Quality Assurance by Question Design in Online Teaching

Yousong Luo
yluo@rmit.edu.au
Department of mathematics, RMIT University
GPO Box 2476V, Melbourne, Vic. 3001, Australia

Abstract

Two key issues relating to online assessment are addressed in this article: *fairness* and *quality*. By fairness we mean that the score has to be an accurate reflection of student's achievement and by quality we mean that the objective of each study topic has to be achieved through the assessment mechanism.

1 Overview

Assessment is an important part of the teaching process. It is not only an indicator of the student learning outcome but also a significant learning activity through which students practice, obtain diagnostic feedback, carry out further study or seek further tutorials when necessary, improve and finally achieve better outcome.

Online assessment tasks are usually carried out in a non-supervised environment to meet the need of flexible delivery. Also, as a practice of student centered learning strategy, students should be encouraged to work in teams and to write their reports individually. Therefore, to realistically reflect student's individual achievement and to avoid plagiarism, any online assessment tasks must be *individualized*.

Most of the popular online testing systems support questions of the following types: *multiple choice, multiple answer, text match* and *numeric answer*. Obviously the question types just mentioned are far from satisfactory in testing the understanding of advanced mathematical concepts and methods. This implies that teaching quality will be severely affected by simply using the existing system tools without modification.

In this article we discuss the possibility of using a generic online testing system with capabilities of generating random numbers and limited scientific evaluation power, such as evaluation of exponential, logarithmic and trigonometric functions, to design online assessment tasks which will ensure equal or better teaching quality than the traditional pen and paper assessment.

2 Background

The Distributed Learning System (DLS) in RMIT university include the following tools: email, discussion lists, virtual chat, web-based shared workspace, joint document storage facilities and a variety of web-based testing and surveying capabilities. The DLS were introduced and widely implemented in various courses among the faculties within RMIT since 1998. The detailed features of various system tools included in the DLS are described in details in Fitz-Gerald [1]. Some mathematics courses offered by the department of mathematics are also currently delivered fully or partially online via the DLS. The following is a brief summary of how these courses are delivered:

- Objectives are clearly described. Both week-by-week and topic-by-topic learning activities for achieving these objectives are recommended;
- Supporting materials are provided online and other external resources are linked;
- Student-student and student-teacher communication channels are setup and discussions are scheduled;
- Online components of assessment and tutorials are set around each objective.

The last item mentioned above is one of the key elements of Online delivery and the major system tool used in mathematics courses at RMIT for that purpose is WebLearn. WebLearn is a web based testing package developed by the Distributed Computing Research Group within the Department of Computer Science at RMIT.

The author is currently using WebLearn in his teaching of five subjects, four of which are first year *calculus* for students studying applied science and the remaining is first/second year *linear algebra* for students taking mathematics as their major. In one of the subjects, for example, the component of WebLearn tests replaces a part of the traditional pen and paper continues assessment in the following way

	Option A	Option B	Option C	Option D
Traditional Tests	10%	0%	10%	0%
WebLearn Tests	10%	10%	0%	0%
Examination	80%	90%	90%	100%

Obviously these options are to accommodate students' various learning patterns and their degree of engagement with the subject material.

3 Fairness, Efficiency and Quality

In choosing an online testing system the following main issues were brought into consideration.

Fairness It is commonly recognized that an online testing system should have the features of producing similar but individualized problems for each student. This is usually done in two ways: either the instructors create a large quantity of similar questions of the same type and the system choose one randomly for each student, or the system can generate random variables and evaluate students' input against the random variables. In the later case instructors need only to create one question with some random variables embedded in it. Most of generic testing systems have this capacity.

Efficiency Here we are more concerned with the efficiency of creating problem sets on the instructor's side. Of course, there are many other aspects, such as efficient delivery, which means minimum hardware and software requirements on students' side.

Our experience shows that creating a large quantity of similar questions of the same type is very time consuming. Once the question banks are created they can be used year after year, but to update or insert new topics it will take the same amount of time again. Using random variables in questions seems to be much more efficient in authoring, but there are many things which can not be tested by numerical answers.

The other problem of authoring online tests is the skills involved when a testing system is used. Some excellent packages are particularly designed for handling mathematics problems but one has to learn a specific language to compose the questions. Acquiring new skills to already heavily loaded teachers is a burden and therefore it is a disincentive from using it.

Quality There are many aspects regarding quality online teaching. What we are concerned with here is: when traditional assessment methods are replaced with online assessment methods how can we control the balance of gain and loss so that curriculum objectives can be achieved to the same standard or better.

Teaching quality is measured by learning outcomes. Research shows that "learning outcomes are determined by a whole complex of factors: fixed student-related factors such as ability; teaching-related factors such as curriculum, methods of teaching and assessing; and the student-centered activities which are engaged by students to achieve an outcome" (Biggs, [2]). Assessment tasks are significant parts of students' learning activity and how well they are engaged in these will directly affect their learning outcome.

Let us take teaching engineering mathematics as an example and make a comparison of the strengths and weaknesses between traditional and online assessment. In a traditional class room environment a set of problems is handed out to students and students are then expected to write a solution with full working out demonstrating their understanding of concepts and mastering of methods. Difficulties are addressed and tuition is given by teachers and friends on the spot. When students' work is marked, comments on mistakes among the steps are fed back to the students and a score for partially correct solution is recorded. The advantages of doing so are that:

1. through face-to-face contact teachers can actively and emotionally motivate students who are less interested in the subject;

2. students can directly ask instructors questions and therefore instructors have exact knowledge about where students have troubles so that explanation or demonstration can be provided interactively;

3. scores for partially corrected solutions are given so that students may still obtain higher

grades even though they made minor mistakes that leads to wrong final answer to a question;

4. we believe that, direct contact with the aid of pen and paper is still the most effective way for instructors to communicate complicated mathematical idea with students.

The drawbacks of this kind of assessment tasks are that:

5. contact hours are limited by class time;
6. the number of students contacted is limited due to instructor-student ratio and students' personalities (some students are less active in talking to instructors);
7. marking students' work takes time and hence comments written on students' work may be fed back to students far too late;
8. although the tasks are carried out under the supervision of instructors, copying is still unavoidable due to the same task problems.

In an online environment the opposite appears to be the case. Corresponding to (5)-(8), online assessment shows strength in all areas where traditional assessment has weakness:

(5') contact hours are unlimited;

(6') each student has the equal opportunity in communicating to 'instructors'-computers;

(7') feedback on students performance is instant;

(8') individualized tasks avoid students from copying;

in addition to all these,

(9') students are allowed to attempt the question several times so that they have opportunity to correct mistakes.

Meanwhile, corresponding to (1)-(4), online assessment shows weakness in all areas where traditional assessment has strength:

(1') it may be difficult to motivate students;

(2') students' feedback can not reach instructors immediately so that assistance can not be provided to the students who have troubles with their work;

(3') scores are only given to the answers that are evaluated by the system and no score at all for all work in between;

(4') complicated mathematical ideas involving symbolic calculation can not be easily communicated in text.

Before answering these questions we note that in our current teaching a large percentage of assessment is still allocated to the traditional pen and paper examination (e.g. as shown in the table in Section 2). This implies that we have to prepare students for the final examination by giving them enough practice in continuous assessment tasks. It is dangerous and unfair to let students only click on a mouse button during the semester and at the end give them a written examination.

Our resolution to the problems addressed in the weakness list above is as follows.

To ensure students become engaged in a task, we believe that questions demanding some inputs are more effective than questions of multiple choice. Obtaining the answer to a mathematical question generally requires a few steps of working so that when several answers are given in a multiple choice question some less motivated students are likely to choose an answer by guessing rather than working it out. When such an answer is submitted and feedback indicates that the answer is wrong, the student may not care about it very much because no work and effort are put into it. A question of numeric answer, for example, will mentally rule out the possibility of guessing so that students either have to actually work it out or forget about it completely. When an answer involving a lot of work is submitted and the system marks it wrong the student is likely to re-submit after the mistakes are found and corrected based on the work and effort they have already put into this task. This cycle may go on several times during which the student may consult fellow student or contact the instructors for a resolution if they still be marked wrong. Eventually the learning objective is achieved.

To compensate students' loss of marks for partially correct solutions, and to test students' ability of applying a method rather than an answer alone, questions should be designed to take answers at each key step, e.g. splitting a question into several sub-questions. To provide assistance along the way hints and other supporting materials should be fed back at each step when students' entered a wrong answer instead of only giving feedback at the end.

The ideal testing system for advanced mathematics problems should not only be able to display typeset quality mathematical symbols and expressions but also be able to handle numeric, symbolic, and string answers, such as

WeBWork (`http://webwork.math.rochester.edu`).

However, from the students' point of view, it is frustrating that a syntax error may occur when inputting a correct symbolic answer to a question. Unless it is covered within the curriculum syllabus it seems to be an extra load for students to learn the syntax. Also, from the teachers point of view, additional skills are very often needed to create questions for a complicated testing system.

The WebLearn system we are currently using serves our needs very well. First of all it is simple. Writing questions for WebLearn doesn't need any additional skill and software but a text editor. For display it supports all standard web publishing protocol. For example we write questions in plain HTML (with imbedded graphs when necessary). For answer handling, apart from all the standard features that a generic testing tool has, it can perform basic operations like addition, subtraction, multiplication, division and power and can also evaluate elementary scientific functions such as trigonometric, exponential and logarithmic functions. With this limited power of evaluation we can still build questions with very complicated answers. Suppose that we create a question with several random variables to be generated by the system and that the answer to the question is a symbolic expression. We can partially give the answer to the students at the end of the question leaving a few places for students to enter numeric answers. At the back of the question we then write the answers for the missing parts only as functions of the random variables. In such a way students can only enter the correct answers until they have fully worked out the symbolic answer on their script paper. On the other hand the partially given answer also serves as a guide that tells students what they should aim for. Sometimes a question has more than one correct answer or actually infinitely many answers. In such a case we may specify a form of answers in the question and again let students work out the missing parts of the answers. This time, not only do the answers to the question have to be found but also equivalence among different answers has to be understood by the students. By going through the procedure of obtaining other solutions from one found solution students will definitely need a better understanding of the deep mathematical concepts involved in the given problems.

Many universities around the world have developed similar web based testing tool as our WebLearn as described in Hewett [3]. Most of them have the features of multiple choice, numeric answer and text match. Developing a brand new, powerful and specific testing tool to handle advanced mathematics problems is very expensive, but a little extension of the existing tools to include evaluation of most common scientific functions is by no means a luxury and such a little extension can solve many big problems.

4 Examples

In this section we collect some sample questions used in our teaching.

4.1 Numerical integration

The following question is designed for and used in a course of first year calculus.

Objective: To use Simpson's rule to approximate a given definite integral.

Question type: Numeric with random variable generation and simple scientific evaluation.

Philosophy of question design:

1. To test the method rather than the answer;
2. To design the question and answer in the way so that (a) evaluation of student's input will be simple and within system's capacity; (b) author's work is minimum and author's solution is readable and easy to be edited.

Question 1 *Use the Simpson's rule with* 4 *sub-intervals (strips) to approximate the integral*

$$\int_0^a \frac{1}{b+cx+x^2}\,dx$$

Discussion: At a glance this seems to be trivial because we can simply let a, b, c be random (with some constrains such that $4b < c^2$) to make the question individual and write the answer in terms of the random variables. The exact answer then is

$$\frac{2}{\sqrt{c^2-4\,b}}\left[\operatorname{arctanh}\left(\frac{c}{\sqrt{c^2-4\,b}}\right) - \operatorname{arctanh}\left(\frac{2\,a+c}{\sqrt{c^2-4\,b}}\right)\right].$$

The exact answer is not too complicated to write but if we use it as the standard answer students may use computer packages or powerful calculators to by pass the Simpson's rule. The other problem is that the answer requires the evaluation of arctanh and this is testing tool dependent.

The Simpson's approximation with 4 sub-intervals is

$$\begin{aligned}
&\big(75\,da^8 + 4496\,cda^6 + 926\,ca^7 + 96\,d^4a^5\\
+\;& 240\,d^3a^6 + 210\,d^2a^7 + 3328\,cd^3a^4 + 6880\,cd^2a^5\\
+\;& 24576\,c^3da^2 + 15616\,c^2d^2a^3 + 22784\,c^2da^4 + 12288\,ac^4\\
+\;& 9\,a^9 + 18944\,c^3a^3 + 7888\,c^2a^5\big) \,/\, \big(1350\,cda^7 + 18540\,c^2da^5\\
+\;& 3690\,c^2a^6 + 1728\,cd^4a^4 + 4320\,cd^3a^5 + 3780\,cd^2a^6\\
+\;& 14400\,c^2d^3a^3 + 29088\,c^2d^2a^4 + 46080\,c^4da + 40320\,c^3d^2a^2\\
+\;& 57600\,c^3da^3 + 18432\,c^5 + 162\,ca^8 + 34560\,c^4a^2 + 19656\,c^3a^4\big).
\end{aligned}$$

The answer for the approximation is so unreadable that if any error occurs it will be difficult to find out what goes wrong. A few minutes of careful thought may solve this problem easily. If we only let a be random and let $b = 2a^2$, $c = 2a$ the question will still look as random as before (the upper limit of integral, the coefficients of x^2 and x have different values). However the integral becomes

$$\int_0^a \frac{1}{2a^2+2ax+x^2}\,dx = \frac{1}{a}\int_0^1 \frac{1}{2+2t+t^2}\,dt.$$

The answer by a direct evaluation is approximately $0.321750/a$ and the answer by Simpson's rule is approximately $0.321747/a$. Both are extremely simple. To avoid students from by passing the Simpson's rule we then choose the later as the correct answer and set an appropriate tolerance to rule out the former.

4.2 Linear algebra, answers are not unique

The question is about finding a basis for the subspace

$$S = \{f = ax^2 + bx + c : f(0) + f(1) = 0\}$$

of $P_2(R)$ (the space of real polynomials of degree less than or equal to 2). Obviously the subspace is of dimension 2 and any two linearly independent polynomials in S will form a basis. To produce individual questions we let A, B random and we present the question to students in the following way:

Question 2 *S is subspace of $P_2(R)$ defined by*

$$S = \{f = ax^2 + bx + c : Af(0) + Bf(1) = 0\}.$$

Find the coefficients of s and t in

$$g = -x^2 + sx, h = tx^2 + 1$$

such that $\{g, h\}$ forms a basis for S.

When students use the method they have learned to solve the problem they may obtain all kinds of answers. Then they have to link the answer they obtained to the answers given. By doing so they have not only learned how to solve this problem but also understood the concepts of *basis* and *linear dependence*. Some students may skip any method to determine s and t directly by substituting $g(x)$ and $h(x)$ into the condition $Af(0) + Bf(1) = 0$. Being able to do this shows that they have already understood the topic more than enough. In such a case should we care about which method they use?

4.3 Differential equations, symbolic answers

This is a typical example where we use a numeric answer to replace a symbolic answer. The aim of the question is to test if students know the method of solving a separable equation. We present a question in the following form to the students.

Question 3 *The solution to the differential equation:*

$$\frac{dy}{dx} = \frac{x^2 + 2}{y - y^2}$$

with the initial condition $y(0) = 0$ can be written in the form:

$$-y^3 + ay^2 = x^3 + bx$$

where a and b are some constants. Find the solution in the form given above and enter the value of a and b.

Here we did not show the random variables in the question but we can put them wherever we like. At the back of the question we then write the standard answers for a and b only rather than the solution $y(x)$. This question has the same effectiveness as the the written version of the same question. The solution given is in implicit form. Therefore it is more difficult to determine a and b by substituting the given solution into the differential equation and the initial condition than to solve the differential equation. On the other hand, if a student use a computer package, for example Maple, to obtain a solution they will still have a great trouble to tell the values of a and b. Hence the method is tested rather than the answer.

For all questions above we allow students to have multiple attempts so that they can learn and maximize their score.

5 Conclusion

As mentioned in Section 2, in one of our subjects several options of assessment methods are given to the students as shown in the table. A statistical analysis of the same subject shows that students' scores in all options are extremely consistent. This implies that those students who achieve good scores in WebLearn perform equally well in written examination. The same analysis performed some years ago, when lower quality continuous assessment tasks was used, showed the opposite, namely that the score of continuous assessment was higher than the score obtained in the examination.

Finally here are some student's comments:

- "It helped me understand the materials, give me more WebLearn and less written assignments."
- "It helped me develop my skills."
- "When you are wrong it makes you want to redo it."

References

[1] G. F. Fitz-Gerald, *Mathematics Delivery using Online Destributed Learning Systems at RMIT,* Proceedings of the Fourth Asian Technology Conference in Mathematics, (1999), 321-328

[2] J. Biggs, *What the student does: teaching for enhanced learning*, Higher Education Research & Development, **Vol. 18**, No. 1, (1999), 57-75

[3] M. Hewett, & T. Pran, *The use of diagnostics in web-based learning for introductory statistics courses*, Gazette of The Australian Mathematical Society, **Vol. 27**, No. 4, (2000), 161-167