Threshold tests as a way to encourage long-term, self-regulated learners in engineering

Danica Solina[1], Chris Wong[2], Kate Crawford[3], Elaine Huber[4]

[1]School of Mathematical and Physical Sciences, University of Technology Sydney, Australia, [2]Faculty of Engineering and IT, University of Technology Sydney, Australia, [3]Eviva Pty Ltd, Australia, [4]The University of Sydney Business School, University of Sydney, Australia.

Abstract

Traditionally, educators evaluate the effectiveness of a new assessment in a pre-requisite early subject via pass rates, grades and student satisfaction. It may be more appropriate to measure the impact on student results in later stage subjects. In this work we report on the impact of changing assessment, over 10 years, at a metropolitan Australian university for an initial calculus subject (Mathematics 1) on two follow-on subjects: Mathematics 2 and Fundamentals of Mechanical Engineering. Earlier research found that Online Mastery Tests can harm later learning outcomes even though failure rates drop within the pre-requisite subject. Here we show that Paper Threshold Tests, requiring greater engagement, metacognitive strategies result in fewer passes in the pre-requisite subject however, they were also resulting in a major inversion of the grade distribution toward higher grades in the follow on subjects where 60% of students now obtain marks > 75%.

Keywords: *Assessment; mastery learning; progression analysis.*

1. Introduction

A university education serves multiple purposes including providing a strong foundation in the field of choice, transferable practical skills, critical thinking skills, and networking opportunities. The underlying objective is the creation of lifelong learners who have the capability to keep up with advancements in their field and to adapt to changes in their careers (Nilson, 2013).

However, many students commence university only having mastered surface learning techniques, resulting in an erosion of learning and decay of knowledge/ learning loss between semesters (Dills et al., 2016). To combat this, assessments should be designed to encourage students to develop deep learning approaches to maximise knowledge retention for application in the subsequent subjects in the next semesters.

As university subjects often have large enrolments (n >500), it is not feasible to personalise subjects for each student. Instead, assessments should be structured to allow students to develop the learning skills to progress through Bloom's taxonomy (Bloom et al., 2001) moving from recall, to understanding, application, evaluation and creation.

This work examines the impact of the type of assessment used in a pre-requisite mathematics subject at an Australian university on students' long-term retention of concepts and their ability to transfer their learning to subsequent subjects. The focus will be on the comparison of Online Mastery Tests (OMTs) and Paper Threshold Tests (PTTs) in the first-year calculus subject, Mathematics 1 (Math1) and their impact on the grade distribution in the two following subjects, Mathematics 2 (Math2) and Fundamentals of Mechanical Engineering (FME). The investigation covers a ten-year period, with the latter subject (FME) having remained unchanged in delivery or assessment. This work had human ethics clearance (HREC 17-1158).

2. Assessment structures

2.1. Math1 assessment and within subject performance

Math1 is structured to run over 12 weeks where there are 11 weeks of lecture material and 11 weeks of associated tutorials, and in some years, computer labs. Math1 is offered in all three study sessions. Session 1(S1) has the largest intake and consists primarily of students attempting the subject the first time. Session 2 (S2) includes repeat students and students who had completed Foundation Mathematics (FM) before attempting Math1. Session 3 (S3) was introduced in the summer of 2017 to give students a chance to catch up on Math1 if they had either failed or completed FM in S2.

Between 2008 and 2020, assessment in Math1 was altered in an effort to improve learning outcomes and pass rates. The first three iterations included a 40% Compulsory Final Exam (CFE), one (later two) Class Tests (CT), weekly review questions called Routine and Review Sheets (RRS) and Mathematica Computer Lab Participation marks (CLP) .

The fourth iteration of Math1 assessment was the most significant change, which introduced Online Mastery Tests (OMTs) through the Webassign the online platform. These tests were a combination of Mastery learning (Block & Burns, 1976) and Keller Plan/Personalised system of instruction (Rae, 2011). The concept behind Mastery Tests is that if students have a strong grasp of fundamentals, then they can tackle more complex concepts. These tests can be repeated (within reason) to give the student the time to attain proficiency.

It was required that students obtain 80% in each of four OMTs. Students were given three opportunities to attain the 80% requirement in each of the four OMTs. The first OMT was on assumed knowledge and was worth a nominal 5%. The following three tests covered 2 to 3 weeks of content over the first 8 weeks. It was possible to attain a pass in Math1 by attaining the minimum requirement in each of the OMTs and an Optional Final Exam (OFE). OMTs only required a final answer to be submitted whilst OFEs required fully worked solutions.

In S2 2017 the OMTs were replaced with the paper equivalent, Paper Threshold Tests (PTTs). The Math1 assessments and weighting are listed in Table 1.

Table 1. Assessment iteration in Math1, session and years used and assessment breakdown.

Iteration, Session & Year	Assessment Breakdown and percentage weight		
	Quiz/Assess	Mastery test	Final Exam
1. S1 and S2 2008	5% RRS, 25% CT, 10% Assignment	-	60% CFE
2. S1 2009	10% RRS, 25% CT, 5% CLP	-	60% CFE
3. S2 2009 – S1 2014	10% RRS, 25% 2×CT, 5% CLP	-	60% CFE
4. S2 2014 – S2 2016		62.5% OMT	37.5% OFE
5. S1 2017		48% OMT	52% CFE
6. S2 2017 – S3 2020		50% PTT*	50% CFE*

Note: * Note weightings changed to 50/50 Autumn2018).

The grade distribution in Math1 for each assessment iteration is given in figure 1 where F:Fail (<50%), P: Pass(50–64%), C:Credit (65–74%), D:Distinction (75–84%) and H: High Distinction (85–100%).
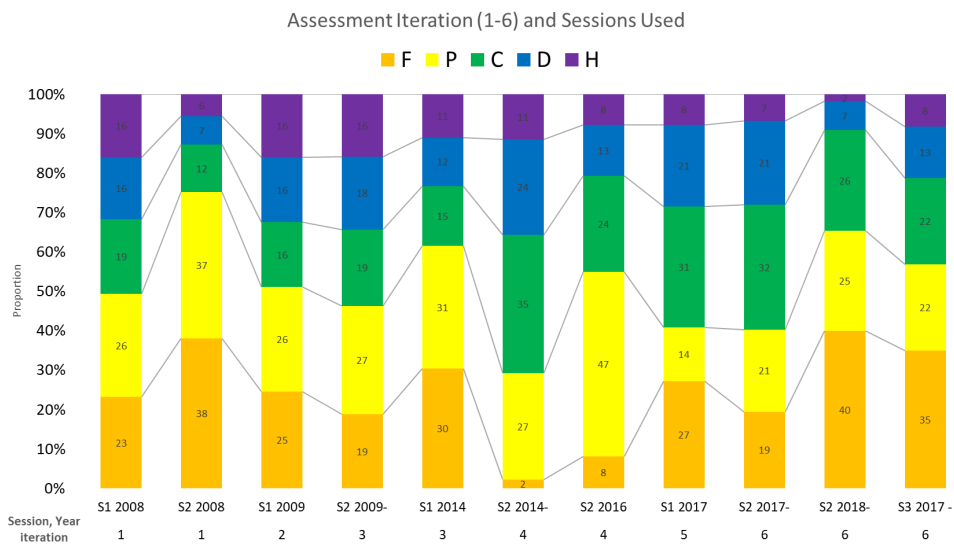
Figure 1. Grade distribution in Math1 with assessment iteration (as outlined in Table 1) and session

The introduction of OMTs in iteration 4 (S2 2014) resulted in a significant drop in fails. Although this decrease seemed positive, tutors began to report that students were just pattern matching instead of fully understanding the material. This minimal approach carried over into Math 2 as well, according to a progression analysis by (Coupland et al., 2017).

Considering these findings, Math 1 assessment underwent a 5th iteration, in which a minimum requirement was introduced for the Compulsory Final Exam (CFE) - now requiring fully worked solutions.

Figure 1 shows the reintroduction of a final exam (iteration 5-S1) resulted in a significant increase in fails. The graphed data is the result after a supplementary exam was given. Pre-supplementary exam the failure rate was 46%. It was concluded that OMTs were not supporting learning of the subject material. Discussions with 80% (n=100) students who attempted the supplementary exam showed that OMTs did not align well with final exam expectations. Students believed they were doing well due to high OMT scores (80%+) but did not invest extra effort on the final exam. Furthermore, OMT only required final answers whilst the final exam required fully worked solutions.

In iteration 6 (S2 2017), OMTs were replaced with PTTs. It should be noted the only other subjects taken in common during their starting semester, Chemistry and Physics, had little change in delivery at the time PTTs were introduced.

As with the first OMT, the first PTT also examined assumed knowledge. The first test was administered in weeks 1 and 3 to give students a chance to either revise or withdraw before

census date in week 4. A score of 80% was required to pass. The questions were structured to reflect the sequence in which the material would be covered in the subject. Throughout the semester, the test was referred to as a reminder of the assumed knowledge.

The last three PTTs examined 2 to 3 weeks of content at a time, covering the first 8 weeks. Tests had similar questions to the OMTs, however students were required to show complete working/reasoning where a question was given more than 1 mark.

These tests examined general procedural knowledge and application. Application was included to place the mathematics in context and to break the habit of using only "x, y, z" as variables. These tests had a threshold requirement of 70%. As with the OMTs, students were allowed to sit the tests three times in order to achieve the required mark. Tests were administered in the second hour of a two hour tutorial, with feedback given a week after the first attempt of each test.

From figure 1, it is clear that the introduction of the PTTs resulted in a significant increase in failure rate. However, grade distribution within a subject and low failure rate may not accurately measure the quality of the learning outcomes.

2.2. Progression Analysis

An alternative approach is to examine the impact of Math 1 performance on the subsequent subjects. A typical mechanical engineering student math program is as follows:
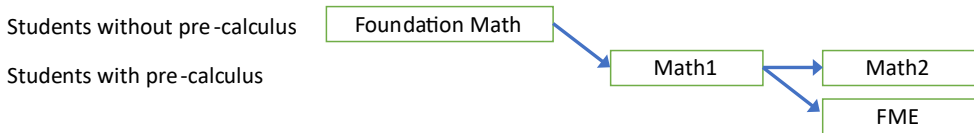
Math2 has undergone a number of assessment iterations. For clarity, we have labelled these iterations A, B and C. These are listed in Table 2 and Figure 2.

Table 2. Assessment iteration in Math2, session and years used and assessment breakdown.

Iteration, session, years	Assessment Breakdown and percentage weight		
A. S1 2008 – S3 2013	30% CTs, 10% Assignments		60% CFE
B. S1 2014 – S1 2017		62.5% OMTs	37.5% OPE
C. S2 2017 – S3 2020		52% (50%*) OMTs	48% (50%*) CFE

Note: *weight change in S2 2017

The grade distribution data in Math1 from 2008 to 2018 is presented in figure 2. The data is grouped based on differences in assessment and iteration. Note that fails in Math2 were high before the inclusion of OMTs (iteration 3B).

The introduction of a compulsory final exam in Math1 (iteration 5C) shows a shift to higher grades (D/H) in Math2. It is unclear however whether this shift occurs because of the introduction of the compulsory final exam in Math 1 (iteration 5C) or the introduction of the compulsory final exam in Math 2 (iteration 5C/6C).
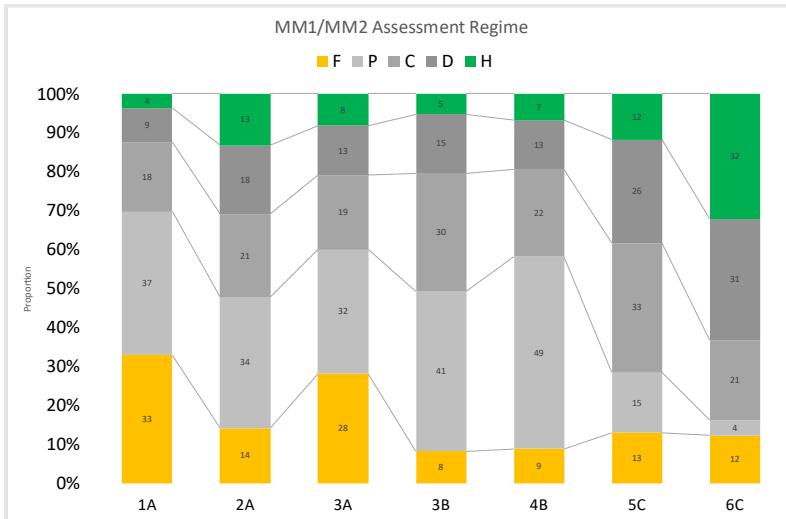

Figure 2. Grade distribution in Math2 with both Math1 assessment (numbers) and Math2 assessment (letters)
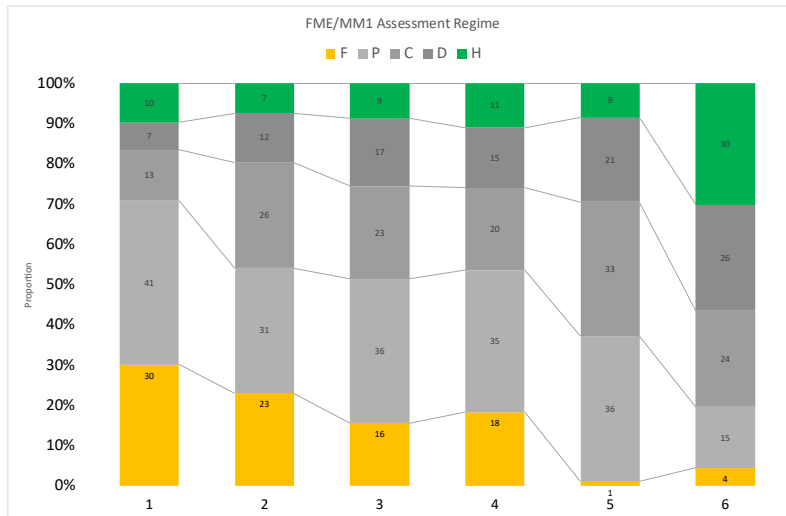

Figure 3. Grade distribution in FME with Math1 assessment iterations

A better indicator is FME, where the subject assessment has not changed appreciably during the time of analysis. The inclusion of the final exam in Math1 (iteration 5) resulted in a dramatic reduction in FME fails and the introduction PTTs in Math 1 (iteration 6) resulted in significant increase in higher grades (D/H) as per Figure 3.

The introduction of PTT's in Math1 lead to a significant improvement in student performance in both Math2 and FME, as shown by the grade distribution in Figures 2 and 3. This resulted in over 60% of students in Math 2 and 55% in FME achieving a Distinction or higher and demonstrates successful transfer of knowledge applied to subsequent subjects.

Furthermore, the delivery of Math 2 was almost identical for 5C and 6C highlighting the impact of the effect of Math1 assessment on Math2. Moreover, the results in FME indicate that students were competently applying the mathematics that they learned in Math1 to a subject in the Engineering Discipline.

This grade inversion was found to be robust for students coming from Math1 from S2 2017 to S12019 into Math2 and FME from S1 2018 to S2 2019 indicating the learning outcomes were not a one-off phenomenon and skills were transferable to follow-on subjects.

The most important aspect of PTTs is that they were only used in a first semester subject, yet the learning outcomes and skills were carried into follow on subjects without PTTs. This suggests that the assessment methods used in the first semester play a crucial role in determining a student's approach to learning in subsequent semesters.

PPTs effectively promote deep learning by encouraging the use of metacognitive strategies instead of surface rote learning (McGuire, 2018). Use of application problems require students to identify important information, use mathematical concepts, and demonstrate their thought process through clear worked solutions and evaluation of the results.

Furthermore, the timing of the three tests follow Nilson's (2013) recommendation for the development of self-regulated learners. The first test provides an initial evaluation of the student's knowledge and effectiveness of their learning approach, with feedback in the following week to encourage an early start. The second test two weeks later allows the student to reflect on the change in their learning approach. The final test at the end of semester offers a chance to refine their learning approach, helping them to build and retain their mathematical skills.

3. Concluding remarks

We demonstrated the effectiveness of PTTs in a first-year calculus subject by analysing progression data to improve learning outcomes and ensuring long-term retention of knowledge. The results showed the importance of implementing PTTs in early semester subjects which can have a positive impact on subsequent subjects.

However, there are drawbacks to using PTTs. Creating 12 distinct PTTs, solution sets, and marking criteria is a time-consuming process, and more versions are required as class size grows to maintain academic integrity. This requires more institutional support for this form of assessment such as appropriate workload allocation and additional resources.

PTTs were difficult to administer during COVID (2020) where S1 had 800+ & S2 300+ students. Due to cost of invigilation for multiple tests, a PDF download/upload approach was taken with reliance on student honesty. Unfortunately, some students cheated, leading to changes in assessment for 2021, and the impact is yet to be analysed.

Learning conditions and quality at many universities have declined due to cost constraints. From this research, investing more resources in first year can have a significant impact on student success in later years by creating independent learners with better long-term outcomes.

References

Block, J. H., & Burns, R. B. (1976). Mastery Learning. *Review of Research in Education*, *4*, 3–49. https://doi.org/10.2307/1167112

Bloom, B., Airasian, P.W., Cruikshank K. A., Mayer, R. E., Pintrich, P.R., Raths, J. & Wittrock, M. C. (2001). *A taxonomy of learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives.* New York, NY: Longman.

Coupland, M., Solina, D., & Cave, G. E. (2017). Mastery Learning: Improving the Model. In A. Downtown, S. Livy, & J Halls (Eds), *40 years on: We are still learning! Proceedings of the 40th Annual conference of the Mathematics Education Research Group of Australasia* (pp. 189-196). Melbourne: MERGA. https://eric.ed.gov/?id=ED589527

Dills, A., Hernández-Julián, R., & Rotthoff, K. W. (2016). Knowledge decay between semesters. *Economics of Education Review*, *50*, 63–74. https://doi.org/10.1016/j.econedurev.2015.12.002

Nilson, L. B. (2013). *Creating Self-Regulated Learners: Strategies to Strengthen Students' Self-Awareness and Learning Skills*. Stylus Publishing, LLC.

McGuire, S. Y. (2018). *Teach Yourself How to Learn: Strategies You Can Use to Ace Any Course at Any Level*. Stylus Publishing, LLC.

Rae, A. (1993). Self-paced learning with video for undergraduates: A multimedia Keller Plan. *British Journal of Educational Technology*, *24*(1), 43–51. https://doi.org/10.1111/j.1467-8535.1993.tb00640.x