Journal of Distance Learning, 2009, 13(1), 26–40

Formative Assessment: Evaluating the Effectiveness of Online Quizzes in a Core Business Finance Course

Hamish D. Anderson
Department of Economics and Finance
College of Business
Massey University
Palmerston North, New Zealand

Introduction

Action without reflection can result in unproductive endeavour, and reflection without action is unlikely to produce *any* endeavour. Feedback provided through formative assessment practices is critical to this action–reflection process and has been referred to as the "life-blood of learning" (Rountree, 1987). Feedback enables students to reflect on and monitor their own progress and then redirect their study endeavours to where they are most productive. If it is correctly structured, assessment should also inform teaching and improve learning (McTighe & Ferrara, 1998). Further, the act of reflecting on students' formative assessment should show teachers where their teaching endeavour is most effective.

Over the last decade, the World Wide Web has become firmly established in the commercial world and is increasingly prevalent in education. The potential benefits that the internet provides to distance education—of flexibility, timeliness, and logistical advantages—are clear. Developed and delivered appropriately, web-enhanced (or blended) education can also yield benefits of engendering active learning, and encourage reflection and application (Eastmond, 1998).

The results of early research into the effectiveness of computer-based learning aids in finance and accounting courses were mixed. Er and Ng (1989) concluded that computerised instruction yielded little benefit for teaching accounting concepts, while Kachelmeirer, Jones, and Keller (1992)

found that similar learning aids did help teaching of accounting courses. Marks (1998) found that introductory graduate finance students with access to computerised formative testing achieved higher summative assessment grades than those students who did not have access to the software.

Gretes and Green (2000) examined the use of computer-based assessment for student practice and feedback, and found a significant one-half letter grade difference between students who used computer-based practice tests and those who did not. The significant positive effect on student performance remained after controlling for students' earlier academic performance. Gretes and Green (2000) also found a positive relationship between the number of practice tests attempted by students and course grades. Qualitative feedback from students of online and computer-based assessment showed improved student motivation and increased incidence of revision and enhanced engagement with the subject material (Thelwell, 2000; Marriott & Lau, 2008).

According to Bloom et al's (1956) taxonomy, the lower learning objectives (such as knowledge, comprehension, and application) deal with the fundamental concepts that students need to know about a new topic, and how to place the basic information in its correct context. Mastery of these lower learning objectives allows the learner to progress onto the higher learning objectives involving analysis, synthesis, and evaluation. The formative assessment in the core Business Finance course in this case study has always sought to allow students to achieve this progression.

This article outlines the traditional pen-and-paper and online formative assessment models that have been used in the Business Finance course. The effectiveness of these formative assessment models is then analysed qualitatively and quantitatively, and a final word is provided on the administrative consequences of the change to an online formative assessment model.

Background

This article outlines a major change in the formative assessment system used in a second-year core Business Finance course at the Manawatu Campus of Massey University in New Zealand. Before 2002, the formative assessment system for the Business Finance course relied on students completing answers to regular problem sets. From 2002 onwards, mastery online quizzes, which covered specific learning outcomes, became the main formative assessment tool.

The change in formative assessment practice was driven largely by student feedback and rising administrative costs (both money and time). The Business Finance course is required or highly recommended for a number of majors—including finance and accounting—and therefore has substantial student numbers. The course is offered in distance and internal modes in Semester 1 and again in distance mode in Semester 2. Over the past 8 years the Business Finance cohort has averaged 484 students per year, split (on average) between 167 internal students and 317 distance students.

The rest of this article outlines the formative assessment system before 2002 and then details the change to mastery online quizzes. I then examine the links between the mastery quiz performance and final assessment for 135 students who completed the Semester 1, 2008 distance offering. In particular, I analyse the late adopters of the mastery quizzes and track their performance in the final examination, as well as examining the effect of mastery quizzes on the borderline and academically weaker students.

The old formative assessment system

Because the learning outcomes of Business Finance are very sequential, students need to master one topic before moving onto the next. Repetition of problems and exercises is also an important learning method in financial mathematics. As such, formative assessment is a critical learning strategy which enables students to attain a comprehension level that, in turn, enables them to successfully negotiate the higher learning outcomes of analysis, synthesis, and evaluation.

Before 2002, the main formative assessment tool for both internal and distance offerings of this course were regular problem sets for which students submitted answers by specified due dates. Because of large class sizes, limited individual feedback was written; however, full worked solutions were provided. To encourage sequential learning in a large undergraduate class, students received a small participatory mark for each submitted problem set. This mark contributed to their final grade. As long as students submitted on time and had at least attempted most of the questions, they received the participatory mark. Many students appreciated the formative assessment, as evidenced by the following comments:

> Part of the paper is to complete regular tutorials. I have found that this has been instrumental in my learning—it really does facilitate the understanding of finance (SECAT,[1] 1998).

> I really liked how I had to hand in tutorials. Kept me up to date and on top of my studies!! (SECAT, 1999).

However, for every positive comment there was at least one negative comment. This was particularly true for distance students who have additional work and personal commitments, and whose study habits are not spread evenly throughout a semester. In fact, lack of flexibility was the most quoted reason for distance students not fully using the formative problem sets. The feedback below highlights the tensions in the assessment model.

> The problem sets were helpful and helped me pinpoint any areas where I had problems. However, the process is too inflexible and most extramural students study this way for the flexibility! (SECAT, 1998).

> Although I agree that ongoing tutorial assignments during the course keeps you up to date and tests [sic] your understanding, I found little time to initiate my own learning techniques (SECAT, 2000).

A review of student and marker feedback in 2001 highlighted that the problem-set system had a number of weaknesses that reduced the effectiveness of determining whether students were achieving the course's learning outcomes. These weaknesses included the following:

- The problem sets were an inflexible learning tool for students.
- Because of limited resources, it was only possible to mark and provide individual feedback on one question per submitted problem set.
- The feedback to distance students was typically given up to 10 days after they had completed and posted their answers.
- There was evidence of plagiarism and copying of written answers amongst students.
- There are increasing numbers of test-bank solution manuals on the internet, and students were increasingly copying and pasting solutions from websites.

[1] SECAT is Massey University's independent student evaluation tool that examines course design and teaching performance.

When I analysed the 2001 distance cohort, there was no significant difference in final examination marks between students who completed all formative assessment (53.8 percent average) and those who didn't attempt any formative assessments (51.8 percent average). Further, analysis revealed virtually no relationship between a student's performance in the formative assessment and their performance in the summative assessment (the correlation was 0.05). In other words, students could not use the formative assessment to gauge whether they had achieved the course's learning outcomes.

At the same time, projected class rolls were expected to grow dramatically between 2002 and 2005. Most of the total resources expended in running the problem-set formative assessment system were variable in nature. The time spent compiling the problem sets and solutions, marking, collating, recording, and distributing had to be repeated for every offering of the Business Finance course. Further, as class rolls increased, so did most of the financial costs and the time spent administering the system.

The limitations of the problem-set system and the challenges facing the course (highlighted above) led to a rethink of the formative assessment structure. From Semester 1 in 2002, the course changed to online mastery quiz testing delivered through WebCT.

Structure of mastery quizzes

The WebCT quizzes feature allows the instructor to develop quizzes from a number of question types including multiple-choice, calculation, short-answer, matching, and paragraph. Calculated questions allow the creation of questions in which the numerical inputs are randomly generated numbers. A new set of randomly generated variables is provided each time a calculated question is presented to a student. The student completes the calculation and enters their answer directly into the online quiz. The instructor can include a variance for rounding errors in the answer, and they can also indicate the number of decimal places required in the answer. The advantage of calculation questions over multiple-choice questions is that students cannot simply guess the answer. They need to understand and apply the concepts rather than relying on the odds of guessing correctly.

WebCT allows the instructor to add a set of questions (typically on the same topic) from which they then choose how many they want to randomly select.

For example, two questions could be randomly selected from a set of 20 questions every time a student attempts the quiz. Each quiz is then made up from a number of question sets. Quiz difficulty must be controlled for fairness and to build student confidence as they progress through their studies. This control is achieved by ranking each question by level of difficulty (e.g., easy, medium, and hard) so that all questions contained in a question set cover the same topic and have the same overall level of difficulty. Each randomly generated mastery quiz therefore has a consistent level of difficulty.

The Business Finance mastery quizzes predominantly use calculation, matching, and multiple-choice questions. There are five mastery quizzes, which are drawn from a test bank of over 1400 questions (in 2002 the test bank had a total of about 800 questions). Each of the five mastery quizzes covers specific topics that are linked to the learning outcomes of the course. Each of the five mastery quizzes contains 15 questions that cover approximately 2 weeks' work (Quiz 1 covers the first 2 weeks' work etc). Students are considered to have mastered the material if they achieve 10 or more out of 15 on any quiz attempt.

Students receive a small participatory mark for each of the five quizzes they master. Unlike the problem-set system, students must now demonstrate they understand most of the material (rather than simply attempting to answer a question) before they receive the participatory marks. Students can complete each quiz as many times as they like right up to the final examination. At the end of the semester, each student's highest quiz mark for each of the five quizzes is downloaded from WebCT. If their highest mark is more than 10 out of 15 for a quiz, they are awarded the participatory mark for that particular quiz.

The value of the online quizzes and the huge question database is the ability for students to retest themselves multiple times throughout a semester, and to receive instant feedback. I encourage students to use the learning tool to identify the areas they have yet to master. This strategy targets their learning and revision most effectively. From the first time the quizzes were offered, student feedback has been overwhelmingly positive.

A small sample of the feedback is shown below:

> Mastery quizzes were fantastic, they were difficult enough to ensure you had to know the work to pass (SECAT, 2002).

> On-line tests fit well with the way I study, being able to test my knowledge and identify areas I'm deficient in, complete further study and then retest myself again (SECAT 2003).
>
> The online quizzes have been great, especially since I can do them over and over to aid my learning (SECAT, 2007).
>
> WebCT quizzes are excellent—invaluable (SECAT, 2008).

The student feedback provides qualitative evidence that the change in formative assessment helps student learning and provides increased flexibility. However, I also quantify the effect on student learning by comparing a student's mastery quiz performance with their internal assessment and final examination results each year. In the following sections I examine how the formative assessment tool affects student uptake. I then examine the performance of users and non-users of the mastery quizzes in the 2008 Semester 1 distance offering. I also analyse the benefits of the quizzes for those students who were late adopters of the mastery quizzes as well as those students who are borderline or academically weaker.

Improved student use of formative assessment tool

One of my primary goals of the mastery online quizzes was to increase the flexibility of the formative assessment tool, particularly for distance education students. To examine this issue, I analysed student uptake of formative assessment under the old tutorial problem-set system and compared it with online mastery quiz use. Three years of data from 1999 to 2001 are used for the problem-set system, and data from 2002 to 2008 is used for the mastery quizzes.

Under the old system, only 59 percent of distance students attempted all formative assessment items. Since 2002, 92 percent of students in distance learning cohorts have attempted all five quizzes, while 81 percent mastered all five quizzes at least once. Although it was expected that distance students would experience the biggest benefits of the more flexible learning tool, internal student use also improved from 71 percent to 86 percent. The significant increase in students' use of the formative assessment supports qualitative student feedback that the quizzes are a flexible learning tool.

How do users and non-users perform?

To examine the effect of the mastery quizzes on student learning, I examined the 2008 Semester 1 distance cohort. The assessment structure was broken into two internal summative assessments (assignment and test) and a final examination. The content of the internal summative assessment related to mastery quizzes 1 to 3, and the final examination covered material from the entire course. Of the 162 students enrolled, 135 students completed the final examination. I restricted my analysis to the 135 completing students to enable tracking of student progress through the semester. Of the 135 who completed the course, 15.6 percent achieved a final grade in the A range, 31.9 percent were in the B range, and 32.5 percent were in the C range. Ten percent of the class received a restricted pass,[2] and the remaining 10 percent received D or E grades.

The following analysis compares the 2008 distance cohort results for the internal test with mastery quiz 3, which covered the same learning outcomes. Table 1 presents the average test marks achieved by those who did not complete mastery quiz 3, those who used it but didn't master the material, and those who mastered the quiz (i.e., scored 10 or more out of 15 at least once). The average grade for those who didn't attempt the quiz was D. Those who attempted but didn't master achieved an average R (restricted pass) grade, and those who mastered the quiz 3 material achieved a B- average. Table 1 also highlights that both the worst (minimum mark) and best (maximum mark) summative assessment marks improve as students move from the 'Didn't attempt' category through to the 'Mastered' category.

Table 1 Student use of online mastery quizzes and student performance in the summative test

Students who:	Number of students	Average test mark	Minimum mark	Maximum mark
Didn't attempt	35	43%	20%	73%
Attempted but didn't master	23	47%	23%	77%
Attempted and mastered	77	64%	30%	93%

[2] Restricted passes count as a credit towards most Massey University degrees but do not meet pre-requisite requirements for enrolment in higher level courses. They are typically awarded for overall final marks of 47 to 50 percent.

Each year I use similar results to explain risk and return, which is a key finance concept. I re-label each category as an investment opportunity and ask students to rank and choose the best investment. When the class has chosen the investment with the highest return and lowest risk, I point out that everyone can earn that return by simply investing time in completing the mastery quizzes!

On closer inspection

Table 2 digs a little deeper to look at the 77 students who mastered quiz 3. The first column shows the highest score in any quiz 3 attempt, and the second column shows how many students achieved this as their highest score. Those who achieved 10 or 11 out of 15 achieved an average of C+, compared to a B grade for quiz marks of 12 to 14. The average for those getting 15 out of 15 for one of their quiz attempts was B+.

Table 2 Highest mark for students mastering the online quizzes, and their performance in the summative test

Highest mastery quiz mark	Number of students	Average test mark	Minimum mark	Maximum mark
10	10	57%	37%	73%
11	15	58%	27%	67%
12	15	65%	47%	87%
13	13	65%	47%	90%
14	13	67%	50%	80%
15	11	72%	50%	93%

A similar result is found when comparing students' average mastery quiz and final exam marks (see Table 3). The average mastery quiz mark for each student is calculated using their best quiz attempt mark for each of the five quizzes. The average final exam mark increases from 45 percent for those not mastering the quizzes through to 69 percent for those averaging 14 or higher for each quiz. The minimum final examination mark also improves as the average quiz mark improves. The correlation between students' average mastery quiz mark and their final examination marks is 0.54. The substantially higher correlation between the online quiz for formative assessment and the summative assessment, compared with the previous problem-set formative assessment system, is further proof that the online quizzes are more useful to students for gauging their achievement of the course's learning outcomes.

Table 3 Average score across all mastery quizzes, and performance in the final examination

Average mastery quiz mark	Number of students	Average final examination mark	Minimum mark	Maximum mark
Below 10	25	45%	30%	63%
10 to 11	22	50%	34%	75%
11 to 12	37	55%	24%	83%
12 to 13	28	62%	43%	87%
13 to 14	12	64%	52%	89%
14 to 15	11	69%	48%	83%

Participation in the quizzes, mastering them, and attaining higher marks, all point to better student performance in summative assessment. However, from the above data and discussion it is not possible to determine whether the mastery quizzes lead to higher test grades, or whether the average quiz results simply reflect the level of understanding of the material. In other words, the academically better students may simply be more likely to complete and master the quizzes. In order to focus more closely on this issue, I first examined the academically weaker students. I then examined the students who used the mastery quizzes before the test and tracked their performance through to the final examination.

Dear Sir, I'm struggling. What can I do to improve my chances of passing?

I analysed the borderline and weaker students in the class using 43 students who achieved 50 percent or less in the internal summative assessment. By the end of the semester, 28 of these students had mastered the online quizzes (see 'Users' in Table 4). The other 15 students did not use or master the quizzes throughout the entire course (see 'Never mastered' in Table 4).

Interestingly, Panel A of Table 4 below shows that those who 'Never mastered' the quizzes performed marginally better in internal assessment than the 'Users' (44.1 percent compared to 41.4 percent average). However, the 'Users' turned this around in the final examination by significantly outperforming the 'Never mastered' group by 8 percent. As shown in Table 4, the improvement in the internal summative assessment performance for the

'Users' is significant (at the 1 percent level using both the parametric t-test and non-parametric Wilcoxon Rank test).

Table 4 The effect of online mastery quiz use and the performance of borderline and academically weaker students

Panel A	Sample size	Internal assessment	Final examination
Users	28	41.4%	50.7%
Never mastered	15	44.1%	42.7%
Difference		-2.7%	8.0%
t-statistic		-1.80*	2.51***
Wilcoxon Rank Test		-1.51	2.63***
Panel B	**Sample size**	**Internal assessment**	**Final examination**
Late starters	16	38.5%	49.6%
Never mastered	15	44.1%	42.7%
Difference		-5.6%	6.9%
t-statistic		-2.19*	2.31**
Wilcoxon Rank Test		-2.41**	2.38**

*** statistically significant at the 1% level
** statistically significant at the 5% level
* statistically significant at the 10% level

However, some of the 'Users' had mastered the relevant quizzes before they completed the course's internal assessment, and these students could bias the findings. I therefore excluded the early users from the sample and ran the analysis again. The second analysis left 16 'Late starters' who had not mastered quizzes 1 to 3 before they completed their internal assessment but who went on to master all of the quizzes in the 4 weeks leading up to the final examination.

The average internal assessment mark of 38.5 percent for the 'Late starters' is now significantly worse than the 'Never mastered' group who achieved 44.1 percent. So while the 'Never mastered' were academically weak students, the 'Late starters' were significantly worse before they completed the mastery quizzes. However, the 'Late starters' managed to turn their performance around in the final examination by significantly outperforming the 'Never mastered' group by 6.9 percent. Put another way, the average 'Never

mastered' student started, continued, and finished the Business Finance course as a D-grade student. In contrast, the average 'Late starter' converted their E from internal assessment to a C grade in their final examination.

Therefore—at least for the borderline and weaker students—using the formative assessment improves performance in the summative assessment. Further, using the mastery quizzes is better late than never for these borderline and academically weaker students. So, to answer the question "What can I do to help my chances of passing?" I say, "Use the mastery quizzes!"

Step right up. Have I got a deal for you!

I use summary statistics similar to those in Table 1 to shamelessly promote the mastery quizzes as a self-help learning tool. The promotion appears to help, with 20 of 38 students who did not use the mastery quizzes before the first internal assessment then going on to use and master the online quizzes before the final examination. The first internal assessment averages for these two groups are virtually identical (56.5 percent for those who 'Never mastered' and 57.2 percent for the 20 who mastered after the assessment due date). However, the average final examination marks for those who decided to use the mastery quizzes was 53.7 percent compared with only 42.7 percent for those who didn't master the quizzes. While there will be some crossover between this group and the 'Late starters' subgroup discussed previously, this result provides further evidence that student use has a positive effect on student performance in summative assessment.

Even if the mastery quizzes simply reflect—rather than improve—the level of students' understanding of the material, the quizzes are clearly a very powerful tool for students to assess their overall understanding of the course's learning outcomes.

What do I get out of it?

The mastery quizzes provide me with a wealth of summary statistics which highlight the areas and questions that students find difficult. This is the single most important feature of the quizzes for me, because I can pinpoint exactly which topics and concepts students struggle with. I can then target my teaching where it is most beneficial. Students are often surprised at my ability to target my teaching—as though I have a higher power!

> Hamish pitched the sessions at exactly the right level for the class—he seemed to know which areas and formulas would cause the most concern (SECAT, 2005).

From an administrative point of view, the savings are considerable. After the initial investment of time in creating the test bank of questions, I now benefit from the way the system runs, marks, provides feedback, and automatically records progress. All of these activities were very labour intensive under the problem-set formative assessment system.

The mastery quizzes had a large fixed administrative cost component of both time and resources, whereas the problem-set administrative costs were largely variable in nature. For the problem-set system, a normal semester offering would involve approximately 24 hours to prepare questions and solutions sets, and 84 hours of marking, writing limited individual feedback, and recording. The new system takes approximately 1 hour to reset and update each offering. Therefore, the savings have been over 100 hours per semester, or 1400 hours between 2002 and 2008! However, I do set aside time at the beginning of each year to write and add adaptations of the previous year's summative assessment questions to the mastery quiz test bank.

As a final note, I have been able to share the mastery quizzes and large test bank with other teachers within my department. The Semester 2 distance offering is typically taught by another teacher, and I have another colleague who teaches the Wellington campus offering. These teachers appreciate the personal savings in administration costs.

Conclusion

The mastery quizzes are a flexible on-demand learning tool that provides instant feedback to students. The shift to mastery quizzes has been a success in terms of qualitative student feedback, effect on student performance, and reduction in variable administration costs. In particular, the quizzes more clearly demonstrate students' level of understanding of a particular area, and enable them to target their study more effectively.

The online quiz delivery system is suitable for many types of courses, and is particularly relevant for those courses that have sequential learning outcomes. Students gain significant learning benefits, and teachers and education providers gain benefits in terms of time and financial resources. However, these benefits must be weighed against the upfront cost of building

the initial test bank. The size of the test bank is important if we want students to be able to retest themselves multiple times in a meaningful way. However, if you ask me whether I would embark on the change in formative assessment all over again, the answer would simply be: "Absolutely".

References

Bloom, B. S., Engelhard, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. New York: David McKay.

Eastmond, D. V. (1998). Adult learners and internet-based distance education. *New Directions for Adult and Continuing Education, 78*, 33–41.

Er, M. C., & Ng A. C. (1989). The use of computers in accountancy courses: A new perspective. *Accounting and Business Research, 19*, 319–326.

Gretes, J. A., & Green, M. (2000). Improving undergraduate learning with computer-assisted assessment. *Journal of Research on Computing in Education, 33*(1), 46–54.

Kachelmeier, S. J., Jones, J. D., & Keller, J. A. (1992). Evaluating the effectiveness of a computer-intensive learning aid for teaching pension accounting. *Issues in Accounting Education, 7*(2), 164–178.

McTighe, J., & Ferrara, S. (1998). *Assessing learning in the classroom*. Washington, DC: National Education Association.

Marks, B. R. (1998). An examination of the effectiveness of a computerised learning aid in the introductory graduate finance course. *Financial Practice and Education, 8*(1), 120–126.

Marriott, P., & Lau, A. (2008). The use of on-line summative assessment in an undergraduate financial accounting course. *Journal of Accounting Education, 26,* 73–90.

Rountree, D. (1987). *Assessing students: How shall we know them?* (2nd ed.). Kogan Page: London.

Thelwell, M. (2000). Computer-based assessment: A versatile educational tool. *Computers and Education*, *34*(1), 37–49.

Biographical notes

Hamish teaches finance at Massey University in undergraduate and executive postgraduate courses. In 2008 he was awarded the National Tertiary Teaching Excellence Award for Sustained Excellence. A key theme running through his research is the effect of regulation or market structures on investor and management behaviour.