (2001) *Applying the Structure of the Observed Learning Outcomes (SOLO) Taxonomy on Students' Learning Outcomes: A Comprehensive Review*. Unpublished Manuscript: Hong Kong Polytechnic University.
- CRAWFORD, K. (1983) Past experiences & mathematical problem solving. *Proceedings of the 7th Annual Conference of the International Group for the Psychology of Mathematics Education*. Haifa, Israel.
- CRAWFORD, K. (1986) Cognitive & social factors in problem solving behaviour. *Proceedings of the 10th Annual Conference of the International Group for the Psychology of Mathematics Education*. London, UK, pp. 415–421.
- CRAWFORD, K., GORDON, S., NICHOLAS, J. & PROSSER, M. (1993) Learning mathematics at university level: initial conceptions of mathematics. *Proceedings of the 16th Annual Conference of the Mathematics Education Research Group of Australasia (MERGA)*. Brisbane, Australia, pp. 209–214.
- CRAWFORD, K., GORDON, S., NICHOLAS, J. & PROSSER, M. (1994) Conceptions of mathematics & how it is learned: the perspectives of students entering university. *Learn. Instr.*, **4**, 331–345.
- CRAWFORD, K., GORDON, S., NICHOLAS, J. & PROSSER, M. (1998a) Qualitatively different experiences of learning mathematics at university. *Learn. Instr.*, **8**, 455–468.
- CRAWFORD, K., GORDON, S., NICHOLAS, J. & PROSSER, M. (1998b) University mathematics students' conceptions of mathematics. *Stud. High. Educ.*, **23**, 87–94.
- D'SOUZA, S. & WOOD, L. (2003) *Tertiary Students' Views about Group Work in Mathematics*. Retrieved 26 September 2012, from *Proceedings of the Australian & New Zealand Associations of Research in Education Joint Conference*. Auckland, New Zealand. <http://www.aare.edu.au/03pap/dso03154.pdf>.
- DARLINGTON, E. (2013) The use of Bloom's Taxonomy in advanced mathematics questions. *Proceedings of the British Society for Research into Learning Mathematics* (C. Smith ed.). Bristol, UK, pp. 7–12.
- DREYFUS, T. (1999) Why Johnny can't prove. *Educ. Stud. Math.*, **38**, 85–109.

- DUVAL, R. (2006) A cognitive analysis of problems of comprehension in a learning of mathematics. *Educ. Stud. Math.*, **61**, 103–131.
- EDWARDS, P. (1996) *Implementing Diagnostic Testing for Non-Specialist Mathematics Courses*. London: Open Learning Foundation.
- ELIA, I., VAN DEN HEUVEL-PANHUIZEN, M. & KOLOVOU, A. (2009) Exploring strategy use & strategy flexibility in non-routine problem solving by primary school high achievers in mathematics. *ZDM Math. Educ.*, **41**, 605–618.
- ETCHELLS, T. & MONAGHAN, J. (1994) Hand held technology: assessment issues. *Acta Math. Univ. Comenianae*, **3**, 119–130.
- GALBRAITH, P. L. & HAINES, C. R. (1995) Students' mathematical characteristics: some developmental skills for undergraduate programs. *Bull. Inst. Math. & Appl.*, **31**, 175–179.
- GILLARD, J., LEVI, M. & WILSON, R. (2010) Diagnostic testing at UK universities: an email survey. *Teach. Math. Appl.*, **29**, 69–75.
- HUGHES, R., MONAGHAN, J., SHINGADIA, E. & VAUGHAN, S. (2006) Revisiting routine questions. *Teach. Math. Appl.*, **25**, 90–96.
- HUNTLEY, B., ENGELBRECHT, J. & HARDING, A. (2009) Can multiple choice questions be successfully used as an assessment format in undergraduate mathematics? *Pythagoras*, **69**, 3–16.
- IANNONE, P. & SIMPSON, A. (2011) The summative assessment diet: how we assess mathematics degrees. *Teach. Math. Appl.*, **30**, 186–196.
- KADIJEVIĆ, D. (2002) TIMSS 2003 mathematics cognitive domains: examination, operationalization & implications. *Z. Inst. Ped. Instr.*, **34**, 96–102.
- KILPATRICK, J. (1993) The chain & the arrow: from the history of mathematics assessment. *Investigations into Mathematics Education: An ICMI Study* (M. Niss ed.). Dordrecht, the Netherlands: Kluwer Academic Publishers, pp. 31–46.
- KREITZER, A. & MADAUS, G. (1994) Empirical investigations of the hierarchical structure of the taxonomy. *Bloom's Taxonomy: A Forty-Year Retrospective* (L. Anderson & L. Sosniak eds). Chicago, IL: National Society for the Study of Education, pp. 64–81.
- LAWSON, D. (2003) Changes in student entry competencies 1991–2001. *Teach. Math. Appl.*, **22**, 171–176.
- LEINBACH, C., POUNTNEY, D. C. & ETCHELLS, T. (2002) Appropriate use of a CAS in the teaching & learning of mathematics. *Internat. J. Math. Ed. Sci. Tech.*, **33**, 1–14.
- MAGUIRE, S., EVANS, S. E. & DYAS, L. (2001) Approaches to learning: a study of first-year geography undergraduates. *J. Geogr. Higher Educ.*, **25**, 95–107.
- MALABAR, I. & POUNTNEY, D. C. (2002) Using technology to integrate constructivism & visualisation in mathematics education. *Proceedings of the 2nd International Conference on the Teaching of Mathematics at the Undergraduate Level* (I. Vakalis, D. Hughes Hallett, D. Quinney & C. Kourouniotis ed.). Heronissos, Crete.
- MATHSTEAM (2003) *Diagnostic Testing for Mathematics*. Birmingham: Learning & Teaching Support Network.
- MONAGHAN, J. (1999) *An investigation into students' global A-level performance & their performance in routine & non-routine questions*. Retrieved 22 August 2013, from <http://www.education.leeds.ac.uk/research/projects/global-alevel-mathematics-performance>.
- NOVAK, J. D. (1978) An alternative to Piagetian psychology for science & mathematics education. *Stud. Sci. Educ.*, **5**, 1–30.
- ORMELL, C. P. (1974) Bloom's taxonomy & the objectives of education. *Educ. Res.*, **17**, 3–18.
- PÓLYA, G. (1945) *How To Solve It*. New York, NY: Anchor Books.
- POUNTNEY, D., LEINBACH, C. & ETCHELLS, T. (2002) The issue of appropriate assessment in the presence of a CAS. *Int. J. Math. Educ. Sci. Tech.*, **33**, 15–36.
- PRING, R. (1971) Bloom's taxonomy: a philosophical critique. *Cambridge J. Educ.*, **1**, 83–91.

- QUALIFICATIONS & CURRICULUM AUTHORITY. (2007) *Evaluation of Participation in GCE Mathematics: Final Report*. Retrieved 26 February 2013, from <http://www.ofqual.gov.uk/719.aspx>.
- RESNICK, L. B. (1987) *Education & Learning to Think*. Washington, DC: National Academy Press.
- ROMBERG, T. A., ZARINNIA, E. A. & COLLIS, K. F. (1990) A new world view of assessment in mathematics. *Assessing Higher Order Thinking in Mathematics* (G. Kulm ed.). Washington, DC: American Association for the Advancement of Science, pp. 21–38.
- SELLEN, J., MASON, A. & SELLEN, A. (1989) Can average calculus students solve nonroutine problems? *J. Math. Beh.*, **8**, 45–50.
- SELLEN, J., SELLEN, A. & MASON, A. (1994) Even good calculus students can't solve nonroutine problems. *Research Issues in Undergraduate Mathematics Learning, MAA Notes*, **33**, 19–26.
- SMITH, G., WOOD, L., COUPLAND, M., STEPHENSON, B., CRAWFORD, K. & BALL, G. (1996) Constructing mathematical examinations to assess a range of knowledge & skills. *Int. J. Math. Educ. Sci. Tech.*, **27**, 65–77.
- TALL, D. & RAZALI, M. R. (1993) Diagnosing students' difficulties in learning mathematics. *Int. J. Math. Educ. Sci. Tech.*, **24**, 209–222.
- TAVERNER, S. (1997) Modular courses - a drip feed approach to teaching & learning? *Teach. Math. Appl.*, **16**, 196–199.
- WILLIAMS, J., HERNANDEZ-MARTINEZ, P. & HARRIS, D. (2010) Diagnostic testing in mathematics as a policy & in practice in the transition to higher education. *Paper Presented at the Conference for the British Educational Research Association*. University of Warwick, Coventry, UK.
- WOOD, L. (2011) Power & passion: why I love mathematics (B. Barton & M. Thomas, eds). *CULMS Newsletter*, 5–14.
- WOOD, L. & SMITH, G. H. (2002) Perceptions of difficulty. *Proceedings of the 2nd International Conference on the Teaching of Mathematics at Undergraduate Level* (M. Boezi ed.). Hersonissos, Crete.
- WOOD, L., SMITH, G. H., PETOCZ, P. & REID, A. (2002) Correlations between students' performance in assessment & categories of a taxonomy. *Proceedings of the 2nd International Conference on the Teaching of Mathematics at the Undergraduate Level* (M. Boezi ed.). Hersonissos, Crete.

Ellie Darlington is a research officer in assessment research and development at Cambridge Assessment who has recently completed a doctorate in mathematics education at the University of Oxford.