

How well do engineering students retain core mathematical knowledge after a series of high threshold online mathematics tests?

MICHAEL CARR^{1,*}, MARK PRENDERGAST², CORMAC BREEN¹ AND FIONA FAULKNER¹
¹Dublin Institute of Technology, Bolton Street, Dublin 1, Dublin, Ireland and ²Trinity College, Dublin 2, Ireland

*Corresponding author. Email: michael.carr@dit.ie

[Submitted September 2015; accepted August 2016]

In the Dublin Institute of Technology, high threshold core skills assessments are run in mathematics for third-year engineering students. Such tests require students to reach a threshold of 90% on a multiple choice test based on a randomized question bank. The material covered by the test consists of the more important aspects of undergraduate engineering mathematics covered in the first 2 years of the Honours degree programme and the 3 years of the Ordinary degree programme. Students are allowed to resit the assessment as frequently as required until they pass. In order to measure the effectiveness of such an exercise, a follow-up assessment was given to students on their first day of their fourth year. A comparison is made with the level of basic mathematical knowledge of these students on their first day in third year, exactly a year previously. For the majority of the students we see a significant decrease in the performance of the students from the beginning of third year to the beginning of fourth year. In addition, students were surveyed for their perception of both how much knowledge had been retained and how effective they felt that this approach had been. Overall the students felt positive about the process of online testing and that it would make it easier for them to regain this information in the future.

1. Introduction

Dublin Institute of Technology (DIT) offers students several different routes into engineering, via direct entry onto an Honours degree (Level 8) or alternatively entry into an Ordinary degree (Level 7) followed by subsequent transfer to third year of the Honours degree (see Llorens *et al.*, 2015 for detail on this system). However, these students (both transfer and direct entry students) have a wide range of mathematical abilities and prior knowledge, and many who transfer across from the Ordinary degree are missing the basic skills required for completion of a mathematics module at this level. The third-year mathematics module includes material on eigenvalues and eigenvectors and their applications to engineering, numerical solution of differential equations (Runge–Kutta, finite difference, etc.), Fourier analysis and vector calculus. Five disciplines of engineering sit mathematics in third year but only

three (mechanical, structural and civil) disciplines sit mathematics in fourth year. For the purposes of this study we will only compare students who sit mathematics in both third and fourth years.

Assessment is a critical catalyst for student learning and there is considerable pressure on higher education institutions to provide more frequent assessment for students (Conole & Warburton, 2005). Due to the dominant impact of technology, the assessment process is gradually shifting from the traditional pen and paper method to computer based and e-assessment (Dube *et al.*, 2009). Such forms of assessment have provided an opportunity to reconsider both the purposes and practices of assessment (Sangwin, 2013). They have become standard practice in many university mathematics departments, and their use is likely to increase significantly over the next decade (Sangwin, 2013). This is in part due to academics having access to open-source algorithmic e-assessment systems, such as STACK (Sangwin, 2004), Numbas (Foster *et al.*, 2012) and DEWIS (Gwynllyw & Henderson, 2009) and also due to the many advantages that computer based and e-assessment affords (Walker *et al.*, 2015). Perhaps, the most valuable advantage is the ability to generate random questions and provide focused and timely feedback to students and staff (Sangwin, 2013). A study conducted by Muwanga-Zake (2006) evaluated the diagnostic value of computer-aided assessment in the process of teaching and learning. Results showed that learners improved due to self-assessment and instantaneous marking which enabled them to diagnose and remedy mistakes themselves. However, one of the difficulties in computer based and e-assessments has been the difficulty in designing a good network of questions (Sangwin, 2013). Previous systems forced the question author to become a computer programmer, in effect writing substantial pieces of computer algebra code to generate the required response processing commands for each question (Sangwin, 2013). However systems such as STACK have made it somewhat easier for teachers to write their own questions and to make personalized mathematics exercise assignments for students (Rasila *et al.*, 2015). The STACK system has currently been adapted for the requirements of engineering mathematics courses at Aalto University (Sangwin, 2013).

High threshold ‘technique-mastering courses’ have been used extensively but mainly with first-year students (Croft *et al.*, 2001; du Preez, 2004; Marjoram *et al.*, 2013; Carr *et al.*, 2013b).

In an attempt to resolve our problem, an Advanced Maths Diagnostic test for third-year students was introduced in 2011. This test covered many of the key concepts from the early years of Engineering Mathematics and is given to the students on their first day of third year. A pass mark of 90% was set in this assessment. Multiple resit opportunities were provided, and a weighting of 10% of the continuous assessment mark for the mathematics module was given to the diagnostic test. The test consists of 10 paired multiple choice questions. Using multiple choice questions it is important that there are at least two questions of each type. This makes it less likely that students are not getting questions correct just by guessing, i.e. if there are paired questions with four multiple choice answers on a topic then there is only a one-in-sixteen chance of someone guessing both questions correctly as opposed to a one-in-four chance for a single question (Lee & Robinson, 2005) with a combination of algebraic and numeric answers (see Table 1).

Students have the opportunity to sit this test once a week. All students are given an initial test multiple choice on paper (+3 for a correct answer, -1 incorrect, 0 no answer). This test was used as our baseline result as all of the students sat the test on the same day under the same conditions. All subsequent repeats were done on WebCT (Web Course Tools a software product which enables online course management) using a multiple choice format under supervised conditions, using six different versions of each question. Students were allowed to use a calculator for both versions of the test. Students receive immediate feedback and may discuss difficult questions immediately after the test with a tutor. Online resources and special classes were also introduced. Some students sat the test up to six times before passing. On average students sat the test three–four times. This procedure has been run

TABLE 1. List of topics covered in the advanced maths diagnostic test (Carr et al., 2013b)

Topic	Sub-topic	No. of questions
Differentiation	Basic	2
	Product rule	2
	Quotient rule	2
	Chain rule	2
Integration	Basic	2
	Substitution	2
Differential equations	First-order ODEs	2
	Second-order ODEs	2
Matrices	Multiplication	2
Complex Numbers	Multiplication	2

for the last few years and normally 80–90% of the students achieve a mark of over 90% in these basic skills. Those who fail to pass the test get 0 out of 10 marks for this part of the module. Our experience has been that the majority of students who fail to pass this component also fail the other components of the module and end up resitting the module.

In an attempt to assess the level of retention of these basic skills the Advanced Diagnostic test is also given to the same students on the very first day of fourth-year maths (third week of September). For comparison this test was also paper based but with blank spaces. This allows for comparison with the performance in third year but also allows us to gain an insight into what particular mistakes students are making. Not all students, depending on their particular discipline of engineering, study mathematics in fourth year so we only do the comparison for those who study mathematics in third and fourth years. These students would not have studied any mathematics since the previous January, a break of 9 months, although in theory they should have been using these skills in their other engineering modules. For most students, this would be the longest break they would ever have had from studying mathematics.

Sample questions

1. Find

$$\int x^3 dx$$

- a) $\frac{x^4}{4}$ b) $4x^4$ c) $3x^3$ d) x^3

2. Evaluate the following integral

$$\int \cos(x+2) dx$$

- a) $-\sin(x+2)$ b) $\sin(x+2)$ c) $2\cos(x+2)$ d) $-2\cos(x)$

3. Find the solution of the following second-order differential equation

$$\frac{d^2y}{dx^2} + 4\frac{dy}{dx} + 3y = 0$$

- a) $Ae^{-3x} + Be^x$ b) $Ae^{-3x} + Be^{-x}$ c) $Ae^{3x} + Be^x$ d) $Ae^x + Be^{3x}$

The complete set of questions used is available in Carr et al. (2013b).

2. Methodology

The methodology for this study involved comparing the results of a paper-based baseline diagnostic test taken by a group of students on their first day in third year and again on their first day in fourth year (in both cases this was given as a pure diagnostic test with no warning and the third-year students were only told afterwards that they would have to keep doing it until they achieved 90%. Some of the students may have been vaguely aware that such a test existed but we are reasonably confident that they would not have done any preparation for such a test on their first day of semester). A questionnaire was also given to the students once they have completed the fourth-year diagnostic test. A total of 83 students took the paper-based diagnostic test in third year, 69 in fourth year and 41 students completed the questionnaire (the day we handed out the questionnaire clashed with project presentations for some students so the 41 are a representative sample). A total of 42 students took the test in both third and fourth years. The questionnaire was designed by the authors and was made up of five questions with the aim of investigating students' opinions regarding the process of continuous diagnostic testing. Following the results of the questionnaire, a focus group was carried out to delve deeper into some of the issues raised. This focus group consisted of four students in fourth year who had undertaken the diagnostic test in both the third and fourth years of their studies. The results have been made anonymous and the participants are denoted by S1–S4.

The diagnostic test is made up of 20 questions based on various mathematical topics. These topics include integration (four questions), differentiation (eight questions), first-order differential equations (two questions), second-order differential equations (two questions), complex numbers (two questions) and matrices (two questions). The test is presented to students in their first mathematics lectures of third and fourth years. This was considered the optimal time to give the test as the majority of students would turn up on the first day of term. In third year students sat a multiple choice test with 3 for a correct answer, 0 for no answer and -1 for an incorrect answer. In fourth year students sat the test with blank spaces for the answer and were awarded 1 mark for a correct answer and 0 for an incorrect answer. No method marks were awarded. By using open answers in the fourth year test we hoped to gain an insight into where the students were going wrong on the more complicated tasks. An analysis of these mistakes is presented in Section 3.4. In hindsight this was not the perfect design as it may somewhat complicate the comparison between years.

3. Diagnostic test quantitative results

Tables 2 and 3 show a breakdown of the students' performance in each question over both years

Comparison of results of all third- and fourth-year students (paired and unpaired):

When we compare the results for the third-year average (46.7%) and the fourth-year average (45.7%) across all students we see a very small but not statistically significant decrease (see Table 2).

Comparison of results of paired third- and fourth-year students ($n = 42$):

If we only compare the results for the 42 students who sat the test on the first day of third year and the first day of fourth year we find the third-year average (51.9%) is significantly larger than the fourth-year average (44.3%).

Using a one-sided paired t -test ($df = 41$, i.e. *we compare each of the 42 students' results from third and fourth years*) this difference is statistically significant at the 5% significance level ($p = 0.013$).

Some of this difference may be attributed to the difference between using constructed-response items in fourth year as opposed to stem-equivalent multiple choice items in third year (Bridgeman, 1992).

TABLE 2. Comparison of mean results per question across all students

Question	Question type	Third-year % mean	Fourth-year % mean	Diff (+/-)
1	Differentiation	100	100	No Diff
2	Differentiation	100	100	No Diff
3	Second-order ODEs	6	8	+2
4	Second-order ODEs	4	0	-4
5	Integration	90	94	+4
6	Integration	89	88	-1
7	Matrices	47	14	-33
8	Matrices	68	8	-60
9	First-order ODEs	16	44	+28
10	First-order ODEs	16	44	+28
11	Chain rule	56	48	-8
12	Chain rule	56	48	-8
13	Product rule	65	48	-17
14	Product rule	32	29	-3
15	Quotient rule	14	61	+47
16	Quotient rule	60	61	+1
17	Complex numbers	36	14	-22
18	Complex numbers	36	9	-27
19	Integration	51	48	-3
20	Integration	-8	48	+56
	Mean difference			-1

TABLE 3. Comparison of mean results per question across 42 paired students

Question	Question type	Third-year % mean	Fourth-year % mean	Diff (+/-)
1	Differentiation	100	100	No Diff
2	Differentiation	100	100	No Diff
3	Second-order ODEs	8	5	-3
4	Second-order ODEs	8	0	-8
5	Integration	94	93	-1
6	Integration	97	90	-7
7	Matrices	56	17	-39
8	Matrices	79	10	-69
9	First-order ODEs	17	36	19
10	First-order ODEs	7	38	31
11	Chain rule	51	40	-11
12	Chain rule	67	38	-29
13	Product rule	76	40	-36
14	Product rule	34	31	-3
15	Quotient rule	25	71	46
16	Quotient rule	74	69	-5
17	Complex numbers	37	7	-30
18	Complex numbers	48	10	-38
19	Integration	63	45	-18
20	Integration	-3	45	48
	Mean difference			-7.6

The main differences in scores between each year were in the second-order ordinary differential equations (ODE), matrices and complex number questions. These results are particularly interesting when they are compared with students' responses to the questionnaire on the diagnostic testing process shown in Figs 1 and 2. Although the results from both tests show a decline in performance from Year 3 to Year 4, the majority of students (71.5%) felt it will be either somewhat easier or easier to restore knowledge as a result of engagement in the continuous diagnostic testing process.

A closer look at each specific topic (Table 1) demonstrates further discrepancies between the quantitative results and students' perceptions of their results. Students were asked to reflect on the diagnostic test they took in third and fourth years and to indicate their level of retention in each of the topics. The two topics which students indicated the highest levels of retention were complex numbers (69.1% somewhat/very high) and matrices (66.7% somewhat/very high). However, these were the topics in which there was the biggest drop in student performance between the third- and fourth-year tests. Of the four questions contained in these two topics there was an average drop of 45 percentage points from the third- to the fourth-year tests. A large number of students (47.6%) were undecided on their level of retention for the differential equation questions. The comparisons of test results for these questions show that students did show good levels of retention for the first-order differential equations but poor retention for the second-order differential equations.

Overall the results of the questionnaire indicate that students have a positive opinion of the diagnostic testing process. As Fig. 2 depicts, the majority of students feel that the continuous practising of skills through the diagnostic testing process has helped their performance in end of term examinations.

Furthermore, 66.7% of students would like the diagnostic tests to be carried out more frequently. Although the pass mark for the diagnostic test is very high (90%), about half (47.6%) of students feel that this mark is just right whereas the other half (50%) feel that it is too high.

To see if the test is suitable for the ability level of our students we carried out a Rasch analysis on the test, i.e. if the mean difficulty level of the questions and the mean ability level of the students are well aligned.



FIG. 1. Fourth-year students' responses on whether it will be easier to restore knowledge from diagnostic testing.

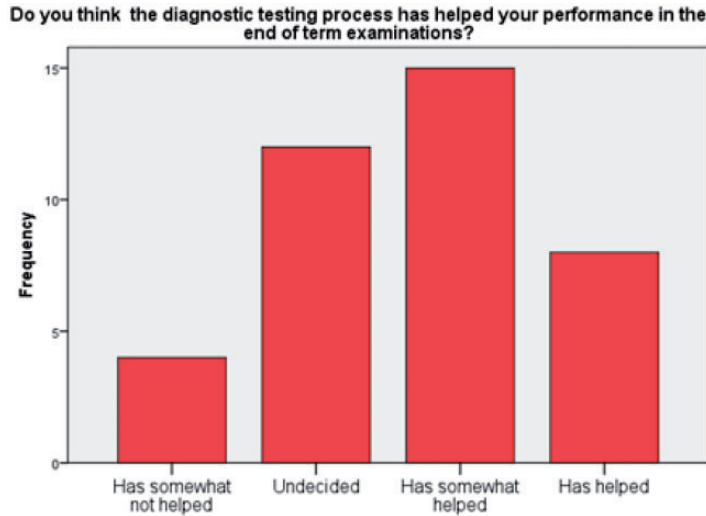


FIG. 2. Students' responses on whether diagnostic testing has helped exam performance.

3.1 Rasch analysis

The Rasch model is an item response model (IRM), which ranks items (e.g. questions in an exam) by difficulty level and student by 'ability' level. IRMs assume that when a student encounters a question in an exam, they have a fixed probability of answering that particular question correctly. The Rasch model makes the further assumption that the probability of getting a question correct is a function of the student's 'ability' and the difficulty of the question alone (and so does not allow for student mistakes or for 'lucky guesses'). Thus this analysis is not appropriate for the multiple-choice question (MCQ) test given in third year but we can apply this technique to the fourth-year results. It is also an assumption of the Rasch model that the test seeks to examine a single facet of a student's understanding, labelled as 'ability'. In the context of this study, we take this 'ability' to be a student's overall mathematical ability. Finally the model assumes that the probability of answering one question does not depend on the answers to other questions (e.g. questions that are structured into parts), which was the case for the diagnostic test used in this study.

Given a particular question and a person, $\{q,p\}$, the dichotomous Rasch model gives the probability that student p answers question q correctly as

$$P = \frac{e^{a_p - d_q}}{1 + e^{a_p - d_q}},$$

where a_p denotes the ability of person p and d_q denotes the difficulty of question q . As the Rasch model allows one to estimate the student ability parameters and the question difficulty parameters on the same unitless scale, one can directly compare these two sets of parameters. It is immediately clear that if the ability of a student matches the difficulty level of a question ($a_p - d_q = 0$) then the probability that the student will answer the question correctly is 0.5. Rasch analysis produces estimates of the parameters a_p and d_q from a set of test results and thereby estimates the probability that a student of ability a will answer a question of difficulty d correctly. For this study we made use of the computer package Winsteps (Linacre, 2016) to perform the Rasch analysis. Winsteps makes use of the joint maximum-likelihood estimation procedure when estimating the parameters.

This allows one to gain insight into the suitability of a set of questions to a particular group of students' ability levels. The results of this analysis for the diagnostic test are represented in Fig. 3. The upper histograms represent the distribution of the student ability parameter estimates and the lower ones the distribution of the question difficulty parameter estimates.

For the fourth-year results the means of the person ability parameter estimates, $-x_p = -0.54$, and of the item difficulty parameter estimates, $-x_q = -0.4$, lay well within one ability parameter standard deviation ($s_p = 1.42$) of each other. In fact 55% of the questions difficulty parameter estimates lay within 1 SD of the mean of the students' ability parameter ($[-1.96, 0.88]$). Therefore it is reasonable to conclude that this diagnostic test was well suited to the range of mathematical ability levels, as measured by this analysis, present in this cohort.

3.2 Analysis of common mistakes

Using the open-ended maths test given to the students in their fourth year we were able to do an analysis of where the students were going wrong. All of the 69 students attempted the test. There are two easy questions on basic differentiation at the beginning of the test and all of the students answered these two questions corrected both in third year and fourth year (see Table 2). In this section we present an analysis for the three question topics that are most difficult: namely second-order differential equations, first-order differential equations and integration by substitution. As these tasks were the most complex we present the information on where they are going wrong.

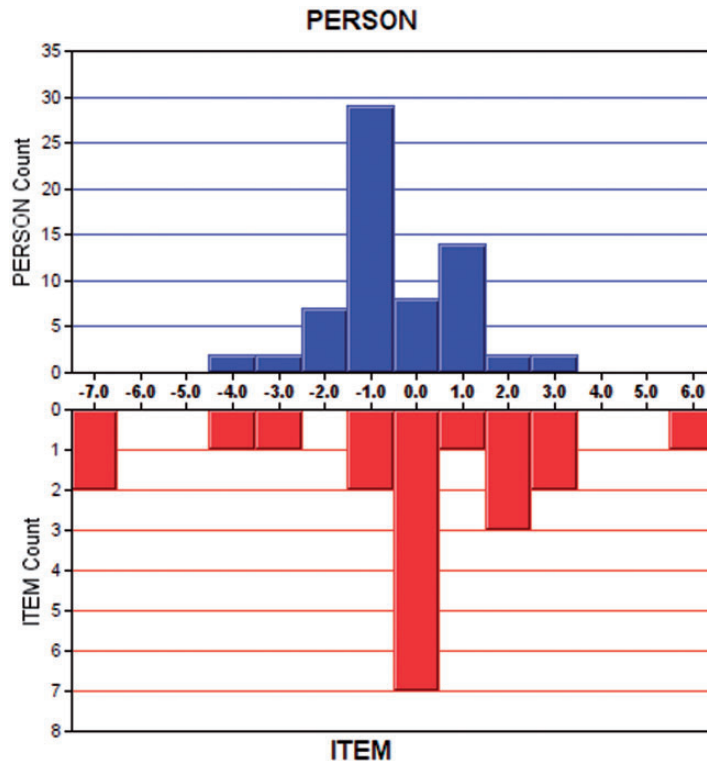


FIG. 3. Person–item histograms for the fourth-year diagnostic test results.

Second-order differential equations (Q1 and Q2)

Q1: Solve $\frac{d^2y}{dx^2} + \frac{4dy}{dx} + 3y = 0$

Q2: Solve $\frac{d^2y}{dx^2} + \frac{6dy}{dx} + 10y = 0$

A total of 69 tests were analysed.

Question 1

There were five correct responses for Question 1. Forty tests were left completely blank. Sixteen responses had resolved the roots of the auxiliary equation but could go no further. Four responses had an auxiliary equation only and the rest of the responses had not made any sensible progress.

Question 2

There were two correct responses for Question 2. Forty-seven tests were left completely blank. Two respondents had resolved the roots of the auxiliary equation but could go no further. Nine respondents had an auxiliary equation only and the rest of the responses had not made any sensible progress.

Question 2 is a slightly more difficult question needing complex roots, but overall fairly comparable to Question 1 in terms of the number of correct answers.

First-order differential equations (Q9 and Q10)

Q9: Solve $\frac{dy}{dx} = \frac{2x}{y+1}$

Q10: Solve $\frac{dy}{dx} = (1+x)(1+y)$

There were nine correct responses to Question 9 (five for Question 10). There were 38 blank responses (36 for Question 10). Thirteen respondents tried to use the quotient rule to differentiate this. Similarly 14 respondents tried to use the product rule to differentiate Question 10.

Integration by substitution

Q19: Evaluate the following integral $\int \cos(x+2)dx$

Q20: Evaluate the following integral $\int x(4x^2 - 7)^3 dx$

There were nine correct responses to Question 19 (six for Question 20). Forty-three students left Question 19 blank (49 left Question 20 blank). Nine students integrated the $(x+2)$ inside the brackets and ignored the cosine in Question 19. Four students integrated each part of Question 20 separately and multiplied the answer together.

4. Qualitative results

In addition to the quantitative analysis above we surveyed students with open-ended questions, i.e. questions that allow the students to insert open responses. The full questionnaire is in the Appendix.

These responses were then analysed and the main themes were extracted. After analysis of the qualitative responses a focus group was held to complement the results of the opened-ended survey.

4.1 Qualitative synopsis of students' responses (n = 41)

Upon analysis of the qualitative data contained in the questionnaire the following findings emerged.

4.1.1 Key points

- 25% of respondents mentioned specific topics of difficulty, the majority of these highlighting ODEs as being an issue.
- 44% of students who expanded on this question referred to ‘poor retention levels’ due to ‘forgetting’ material or ‘needing revision’ but did not mention specific topics.
- 12% of students felt it encouraged them to revise/improve their skills.
- 60% highlighted the ability of the diagnostic test to aid memory and retention of topics. One student saying it helps ‘reinforce existing knowledge’ (Male, Level 7).
- 36% focussed on the ability of the diagnostic test to aid real understanding of the mathematical topics contained on the test by encouraging real engagement and study of material after the test was completed.
- 50% of students stated that they were addressing areas of weakness in the diagnostic test.
- 50% of students maintained that there was ‘not enough time in the day’ (Female, Level 8) to address the areas of weakness highlighted by the test. ‘I have to focus on other areas of study’ (Male, Level 8); ‘I can’t as I have too much of a work load’ (Male, Level 8).
- Two-thirds of the students who responded to this question were positive towards the diagnostic test and its benefits in preparing them for end of term examinations—‘It keeps you in study mode’ (Male, Level 8, Mature Student), whereas 26% of respondents could not see the benefit of the diagnostic testing in preparing them for their end of term examination.

4.1.2 *How could the diagnostic test be improved?* Students’ opinions in relation to how the diagnostic test could be improved centred on the following four themes, starting with the one which was most prevalent:

- (1) 32% of students who responded suggested that the pass mark should be lowered from 90%—‘Maybe reduce the 90% pass rate slightly’ (Female, Level 7).
- (2) 16% of students who responded had an issue with the negative marking and suggested that this be removed from the testing.
- (3) Another 16% requested that more tests be carried out as they found them so helpful.
- (4) 11% of students felt that the correct answer or some form of feedback should be provided when they got questions wrong.
- (5) 5% said the test was fine as it is

In total, 48% of respondents did not provide any suggestions for change at all.

4.2 Focus group

A focus group was held with four volunteer students who had sat the exams in third and fourth years and who had entered DIT via the ordinary degree programme and subsequently transferred to the honours degree. Previous research (Carr *et al.*, 2013a) has shown that this group is the most likely to have weaknesses in their core competencies.

4.2.1 *Do you feel that the core maths regime in third-year maths helped improve your basic maths and/or prepare you for third-year engineering?* All respondents agreed that the regime in third year helped improve basic maths skills. The main benefit in improving basic maths skills

mentioned by all focus group respondents with the exception of one was refreshing basic skills that they did in secondary school and students felt the forum it was completed in was appropriate as if it was carried out in class 'it probably would have wasted time in the class for people who knew it already' (FG; S1).

4.2.2 In our class survey many students state that they do not have enough time to engage in covering basic maths or revise (addressing the areas of weaknesses?) The majority of respondents (all with the exception of one) felt that not having enough time to revise was not a legitimate excuse due to the mathematics support facilities that exist in the college, the approachable helpful nature of the mathematics lecturing staff and the fact that it should be made a priority at this stage in your undergraduate studies; 'You're in third year, you've no excuse' (FG; S3).

4.2.3 How did you feel about the diagnostic testing in general? As demonstrated in the questionnaire the majority of students were happy with the questionnaire although the issue of negative marking 'being a bit harsh' was raised however the same student felt 'you had plenty of chances so you can't really complain' (FG; S1). Students felt over all it was an excellent means of revising skills that you sometimes forgot but should know.

4.2.4 Is there any way the diagnostic testing could be improved? Students felt that the questions in the diagnostic test were not always applicable to their area of engineering but in general the response to this question was positive with little room for improvements and some students even suggesting that it would be beneficially to complete diagnostic testing like this in both semesters of third and fourth years 'to help you along' (FG; S4).

4.2.5 Do you feel your basics were better going into fourth-year engineering than they were going into third-year engineering? Why do you think this was the case? All students felt that their basic maths skills were better going into fourth year than going into third year. They felt the diagnostic test 'forced them to find the time' to revise things in third year that they knew they needed to know but never got around to studying. All students being interviewed who took Ordinary Leaving Certificate maths at second level education in Ireland had never studied integration and mentioned that the test 'forced me to study it' which was good (FG; S3 & FG; S1).

4.2.6 The class average in fourth year was significantly worse that it was at the beginning of third year, why do you think this is the case? All interviewees were surprised by this finding and could not really come to any conclusion as to why this might have been the case other than maybe the first day of term when you are only back 'so you literally don't have a clue' (FG; S2).

4.2.7 The biggest decrease in scores occurred in the complex numbers, second-order ODEs and matrices questions, why was there such a decline in these subjects in particular? Again all of those being interviewed were surprised by the decreases in these areas

with one student expanding on his lack of understand of the decline in performance in matrices in particular due to its ‘methodical’ nature (FG; S1).

4.2.8 Do you believe that the lack of any mathematics modules in Semester 2 had an influence on the results of the diagnostic tests? All students felt that the absence of mathematics modules for an entire semester had been a contributing factor to their decreased performance in the diagnostic test.

4.2.9 Last word on the diagnostic test ... Students were positive about the testing regime helping them remember/revise previous material from an early stage in the year, guaranteed marks going into your final exam, keeping your brain working throughout the term/year along with reiterating what they would change about the test: include a larger bank of questions and possibly remove the multiple choice element and give more time to work out questions.

5. Conclusions

In summary this intervention was effective from the point of view that almost all of the students achieved 90% in an exam in this material at the end of third year (Carr *et al.*, 2013b). In addition it was popular with the students and laid a foundation for retrieving the knowledge in future. This mark of 90% is not maintained and one of the main findings of the study is that there was a significant decrease in the performance of the students from the beginning of the third year to the beginning of the fourth year ($p=0.013$) although this may be explained in part by the shift from an MCQ environment to open-ended questions. This brings forward the issue of students not remembering material over time. Students not being able to retain information such as this is particularly common for knowledge acquired in educational settings, and much of this material is lost within days or weeks of learning (Rohrer & Taylor, 2006). Thus, any learning strategy must be judged at least in part by students’ retention of knowledge after a non-trivial retention interval. For longer term retention this material needs to be ‘over practiced’ over a longer period of time ‘that people remember more mathematics and other high school material when learning occurs in sessions spaced out over several years’ (Bahrick & Hall, 1991). This is an important area of research that has the potential to inform instructional practices in all educational settings (Bahrick, 2000) and raises serious concerns about the introduction of semesterization. For over a century, psychologists have noted that temporally spaced practice leads to more robust and durable learning than massed practice (Cepeda *et al.*, 2006; Rohrer & Taylor, 2006). Furthermore a study conducted by Kwon *et al.* (2005) found that teaching for conceptual understanding can lead to longer retention of mathematical knowledge.

Overall the students felt positive about the process and felt the process had helped them in their third-year exam and the students felt that the process did make it easier for them to retain this information in the future. It does seem that such practice of basic skills only leads to short-term memory gain. In addition doing the more advanced module, i.e. the fourth-year mathematics module should help with consolidation of this more basic material (Bahrick & Hall, 1991).

Work in this area by Engelbrecht *et al.* (2007) found a similar deterioration in long-term retention of basic mathematical knowledge over a 2-year period but interestingly found ‘retention of knowledge and skills is excellent in the case of simple differentiation and integration techniques’. This is similar to what we have found (Table 2).

Achieving 90% on a multiple choice test is not being the same as submitting their answers in numerical/algebraic/text format. Paper and online assessment can differ in several respects. Conole & Warburton (2005) discuss the limitations of MCQs in detail. But item-based testing can be used to examine the ‘full range of learning outcomes in undergraduates and postgraduates, provided sufficient care is taken in their construction’. Bridgeman (1992) has shown performance on stem-equivalent multiple choice items can be easier than constructed response items, and this difference varies from question to question depending on how the questions are constructed. It seems to work very well for some types of questions, i.e. basic differentiation (Table 3) but for more complex tasks such as second-order differential equations we may need to break the process up into several different multiple choice type questions. We can see this from the analysis of the common mistakes made on this question. Almost half of the students who got the question wrong were able to find the roots of the auxiliary equation or at least write the auxiliary equation. With the crutch of multiple choice answers they may have been able to get the question correct.

What we can say definitively is that MCQ tests with a very high pass mark are useful in a certain context. Many of the students in this study had never covered this (more basic) material before, due to choice on exam papers and the 40% pass mark that is the standard in Irish education. Such MCQ tests are certainly not the answer to all our problems and should only ever form part of a module’s assessment programme (in the case under discussion the MCQ tests were worth 10% of the total assessment grade). We would recommend that the majority of (assessment in) any mathematics module should concentrate on conceptual understanding for longer and deeper understanding in line with (Kwon *et al.*, 2005).

REFERENCES

- BAHRICK, H. P. (2000) Long-term maintenance of knowledge. *Oxford Handbook of Memory* (E. Tulving & I. M. Fergus eds). England: Oxford University Press, 347–362.
- BAHRICK, H. P. & HALL, L. K. (1991) Lifetime maintenance of high school mathematics content. *J. Exp. Psychol.*, 120, 20–33.
- BRIDGEMAN, B. (1992) A comparison of quantitative questions in open-ended and multiple-choice formats. *J. Educ. Meas.*, 29, 253–271.
- CARR, M., NI FHLOINN, E. & BOWE, B. (2013a) Core skills assessment to improve mathematical competency. *Eur. J. Eng. Educ.*, 38(6), 608–619.
- CARR, M., NI FHLOINN, E., MURPHY, E. & BOWE, B. (2013b) Addressing continuing mathematical deficiencies with advanced mathematical diagnostic testing. *Teach. Math. Appl.*, 32, 66–75.
- CEPEDA, N. J., PASHLER, H., VUL, E., WIXTED, J. T. & ROHRER, D. (2006) Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychol. Bull.*, 132, 354–380.
- CONOLE, G. & WARBURTON, W. (2005) A review of computer-assisted assessment. *Alt-J*, 13, 19–33.
- CROFT, A. C., DANSON, M., DAWSON, B. R. & WARD, J. P. (2001) Experiences of using computer assisted assessment in engineering mathematics. *Comput. Educ.*, 37, 53–66.
- DUBE, T., ZHAO, Z. & MA, M. (2009) *E-assessment and Design Methodology Management*. Paper Presented at E-Assessment Live 2009 Conference, Loughborough University, UK, 8 July 2009.
- DU PREEZ, A. E. (2004) Format and long-term effect of a technique mastering programme in first year calculus. *Doctoral Dissertation*, University of Pretoria.
- ENGELBRECHT, J., HARDING, A. & DU PREEZ, J. (2007) Long-term retention of basic mathematical knowledge and skills with engineering students. *Eur. J. Eng. Educ.*, 32, 735–744.
- FOSTER, B., PERFECT, C. & YOUNG, A. (2012) A completely client-side approach to e-assessment and e-learning of mathematics and statistics. *Int. J. E-Assessment*, 2, 1–12.

- GWYNLLYW, R. & HENDERSON, K. (2009) DEWIS: a computer aided assessment system for mathematics and statistics. *CETL-MSOR 2008 Conference Proceedings*, The Maths, Stats and OR Network Lancaster, UK. 8th-9th September 2008, David Green (ed) https://www.heacademy.ac.uk/sites/default/files/CETLM_SOR2008_Proceedings.pdf
- KWON, O. N., RASMUSSEN, C. & ALLEN, K. (2005) Students' retention of mathematical knowledge and skills in differential equations. *Sch. Sci. Math.*, 105, 1–13.
- LEE, S. & ROBINSON, C. L. (2005) Diagnostic testing in mathematics: paired questions. *Teach. Math. Appl.*, 24(4), 154–166.
- LINACRE, J.M. (2016) Winsteps® Rasch measurement computer program. Beaverton, Oregon: Winsteps.com.
- LORENS, M., O'SHAUGHNESSY, S., CARR, M., SHERIDAN, D. & BOWE, B. (2015) Ca ira: the need for reform of Engineering education. 43rd Annual SEFI Conference June 29 -July 2, 2015 Orléans, France <http://www.sefi.be/conference-2015/CHAP%202.%20Engineering%20Education%20Research/57160-%20D.%20SHERIDAN.pdf>.
- MARJORAM, M., ROBINSON, P., O'SULLIVAN, C. & CARR, M. (2013) Improving the key skills of engineering students. *Mathematics 41st SEFI Conference, 16–20 September 2013, Leuven, Belgium*.
- MUWANGA-ZAKE, J. (2006) Applications of computer-aided assessment in the diagnosis of science learning & teaching. *Int. J. Educ. Dev. Using ICT*, 2(4).
- RASILA, A., MALINEN, J. & TIITU, H. (2015) Automatic assessment and conceptual understanding. *Teach. Math. Appl.*, 34, 149–159.
- ROHRE, D. & TAYLOR, K. (2006) The effects of overlearning and distributed practice on the retention of mathematics knowledge. *Appl. Cogn. Psychol.*, 20, 1209–1224.
- SANGWIN, C. (2004) Assessing mathematics automatically using computer algebra and the internet. *Teach. Math. Appl.*, 23, 1–14.
- SANGWIN, C. (2013) *Computer Aided Assessment of Mathematics*. Oxford: Oxford University Press.
- WALKER, P., GWYNLLYW, R. & HENDERSON, K. (2015) Diagnosing student errors in e-Assessment questions. *Teach. Math. Appl.*, 34, 160–170.

Appendix. Qualitative questionnaire

1. Reflecting on the diagnostic test you took at the start of this academic year, what was your level of retention in each topic since the last diagnostic test you took in third year.
2. In areas where you had low levels of retention do you think it would be easier to restore your knowledge in the future as a result of your engagement in continuous diagnostic testing. Explain.
3. Are you addressing the areas where you feel you had low levels of retention?
If yes how?
Are you addressing the areas where you feel you had low levels of retention?
If not, why not?
4. Do you think your continuous practising of skills through the diagnostic testing process has helping your performance in the end of term examinations? Explain.
5. How could the diagnostic test be improved?

Dr. Michael Carr is a lecturer in Mathematics and Statistics in the College of Engineering and Built Environment in the Dublin Institute of Technology. His research interests include the first year experience of students, development of core mathematical skills in university students, and mathematics education for engineers.

Dr. Mark Prendergast is an Assistant Professor in Mathematics Education in Trinity College Dublin. He is qualified as a second level mathematics teacher from the University of Limerick in 2003 and completed his PhD in Mathematics Education, also in the University of Limerick, in 2011. His current research interests lie in promoting student interest and achievement in mathematics through effective teaching of the subject and also in examining the effect of 'Project Maths' on International students studying in Ireland

Dr. Cormac Breen is the coordinator of the Mathematics Learning Support Centre (MLSC) in the Dublin Institute of Technology (DIT). He holds a degree and a Masters in Theoretical Physics from University College Dublin (UCD) and completed his PhD in Theoretical Physics, also in UCD, in 2011. He has lectured and tutored in Mathematics in both UCD and DIT, and has been the coordinator of the MLSC since May 2013. He is also the secretary of the Irish Mathematics Learning Support Network. His current educational research interests lie in the area of transitions to third level education, diagnostic testing and the first year experience of third level students.

Dr. Fiona Faulkner is a mathematics lecturer in the Dublin Institute of Technology. She is a qualified secondary school mathematics and Physical Education teacher (BSc 2004- 2008) with a PhD in mathematics education (2008-2012) both from the University of Limerick, Ireland. Her current research interests lie in the areas of mathematics diagnostic testing, out of field teaching, mathematics education for access and international students and teaching mathematics for understanding.