The Summative Assessment Diet: How we assess in mathematics degrees

P. Iannone* and A. Simpson

*Corresponding author: Paola Iannone. School of Education and Lifelong Learning, University of East Anglia, Norwich, NR4 7TJ, UK. Phone: 0044 (0)1603 591007, Email: p.iannone@uea.ac.uk
Abstract

Much research and many papers on mathematics pedagogy have discussed assessment and, in particular, the need to provide a varied diet of methods by which students are assessed for the award of their degree. We explore the mix of assessment methods provided across a range of UK university mathematics departments, examine whether there is a relationship between the mix of assessment methods and type of institution or between the type of assessment and the mathematical topic.

Keywords: Assessment at university, mathematics, UK
1. INTRODUCTION

How we assess and what we assess gives a clear indication of what we value and directs what and how students learn (van de Watering, Gijbels, Dochy and van der Rijt, 2008). There are clear steers from UK government towards coursework based assessment focussed on employability that should apply across all degree subjects (Dearing, 1997; Browne, 2010), but there are also indications from the learned societies in the UK that mathematics necessarily has “distinctive features of … assessment and teaching” which would suggest the need for a particular pattern of assessment (LMS, 2010, p1).

Clearly these recommendations need not be in direct conflict; both the generalist literature and the mathematics specialist literature has made calls for a wider range of assessment types to be used, whether to address the unique nature of mathematics or to widen the range of skills assessed or both (Kneale, 2009; Challis, Houston and Stirling, 2006). However, Birenbaum (1994) notes:

Traditional assessment methods however are still in use alongside the new and renewed ones, making ours the era of "pluralistic assessment". It seems that this will remain the case for at least the near future, as alternative assessment is still in its infancy and suffers from unresolved psychometric and other implementation-related difficulties. (pg. 239)

In this paper we assess the extent of the plurality of assessment in university mathematics in the UK. We review the literature in this area and then present the results of a representative survey of assessment methods in UK mathematics departments. We explore the varied diet within assessment methods and then examine the extent of
assessment methods which might be labelled as ‘innovative’ and the mathematical topics where such methods are used.

We are aware that all mathematics departments provide assessments with a formative function and much effort goes in to providing feedback in both non credit bearing assignments and much credit bearing work and lecturers hope students use this feedback to assess and develop their own understandings. However, much evidence from the research literature shows that the higher the stakes, the more assessment influences students’ approaches to learning and preparing for tasks (Scouller, 1998), so for the purposes of this article, we focus on assessment as credit bearing, summative tasks.

2. BACKGROUND

In the general education literature, there is a long history of research into different assessment methods and much of it points to the introduction of what are termed more ‘innovative’ methods. Birenbaum (1994) provides a scheme for characterising assessment methods across factors such as the timeframe in which it is conducted, the simplicity and authenticity of the tasks, the availability of tools, the evaluator etc.; in all potentially giving over 10 billion combinations for different types of assessment. In practice, van de Watering et al. (2008) conflate these into twelve and Iannone and Simpson (2011) reduce these to eight types of assessment either in common use in mathematics or featuring regularly in the mathematics education literature:

- Closed book examinations
- Dissertation
- Open book examination
Multiple choice test
Oral examination
Regular example sheets
Project
Presentation

Note that neither van der Watering et al. nor Iannone and Simpson address the mode of delivery of the assessment nor whether assessment is individual or groupwork. Clearly assessment can be delivered as a paper and pencil test taken in an examination hall and marked by a qualified member of staff, or can be delivered as a ‘quiz’ sat at a computer and marked instantly by the CAA system (Sangwin, 2004). Similarly a project could be a piece of work which students are required to undertake on their own, or a piece of groupwork where marks are assigned to members of the group using some agreed method (Houston, 2001). In this paper we follow van der Watering et al. and Iannone and Simpson by constraining all the possible variants to the manageable categorisation above.

Within this categorisation, the closed book examination is seen as the ‘traditional’ form of assessment which one might assume has always dominated universities. According to Stray (2001), however, prior to the 18th century assessment in universities was dominated by the public oral examination, and it was only with the start of the “Senate House Examination … later the Mathematical Tripos” (pg.36) that more written examinations and finer gradations in marks were introduced. These began to dominate to the extent that by the end of the 19th century, almost all assessment was undertaken by written (closed book) examination (with exceptions in areas such as languages and
medicine). Stray (2001) attributes this shift to the wish to assess individual cognitive progress, the corruption of the oral examination system and the growing predominance of Newtonian mathematics.

The closed book examination retained an unquestioned near monopoly until well into the latter part of the 20th century. However, in many subject areas, people began to question whether this form of assessment genuinely gave a valid and reliable measure of students’ ability (Cox, 1967) and whether the set of skills which they assess might be too restrictive (Heywood, 1977). There has been extensive research in university assessment since then and much of it has been focussed on varying the diet of assessment.

However, it is not always clear that this research has explored different degree courses in equal depth. For example, Struyven, Dochy and Janssens (2005) review a wide range of literature on student perceptions about assessment in higher education – in thirty articles, which between them detail research involving over 5000 students, not one of the participants was studying mathematics (nor were there any from many other subjects such as chemistry, geography or theology).

That is not to say that the questions about assessment methods have been ignored in mathematics. Griffiths and McLone (1984) produced a critique of the abilities assessed by the then stereotypical university mathematics closed book examination. Hirst and Biggs (1968) described the introduction of a project module arguing that undergraduate mathematics students are not aware of the way in which mathematics is discovered and
that a research project would help the students recognise “the muddy swamp of half-formed ideas which is at its [text-book mathematics] source” (pg.252).

More recently, there have been clear calls to widen the mathematics assessment diet to include such alternatives as posters (Berry and Houston, 1995), multiple choice (Haines and Crouch, 2005) or mathematical writing (McConlogue, Mitchell and Vivaldi, 2010) while others simply call for any move to ‘innovate’ away from “the dominant epistemological paradigm which is one of absolutism of knowledge …[which is]… reproductive and content dominated” (Burton and Haines, 1997, pg.275).

In a more balanced document, Challis et al. (2004) spell out key skills, adapted from Dearing (1997), which a mathematics degree should seek to develop and thus, they argue, should play a part in assessment. These included communication, use of IT, problem solving, group work and improving one’s own learning. While noting that these should be assessed alongside (not entirely separated from) specific competencies in mathematics, they raise the question of the extent to which these skills can be addressed in a diet mainly restricted to closed book examinations.

Some mathematics departments have clearly developed more varied diets: Povey & Angier (2006) report on the assessment methods used at one UK university which include oral presentations, posters, projects as well as a traditional closed book exams. They report:

The criteria for assessment allow a wide range of skills to be acknowledged, for example, posing problems as well as solving them or communicating their mathematics visually or orally. Mathematical imagination is valued. Secondly, the
students have the opportunity to become aware of their own progress and to find out about themselves as learners of mathematics. […] Thirdly, many of the assessments involve negotiation, either with their tutors or with their peers or with both. In some cases, this challenges standard conventions of where authority lies, for example, devising the criteria by which they are to be assessed or deciding, in part, how marks are to be allocated amongst themselves at the end of group projects. (Povey & Algier, pg 45)

Other authors have investigated how projects (often paired with a presentation) can develop some of the key skills required form graduates today. For example, Ramesh (2009) describes how projects, especially when part of the final year assessment, can develop students’ individual thinking and independent work.

Thus, it appears that there are many calls for and attempts to vary the assessment diet in the UK. But, to the best of our knowledge, there is no recent survey which examines that diet, or explores where any variation occurs. Thus, our research questions are:

- What mix of assessment methods do mathematics departments in the UK offer?
- How are different assessment methods used in different topic areas? and
- Are there links between the assessment mix and the types of university mathematics department?

3. METHODS

Much of the work detailed above tends to report on the individual, the new and the idiosyncratic and there is the danger that by seeking out particular assessment methods we give the image of over-representing what is called ‘innovative’. There is also some
suggestion that universities with different aims and different student intakes may take
different approaches to the mix of assessment they provide.

Thus we took the approach of developing a systematic, representative survey. To
provide a measure (albeit a crude one) against which to stratify a sample of universities,
we began with a publicly available ‘league table’ of mathematics departments. Given
the difficulty in obtaining complete data in the full depth required from every
mathematics department, we chose one department from each group of six universities
listed in order down the table (they are coded in order as M1 to M11 in the subsequent
discussion).

For each department we obtained full details for their mathematics degree. Where
departments offer multiple degrees, we chose the one closest to a full time, three year,
BSc in mathematics. We obtained details of the number of credits the students take per
year and the weight of the marks accrued in each year towards the final degree award.
For each university, we obtained details of every module available for students on the
BSc in Mathematics. Many universities allow considerable flexibility to take modules
in other subjects and, where this was the case, these modules were not taken into
account in the analysis – that is, only modules provided by the mathematics department
were considered. For each module we also noted the number of credits the module
bears towards the final degree and whether the module is optional or compulsory.

For each module we then recorded the assessment method taking care to include as
many details as possible. In many cases, modules are assessed by a combination of
assessment methods and here we recorded the weight of each assessment method toward the final mark for the module, the type of assessment methods used and so on.

4. ANALYSIS AND FINDINGS

It was immediately obvious that the modal assessment method across our sample was closed book examination. The complexity of the different choices and options within departments was such that we took the pragmatic decision to make calculations which might over-inflate any instances of non-closed book examinations. Thus, in order to examine the influence of closed book examinations (CBE) we calculated the average percentage of the final degree mark which the students would accrue from CBEs by taking the full range of modules, which we then weighted for the credit awarded for results in different year groups. A summary of these main findings are shown in Table 1.

Note that M10 seemed a considerable outlier in the proportion of CBE contributing to the final degree. On closer examination, however, it was clear that the mathematics department distinguishes ‘class tests’ from examinations, but these are apparently in traditional closed book form but are administered locally by the department, rather than centrally by the university. If we include those under the CBE heading, M10 still has the smallest proportion of the final award coming from CBEs, but the proportion is more in line with that expected from the trend amongst other departments.

<table>
<thead>
<tr>
<th></th>
<th>Weight (Y1/Y2Y3)</th>
<th>Number of modules offered</th>
<th>Modules assessed by 100% CBE</th>
<th>Proportion of final award assessed by CBE</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>(10/30/60)</td>
<td>57</td>
<td>22</td>
<td>79.83%</td>
</tr>
<tr>
<td>M2</td>
<td>(0/40/60)</td>
<td>55</td>
<td>41</td>
<td>88.37%</td>
</tr>
</tbody>
</table>

Comment [1]: We need to make sure all the tables go on one page.
It is clear then, that the assessment diet is relatively restricted, from just under 50% to just under 100% of the final award comes from results students obtain in closed book examinations (and recall that our method under-estimates the proportion of CBE). An explorations of the trends down the table also suggests that there is some link between our (admittedly crude) measure of the background of the university and the reliance on CBE. Indeed, there is a very strong and significant correlation between the ranking and the proportion of CBE contributing to the final degree \( r(9) = -0.845, p<0.001 \) and the institutions coded M1-M6 all have at least one module assessed entirely by CBE with M7-M11 having two or more assessment methods for every module with a CBE component.

Despite this, one has to conclude that the assessment diet, even in these institutions is still very clearly dominated by CBEs.

4.1 Alternatives to the closed book examination
To examine the assessment diet in more depth, we explored the data for each department in more depth; in particular, we looked for situations in which closed book examination was not the sole or main method used. We found some commonalities and some idiosyncrasies, so we categorised the topics which appeared to have less assessment by CBE into:

- Statistics
- History of mathematics and mathematics education
- Projects
- Others

4.1.1 Statistics

Every mathematics department surveyed offers statistics and/or probability modules as part of their main mathematics degree, as either compulsory or optional modules. Across the sample, these modules tended to be assessed by a combination of small projects and CBE, or by coursework and open book exam or CBE. Indeed, in our sample the only use of open book examinations was in statistics modules. Table 2 summarises the assessment of statistics in our sample.

<table>
<thead>
<tr>
<th>Depts</th>
<th>Modules</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>• Probability (3 modules) (87.5% CBE, 12.5% coursework)</td>
</tr>
<tr>
<td>M2</td>
<td>• Probability and statistics (6 modules) (100% CBE)</td>
</tr>
<tr>
<td></td>
<td>• Applied probability (100% CBE)</td>
</tr>
<tr>
<td>M3</td>
<td>• Probability and statistics (4 modules) (95% CBE and 5% coursework)</td>
</tr>
<tr>
<td></td>
<td>• Probability and statistics (6 modules) (95.75% open book, 4.25% coursework)</td>
</tr>
<tr>
<td>M4</td>
<td>• Probability and statistics (2 modules) (90% exam+tables, 10% coursework)</td>
</tr>
<tr>
<td></td>
<td>• Applied probability (100% CBE)</td>
</tr>
<tr>
<td>M5</td>
<td>• Statistics and probability (3 modules) (90% exam unseen, 10% coursework)</td>
</tr>
<tr>
<td>M6</td>
<td>• Probability and statistics (4 modules) (85% exam+tables, 15% coursework)</td>
</tr>
<tr>
<td></td>
<td>• Statistics (7 modules) (64.28% exam, 35.72% coursework)</td>
</tr>
</tbody>
</table>
Departments M4 and M6 state that the use of statistics tables is allowed in some of their exams. They are not described formally as ‘open book exams’ in the official documentation, but the fact that students are allowed access to supporting materials would mean they are classified as open book in our scheme. Department M3 is the only departments to classify some of their exams for statistics as open book exams, whilst departments M6, M5, M7, M8 and M10 describe the assessment generically as “exams” or “unseen exams”, although it is difficult to imagine that many of these exams can be sat without the use of statistics tables. From the documentation provided, it appears as if the exams reported in Table 2 are traditional pen and pencil exams.

4.1.2 History of Mathematics and Mathematics Education.

Three out of the eleven department surveyed offer a history of mathematics module and five offer one or more modules in mathematics education taught within the modules offered by the mathematics department. In all cases these modules are optional and are assessed by a combination of assessment methods which do not always include CBE (see table 2 for details).
### Table 3: Assessment of history of mathematics and mathematics education in the sample

One might argue that the intellectual tradition to which these modules belong bears a closer resemblance to social sciences or humanities than to mathematics and the pattern of assessment here may reflect those traditions.

#### 4.1.3 Projects

All mathematics departments surveyed except M11 offer at least one module assessed entirely by an individual project. This module is often for final year students and can be either compulsory or optional. In Table 3 we have summarised the assessment of the project module across departments:

<table>
<thead>
<tr>
<th>Dept</th>
<th>Year</th>
<th>Modules assessed by projects with no exams</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>Y1</td>
<td>• Experimental mathematics (6 credits) (100% written)</td>
</tr>
<tr>
<td></td>
<td>Y2</td>
<td>• Second year essay (6 credits) (80% written, 20% presentation)</td>
</tr>
<tr>
<td></td>
<td>Y3</td>
<td>• Essay (15 credits) (80% written, 20% presentation)</td>
</tr>
<tr>
<td>M2</td>
<td>Y3</td>
<td>• Project (20 credits) (90% written, 10% presentation)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Project (40 credits) (90% written, 10% presentation)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Project (10 credits) (90% written, 10% presentation)</td>
</tr>
<tr>
<td>M3</td>
<td>Y3</td>
<td>• Group project (10 credits) (65% written, 15% presentation), 10% individual contribution)</td>
</tr>
</tbody>
</table>

* These modules have a placement component and the teacher mentoring the students in school will assess the students.
As we can see from Table 4, in all cases the largest percentage of the marks is given to a written report on a topic in mathematics (normally following individualised supervision with a member of staff). A presentation (if assessed) usually accrues between 10-20%, except M10 where for “Project management for mathematics” the presentation and the written report accrue the same percentage of marks. These presentations often include a question and answer session which may be open to other students and lecturers. In one of the departments surveyed, M3, the project is completed in groups.

4.1.4. Idiosyncrasies

Clearly each individual department has their own unique way of assessing across their own pattern of modules and the previous sections pointed out some commonalities. We list in table 3 the ‘idiosyncrasies’ – that is those modules which do not have a major CBE component which do not fall into our previous categories.
### Table 5: Other modules not assessed by CBE in the sample

As we can see M1 and M10 have a range of modules assessed by combination of methods (which in this case include computer based assessment, small projects, multiple choice tests, worksheets etc) whilst for the other departments the range of modules not assessed by CBE is very small. M6 and M10 have a range of modules.*

* The names of some of the modules in this table have been changed to preserve anonymity of the institutions but reflect the module content.
explicitly aimed at assessing study skills, employability skills and skills related to independent study and in the assessment of those modules there is no CBE component, perhaps reinforcing the belief that employability and communication skills cannot be adequately assessed by CBE.

5. DISCUSSION AND CONCLUDING REMARKS

We addressed the three research questions:

- What mix of assessment methods do mathematics departments in the UK offer?
- How are different assessment methods used in different topic areas? and
- Are there links between the assessment mix and the types of university mathematics department?

It is clear that the modal form of summative assessment (indeed, the majority form in almost all cases) remains closed book examination. In addition, there are clear categories of topic which emerge where CBE is not the main assessment method: statistics, history of mathematics, mathematics education and final year projects.

For statistics, we might argue that the unusual presence of open book examinations indicate a strong wish on the part of lecturers (and perhaps a strong tradition in this field) to focus on process over memory – statistics may require the recollection of many more formulae which may be very much harder (and less valuable) to reconstruct than in pure and applied mathematics. It may also be that the use of technology in handling data has a longer tradition than the use of packages in supporting the development of pure mathematics proofs and calculations.
The history and education modules, one could argue, are based in quite different, more literate traditions, in which the argumentation schemes are distinct from mathematical proof. Again, then, it may be no surprise that these account for a reasonable number of non-CBE assessed modules.

In some of the more idiosyncratic examples, the modules were more directly tied to employment skills and personal development and these tended to be pan-university initiatives manifested in particular forms in the mathematics departments.

Despite this, there is some evidence of core mathematical thinking being assessed without CBEs. The project, which was advocated as a way of giving students some insight into the mathematical research process as long ago as the 1960s (Hirst and Biggs, 1968), can also enable students to provide evidence of communication, organisation and presentation skills. There is also evidence that some very traditional topics, even at very high ranking universities, do permit alternatives to CBE – for example, the linear algebra module in M1.

Recall that, because we wanted a measure of representativeness, we have not deliberately sought the most innovative method, the newest module or the most enthusiastic lecturer. But we do know they exist – the pages of Teaching Mathematics and its Applications and newsletters such as the MSOR Connections regularly report on alternatives to CBEs.

But it is clear that the assessment diet remains CBE heavy. There may, of course, be very good reason for this – closed book exams are relatively easy to set, administer and
mark; they are seen as harder to cheat in than most other forms of assessment and they require a balance of memory, application and understanding which many mathematicians may feel suits the subject. Indeed, Iannone and Simpson (2011) provide evidence that mathematics students often prefer examinations over alternatives such as projects and presentations. As the learned societies in mathematics argue, there may just be something particular about mathematics which fits the CBE model best or, at least, fits it least worst (LMS, 2010); to misquote Winston Churchill (talking of democracy) in most cases closed book examinations may be the worst form of mathematics assessment except all those other forms that have been tried from time to time.

ACKNOWLEDGMENT

We would like to thank the Maths, Stats & OR Network of the Higher Education Academy in the UK for funding the research reported in this paper.

Notes on Contributors:

Paola Iannone is lecturer in educational research in the School of Education and Lifelong Learning in the University of East Anglia. Her research interest is teaching and learning mathematics at university level, with particular focus on assessment and proof and proof production.

Adrian Simpson is reader in mathematics education in the School of Education at Durham University and is also the Principal of Josephine Butler College. His research has primarily focused on students’ thinking across the transition from school to university mathematics, the transition to independent graduate study and mathematical logic and rationality.
REFERENCES


