Undergraduate mathematics and statistics assessment practices in Australia

CRISTINA VARSAVSKY\textsuperscript{a*}, KAREN HOGEBOOM\textsuperscript{a}, CARMEL COADY\textsuperscript{b}, DEBORAH KING\textsuperscript{c}

\textsuperscript{a} School of Mathematical Sciences, Monash University, Clayton, Australia
\textsuperscript{b} School of Computing, Engineering and Mathematics and statistics, University of Western Sydney, Penrith, Australia
\textsuperscript{c} Department of Mathematics and statistics, University of Melbourne, Melbourne, Australia
\textsuperscript{*}Corresponding author: cristina.varsavsky@monash.edu

Keywords: Mathematics and statistics assessment, undergraduate, graduate attributes, threshold learning outcomes

Within the higher education sector, the current focus is on graduate outcomes and the development of threshold learning outcomes for the various disciplines, including mathematics and statistics. We ask the question, how can these outcomes be demonstrated through the assessment of units of study students undertake as part of their undergraduate program? Our study focusses on the types of assessment used across Australia, using a web search methodology to analyse information publically available on the institutions’ websites. The findings confirm that mathematics and statistics students in Australia are still largely assessed through closed book in-semester tests and final examinations, contributing 70% and 30% respectively to the final grade of their units of study. Our results broadly match those of a similar study done in the UK, with some differences, particularly in the percentage contribution of closed book assessment to final grade in statistical units and across year levels. Finally, implications of assessment practices for certification of student achievements are discussed.

1. Introduction

A focus on graduate learning outcomes is the current trend in higher education around the world [1, 2, 3], and Australia is no exception. The establishment of the Tertiary Education Quality and Standards Agency (TEQSA) in 2012 heralds significant changes in the Australian higher education sector with institutions now required to go through a periodic re-accreditation cycle [4]. The implementation phase of the regulatory requirements has been based on defined higher education provider standards. The next stage will also involve course design and learning outcomes standards, and these will soon form a normal part of the re-accreditation process.

Prior to this shift to a regulated higher education sector in Australia, and following similar international developments, the Australian Learning and Teaching Council (ALTC) (now called the Office of Learning and Teaching (OLT)) has invested in developing the Learning and Teaching Academic Standards (LTAS) [5]. As a result of this project, the various discipline groups developed the Threshold Learning Outcomes (TLO) for their disciplines. The mathematics and statistics TLOs [6] have been developed to align with the Science TLOs [7]. Although the learning outcomes standards are still under development, and no final decision has been made on how TEQSA will monitor attainment of these
standards, there are strong indications that institutions will be able to use the discipline TLOs as reference points for demonstrating the quality of their graduates.

The mathematics and statistics TLOs articulate the minimum characteristics a mathematics or statistics graduate should have at the point of graduation. The TLO graduate characteristics are grouped under the following broad headings [6]:

- **Understanding**: Demonstrating a coherent understanding of the mathematical sciences.
- **Knowledge**: Exhibiting depth and breadth of knowledge in the mathematical sciences.
- **Inquiry and problem solving**: Investigating and solving problems using mathematical and statistical methods.
- **Communication**: Communicate mathematical and statistical information, arguments, or results for a range of purposes using a variety of means.
- **Responsibility**: Demonstrate personal, professional and social responsibility.

The question is then, how would institutions demonstrate that their mathematics and statistics graduates have these qualities? We argue that assessment is central to this development. The role of assessment in shaping student learning has been widely advocated for many years [see for example, 8, 9, 10]. Assessment tasks signal to students what they must learn, and formative assessment provides students with feedback as they develop understanding of new concepts. In this new regulated environment, assessment as certification for what students learn, is gaining new importance. If assessment is designed carefully, then students’ assessed work would be the most effective way to produce evidence for attainment of learning outcomes standards.

Within the context of the standards agenda, this paper provides a level of understanding of the assessment practices in undergraduate mathematics and statistics across Australian institutions, which will help to inform the debate around effective and defensible approaches for demonstrating graduate outcomes. This study focuses on the certification aspect of assessment. It extends to Australia the study conducted in the UK by Iannone and Simpson [11, 12]. It asks the questions

1. What are the types of assessment tasks used in mathematics and statistics units of study (subjects) in Australia?
2. What is the contribution of each type of assessment towards the marks and/or grades awarded to students on completion of the units of study?
3. Are Australian assessment practices in undergraduate mathematics and statistics different to those used in the UK?

2. Methodology

A search through institutional websites was the methodology followed to find out what types of assessment tasks are used in Australian institutions in undergraduate mathematics and statistics units of study. This methodology mirrors the approach used in the UK by Iannone and Simpson [11, 12]. In Australian universities it is common practice to make basic information about each unit offering available publically, including learning outcomes, synopsis, prerequisite studies and assessment tasks, together with their contribution to the final marks and grades for the units.

Thirty Australian higher education institutions offer a degree in mathematics or statistics; in most cases mathematics and statistics are offered as a major study within the Bachelor of Science. Our web search resulted in finding information about assessment breakdown for 16 of these institutions, including six of the Group of Eight research intensive universities which produce the majority of mathematics and statistics graduates in Australia.
Only units offered and taught by a mathematics or statistics department were considered. A total of 395 undergraduate units were found, of which 88 were first year units, 124 second year units, and 183 third year units.

3. Findings and discussion

3.1. Closed Book Assessment

Closed book assessment in the form of final exam and test was found to be the most dominant form of assessment in the majority of undergraduate mathematics and statistics units in Australian universities. Just under 90% of the units (345) included a final examination at the end of the semester; of these 38% (128 units) also had an invigilated test during the semester. Final examinations and in-semester tests are taken under supervision, and students are given a limited time to complete them. Figure 1 shows the distribution of the percentage contribution closed book assessment (final examination and in-semester test(s)) makes to the final grade of students undertaking the units. It shows that undergraduate mathematics and statistics units across Australia are largely assessed through a closed book assessment, amounting to between 60 and 80 per cent of the students’ grade for the units. Project units were the only units that did not have invigilated closed book assessment.

![Figure 1. Percentage contribution of closed book assessment to final grade, in 395 undergraduate mathematics and statistics units at 16 Australian universities.](image)

When unit assessments from all universities are pooled and examined by year level, we observe a general trend for the mean percentage contribution of closed book exams to decrease with increasing year level. This trend is not just seen in individual universities, but also across all universities (Figure 2). The mean percentage of closed book assessment decreases by 7% between first year and third year. There are a number of universities which have a final year project unit (6 units in total). These project units have no closed book assessment and hence the inclusion of these units may skew the mean values. Following the methodology of Iannone and Simpson (2011), the six units specifically labelled as a project were removed. Even without these units, the general trend is still decreasing, and the mean percentage of closed book assessment decreases by 8.5% between first year and third year.

This drop through year levels is different to the trend found by Iannone and Simpson (2012). They found that closed book assessment increases with increasing year level in UK
universities. Between first and third year, the mean percentage of closed book assessment increased by 10%. When projects were excluded, the mean percentage of closed book assessment tasks increased by 15% from first year to third year. This is an interesting difference which is difficult to explain with the available information; in the Australian context this may be due to the significant difference between first year and higher year class sizes which make in-semester assessment more manageable.

**Figure 2.** Mean percentage contribution of closed book assessment to final grade in undergraduate mathematics and statistics units in Australia, by year level.

The medians, on the other hand, are remarkably consistent across all year levels (70% for each of the three year levels and for all year levels combined). The spread, as measured by the interquartile range (the middle 50 per cent) increases from first to second year, and then decreases strongly in third year (Figure 3).

### 3.2. The assessment of statistics

Iannone & Simpson [11, 12] found a different pattern of assessment in statistics units as compared to non-statistics units. The observed difference in assessment practices may be due to their definition as to what constitutes an “open” exam whereby “students are allowed access to supporting materials”, which includes the use of statistical tables. They also found that there is a greater use of computer packages in many of the statistical units in their study.

No evidence of a different pattern in assessment practices can be discerned in Australian universities. The study found and examined 72 statistics units, covering all levels of undergraduate statistics. The proportion of closed book assessment is consistent with other non-statistics units. The mean of percentage contribution of closed book assessment to the grade in statistics units is 65.3% and for non statistics units it is 65.9%. Both groups have the same median (70%). The interquartile range of the statistical units shows less spread than non-statistical units. Also, the range of the closed book assessment of statistical units is entirely contained within the range of non-statistical units. (Figure 4). All these features indicate no discernible difference in closed book assessment patterns between statistical and non statistical units. Further, there is no difference in the use of computer based assessment in statistical units when compared to non-statistical units in Australia.
Undergraduate mathematics and statistics assessment practices in Australia

Figure 3. Median percentage contribution of closed book assessment to final grade in undergraduate mathematics and statistics units in Australia, by year level.

Figure 4. Percentage contribution to final grade of closed book assessment in 324 non statistical units and 72 statistical units in Australia.

3.3. University rankings versus proportion of closed book assessment

Plotting the mean percentage contribution of closed book assessment for each university against their ranking within Australia according to The Higher Education rank [13] showed no relationship between the average percentage of closed book assessment and a university’s league table position (Figure 5).

This is in contrast with UK universities, where Iannone and Simpson [11, 12] found a direct relationship between university rank and the contribution of closed book exam to the final grade for the unit. That is, the higher the university ranked, the higher the proportion that closed book assessment contributed to the unit grade. Again, the scope of this study does not allow to elaborate on these differences; further investigation is required.

3.4. Assignments

Assignments were the second most common form of assessment found in our web search. In the Australian context, mathematics and statistics assignments usually consist of a number of problems or exercises. Of the 395 units in the study, 264 (67%) have assignments in some form declared in their assessment schedule. Another 59 units list a “continuous” assessment component, which could include assignments. The mean percentage contribution to assessment, when averaged over all undergraduate mathematics and statistics units is 27.5%. The average contribution of assignments increases by year level, matching the decrease in closed book assessment with a 22.4% in first year, 26.1% in second year and 30.6% in third year.
3.5. Other assessment types
In addition to assignments and closed book exams, we found a diverse range of assessment types being used for the assessment of undergraduate mathematics and statistics units, but each of these appeared only a small number of times.

In addition to six third-year project only units where the entire mark for the unit is based on a project, we found research projects in 45 units (11%) across all year levels. The majority of these also included a presentation element whose contribution to the final mark ranged from 10% to 40%. We also found 13 units not related to a research project that required oral presentations; these were mostly in third year and contributing from 5% to 35% to the final grade (median 10%).

Given the readily available mathematics and statistics software, online testing facilities, and sophisticated learning management systems, it was surprising to find out that only 17 units (4%) explicitly state computer based assessment, either in the form of online quizzes or in the use of software packages. The majority of online assessment occurs in first year units (15 units) with contribution to overall unit assessment ranging from 5% to 30%.

Group projects appeared in only four first year units, contributing from 10% to 30% to the final grade. Twenty units have portfolios; one university alone is responsible for 19 of these portfolios, with 10% to 15% contribution towards the final grade. The remaining assessment task types included essays (5 units, 10 to 30%), oral exam (1 unit, 40%), lab book record keeping (2 units, 20 and 25%), literature reviews (3 units, 40 to 50%), case studies (5 units, 20%), guided discussion (1 unit, 5%), reflection (1 unit, 10%).

3.6. Are TLOs being assessed?
Our study shows that assignments (problem sets), hour-long tests and examinations, which form part of a century long tradition in mathematics and statistics assessment, seem to remain the norm. Can these practices be defensible, considering that “in general, students consider important those aspects of instruction that their teacher emphasizes and assesses regularly” [14]?

Mathematicians are used to formulating the learning outcomes for the students in terms of procedural knowledge (performing correct steps in a valid sequence) rather than student
achievements in using mathematics and statistics (mathematical thinking and analysis) [10]. This results in idiosyncratic assessment practices that favour routine mastery, leaving a gap between what students “know” and what they “can do”. Madison [10] explains this as “tensions between what is practical and is effective in judging student performance and understanding”, and that “being unable to see evidence of value in the hard work of effective assessment, [mathematicians] often rely on the results of practical methods, believing they are measuring similar or highly correlated results”. Bergqvist, on the other hand, leaves open the possibility that “there may be a gap between what the teachers believe is tested [in closed book exams] and what is actually tested” [15].

How can these assessment practices focussing largely on closed book examinations align with the current higher education agenda towards greater accountability and the requirements to demonstrate learning outcomes? Are the mathematics and statistics TLOs assessed in any meaningful way throughout the units students take in completing their program of study? Closed book tests and exams, undertaken within a set limited time, may be appropriate for assessing procedural knowledge, and could, to some minor extent address the “knowledge” and “inquiry and problem solving” dimensions of the TLOs. It is hard to see how these types of assessment could effectively measure performance in the “communication” and “responsibility” aspects. If these are assessed at all, the only place would be through the relatively low stakes assignments. Furthermore, as stated by Houston, “learning mathematics for the principal purpose of passing examinations often leads to surface learning, to memory learning alone, to learning that can only see small parts and not the whole of a subject, to learning wherein many of the skills and much of the knowledge required to be a working mathematician are overlooked” [9].

Universities have been addressing the skills agenda in various ways, including integration of skills in key units at each year level, or the so called “capstone units” at the end of the program where students draw together what they learn from the point of admission [16]. However, in the context of mathematics and statistics, the only units found in our search that could fit the “capstone” definition were research project units. Not all students undertake project units where available; such units are offered only to the very high achieving students, and not all take up this opportunity.

4. Conclusion
A wide range of assessment types are used across Australian mathematics and statistics departments for assessing their undergraduate students. However, the predominant forms of assessment used are closed book exams and in-semester tests and assignments, and the typical contributions these make to the student final grade are 70% and 30% respectively. While assessment in higher education has taken some prominence in recent years, in undergraduate mathematics assessment is still a minority culture backed by centuries of tradition [15]. Our findings largely align with practices in the UK reported by Iannone and Simpson. Although there is evidence of less traditional assessment approaches such as reflection, essay writing and portfolios, these appear to correspond to small pockets of innovation. Finally, it should be noted that this study only looks broadly at the types of assessment tasks students are asked to complete towards the mathematics and statistics units they enrol in, and it does not reveal the specific tasks each of these involves.

Our investigation of assessment types was initiated around the question of whether attainment of the nationally agreed threshold learning outcomes could be demonstrated through assessment tasks our students complete as part of their undergraduate studies. Given the strong focus on final examinations which are likely to test only procedural knowledge, there seems to be a gap between what mathematicians and statisticians think their graduates...
should be able to do (as articulated in the nationally endorsed threshold learning outcomes) and the types of assessment they use to drive learning and to grade their achievements. Conclusions, however, cannot be made based only on the assessment types used. A comprehensive review of the assessment tasks (that is, what students are actually asked to do in the different assessment tasks) is needed to decide if current assessment practices are fit for this new purpose or if a radical change is required.

References


