HIGH STAKES TESTING USING WEBASSIGN

Lisa Townsley
University of Georgia
UGA Mathematics Boyd GRSC Athens, GA 30602
townsley@math.uga.edu

The mathematics department at the University of Georgia (UGA) has a long-standing tradition of standardized, electronic testing for all sections of precalculus. A primary motivation for using online assessment throughout the precalculus course was a lack of uniform assessment of basic precalculus skills. Since some 2000 students enroll in precalculus each year at UGA and then transition to calculus, there is a need for the earned final grade of A (B, C+, etc) to indicate a uniform level of preparation. Using the same homework, online quiz, and test for all students provides that uniformity, but is a challenge to deliver securely and grade consistently. In 2009, the department transitioned to a new assessment platform. WebAssign® is a commercial product developed at North Carolina State University, and it provides enough flexibility to serve the needs of the precalculus curriculum at UGA. In this paper, I will discuss how we address the special concerns that arise when providing online testing.

At UGA, we use WebAssign to deliver homework assignments and "quizzes". We also use WebAssign for testing. Testing is conducted in a lab outside of class time, proctored by student workers. As an instructor of such a course within a multitude of sections, you are held to a strict timeframe, but your only assessment duties are in-class quizzes. Since the primary responsibility is instruction, this course becomes an appropriate venue for graduate students as first-time teachers to develop their teaching skills. In addition, online assessment provides some grading relief to departmental instructors with high teaching loads.

The precalculus course covers five main units. The first unit includes a review of circles and lines, an introduction to function, and modeling with functions. The second unit includes transformations of functions both graphically and algebraically, a study of quadratic functions and their extrema, and modeling and combining functions in an applied setting. The third unit studies exponential and logarithmic functions, their applications, and solving equations involving these functions. The final two units study trigonometric functions, including right triangle trigonometry, radian measure, the functions' graphs and properties, solving trigonometric equations, inverse trigonometric functions, and various formuli and applications. Because testing for each content unit is available on specific days, the teacher is required to cover the mathematics content according to a fairly inflexible schedule.

Students enrolled in Precalculus at UGA are introduced to WebAssign in their first lecture. They self-enroll into their WebAssign course, and homework awaits them. Because the campus provides computer access and WebAssign works well on the

profusion of browsers available on the student's personal computers, access is not an issue. WebAssign allows for posting of syllabi, worksheets, video, and announcements and also has a gradebook, so there is no need to use an additional course management system. At UGA there is also a course website that reminds students of course requirements (ID necessary for tests, test format, test calendar, attendance policy, etc.).

Fig. 1 Math 1113 Course Webpage
http://www.math.uga.edu/116/1113home.htm

The homework format is designed to encourage study and minimize student frustration. Problems selected from the textbook content provided by the publisher include links to "read it" (the text) or "watch it" (see the problem solved in a step-by-step fashion with other numbers in place). The solution key to each homework problem is released after one student attempt at the problem, so a conscientious student can earn a 100% homework grade for just attempting the homework. The webquizzes have a different format: no solutions, hints or links before the due date, and only 3 attempts per problem, with a penalty for errors requiring subsequent attempts. This design should help students prepare for tests—a problem marked incorrect that has another attempt available encourages and rewards student persistence, while the limit on attempts minimizes frustration.

There is a natural segue from the webquiz format to the testing format. Students have 75 minutes to complete 12 test questions, which they may or may not have seen before in a webquiz or homework assignment. This time the student is allowed only two attempts per answer blank. The chief difference is the testing environment. Tests are taken outside of class time—which frees up 5 lecture hours for the instructor. Undergraduates are hired to proctor the tests. Students are partitioned according to sections and may schedule a test time for any one of many blocks on a given test day; the scheduling is completed online

as well. The course webpage and announcements in WebAssign prod students to register for a convenient test time. The scheduling software was created at UGA, and uses the campus student ID system.

Once a student arrives at a test, they surrender a photo ID to prove their identity. Students are also asked to stow their phones, either with the proctors, or in their bookbags, which are also sequestered away from the testing computers. Since the protocol is the same for each test, students don't have any issues with check-in or what to bring/stow. Students login to WebAssign and wait for the password to start their test. The password does *not* provide security—it merely helps one start the student tests simultaneously.

Once the test begins, a timer in WebAssign keeps the students informed as to their available time. Students draw one of a pool of problems for each assessment item, and each problem features randomized numbers. This way, students who violate the academic honesty policy by outside discussion with other students about the test items have minimal information to convey. Students have scratch paper to work on, and the scratch is returned to the instructor in case of questions regarding grading.

Using instant grading via WebAssign is superior to a multiple-choice test for mass-testing. First, on a multiple-choice test, the usual mistakes a student makes become "distractor" answer choices, so students are likely to get problems wrong. This contrasts with our WebAssign-based test, where students with an error are given a second attempt. Second, a WebAssign-based test offers students the opportunity to enter symbolic answers that are algebraically equivalent to the key, whereas on a multiple choice test, the student needs to manipulate his/her solution until it matches one of the choices. A variety of answer formats are available in addition to multiple choice: multiple-select (choosing all correct answers), symbolic answer entry, numeric entry, and true/false (and others we don't currently use). We mark problems by partial submission, since a student should be able to inform herself before working on a new part of a problem. We use WebAssign to provide a quid pro quo for the loss of partial credit on a paper-and-pencil test: partial submission gives students information as to whether they are on track to a correct solution, and instant grading provides the student an opportunity to modify incorrect work for 75% credit on a second attempt.

At UGA we have developed a standard format for questions and for answer types, so that communication of answers via computer during a test is less daunting to the student. Some questions are multiple-choice or True/False, and there is no difficulty with answer format there. For other questions, the answer blank *always* accepts symbolic forms of the answer: 2 or 1/(1/2) or e^(ln(2)) are all marked correct for an expected answer of 2. Some question types disallow decimal forms of the answer, to preclude student use of the calculator solver when we want to test equation solving skills. The student who enters 0.5 in such a question is not permanently penalized, however, because teachers can easily override penalties in WebAssign provided the scratchwork supports the student's claim.

WebAssign is perhaps the industry standard for protected mathematics testing. Other software product vendors insist that security can be maintained with just passwords, but at UGA we don't believe that test integrity can be protected over 4 days of testing with

endlessly changing passwords, nor do we desire the course management protocol that would require. In WebAssign, the assignments can be restricted to (or away from) specific IP addresses. This means that a student can only access the test in a designated testing environment and also cannot access homework and quizzes while working the test. In WebAssign, tests can be timed and password protected as well, with an additional enhancement available that blocks access to any other software. At UGA, we use a desktop calculator emulator and so do not use this particular feature. The computers are set to lock browsers at the WebAssign site, to prevent student searches for test assistance or sharing of test information.

Additional concerns regarding testing security include disability access and students' tendency to share information about a test. Our list of IP-enabled locations includes the testing center for students who need extra time and/or a quiet environment to test. The testing center restricts those computers to testing so that students cannot access test materials at other times. Due to the large student enrollment in precalculus and the fact that we must offer many time slots for students to test at their convenience outside of class, testing takes 4 days in the fall semester and 3 days in spring to complete. To minimize sharing of test information, the pool of available problems for each particular test question must be large. In addition, we remind students that the academic honesty policy requires them to respect the integrity of other students' work. We have prosecuted students for sharing information about particular test questions ("the drawbridge was on the test") and about test format as well ("how many true/false questions were there?").

Once a secure testing environment is created, test design can be considered. Format of a test can either assist a student to show her/his best or it can frustrate her/his best efforts. This is a good reason to stick to the same answer format, in our case allowing symbolic forms of the answer and not requiring simplification. This format is preserved in homework and webquizzes whenever possible, to help the student be reasonably comfortable at a test. We don't randomize the order of questions on a test, but walk the student through the content in the order as presented in class. The exception is to move word problems to the end of a test, since students can get stuck extracting information from a word problem and fail to manage their time well. Since all the questions are delivered at once (not unlike a paper test), students can still pick and choose the order in which they work problems. We try to begin the test with a problem that is easy, such as a multiple choice problem. Many test problems are straightforward, and so are likely to be repeats of homework problems (such as: find the line through a given point perpendicular to a given line). We make sure to include some problems with multiple parts, so that a weak student can earn partial credit for understanding some components of the problem. For example, the original form of the question below was: find the function $f(x)$. Now it has two answer blanks, find a and b. The student can now earn credit for finding the easy constant a.

A function f is of the form $f(x) = a + cos(bx)$ where a and b are constants and $0 < b < 8.$ If $f(0) = 5$ and $f\left(\frac{\pi}{8}\right) = 3.5,$ find a and $b.$

Figure 2. Test Question

Testing on a computer can be frightening for students, and rightfully so: they are expected to think about mathematics and about appropriate data entry simultaneously. This is more challenging than a traditional pencil-and-paper test, and a good reason to adopt the format described above to lessen test anxiety. Teachers can coach their students to exhibit smart test-taking behavior—for example, start by writing down your useful formuli on the scratch paper, then examine the entire test and answer a few of the multiple choice, easy questions. Then, on a question like that below, why would you answer the second part until you have submitted and earned credit for a correct answer to the first part, on which the second answer depends?

Eight years ago a house was worth $290,000. Now the house is worth $390,000. Assume a linear relationship between time (measured in years) and value (measured in dollars).
   (a) Find a formula for the value, $V(t)$, at time t. ($t = 0$ refers to now.)
       $V(t) = $ _____
   (b) What will be the value of the house five years from now? Give an exact answer or round your answer to the nearest dollar.
       $ _____

Figure 3. Test Question

In contrast to the student who is frightened of computerized testing, there are excellent "gamers" taking the test. Good test design should distinguish between students who are taking a test "smartly" as described above, and those who are gaming answers. For this reason, we don't allow the randomized numbers to be displayed in red anywhere in the coursework. WebAssign's default is to distinguish randomized numbers in a problem by displaying them in red. The red color allows students to memorize how the solution form uses the randomized numbers and encourages memorization of specific answer *forms* rather than learning a *process* to complete that problem type. Similarly, we don't display the "name" of a WebAssign problem. The name indicates what section of text the problem comes from, which is information that we would like the student to deduce from the problem itself. Finally, the answer format of textbook questions can lead to gaming. See the two examples below, asking for the same solution set. The first version allows a student to enter "NONE" in the last box, to win points for no understanding related to the question. The student can work backwards with "NONE" until a box is marked wrong, and then know exactly how many correct answers there are. In addition, $\arctan(\frac{1}{\sqrt{3}})$ is an acceptable answer in the first question while second question expects a more complete answer.

Find the solutions of the equation that are in the interval $[0, 2\pi)$. (Enter solutions from smallest to largest. If there are any unused answer boxes, enter NONE in the last boxes. If the equation has no solution, enter NONE for each answer.)
$$\cot^2(\theta) - \sqrt{3}\cot(\theta) = 0$$
$\theta = $ _____ $\theta = $ _____ $\theta = $ _____ $\theta = $ _____ $\theta = $ _____ $\theta = $ _____

Figure 4. Good homework problem, bad for tests.

A gamer's approach to Figure 4: step 1, submit "NONE" in the last answer box. (Get a point!) Step 2, repeat. Eventually I determine that there are 4 numeric answers in $[0, 2\pi)$, without any interface with the quadratic equation.

Find the solutions of the equation that are in the interval $[0, 2\pi)$. If there is more than one answer, enter them in a list separated by commas. Enter exact answers: decimal approximations or symbolic trigonometric expressions such as arctan(5) will be marked incorrect.

$$cot^2(\theta) - \sqrt{3}\cot(\theta) = 0$$

θ = _____

Figure 5. Same problem, coded for tests.
WebAssign allows any order in the list of answers.

Finally, good test design should distinguish the various grade levels of students. We are commited to at least two word problems on each test, and one problem that requires several algebra steps. Skill at both these types of problem is necessary for calculus. However, too many multi-step algebra-intensive problems will not distinguish students, since the good student is likely to make miniscule algebra errors while the weak student may make a severe conceptual error, but both students earn 0 points. The final exam scores are generally 10 points lower than the tests, indicating that students are cramming for tests and temporarily memorizing solution methods. Repeating "vanilla" problems from previous tests on the final exam distinguishes whether the students crammed for tests or actually learned the basic problem-solving skills necessary for calculus study.

Here are some final words on testing security. Since we test over several days, we need a large enough data base of problems in order to make the pools of problems for a particular test question large. We keep a record of which question types were asked each term, to try to avoid predictability. Questions which have been coded by faculty at UGA are stored in "The Attic" a WebAssign course that is a library of these questions filed by section. This library, together with access to all publisher-provided questions, informs the lectures of each instructor. The instructors do not see the tests prior to release on the test day. This insures that no teacher can prefer his/her students over other students enrolled in the course by providing inappropriate advance information.

What was a relatively isolated problem of fraternity members hoarding (paper) exam copies for study by subsequent students is now universal—web postings mean every student has access to old tests. Websites such as www.studyblue.com , www.koofers.com and www.cramster.com are just the tip of the iceberg: students can form open or *closed* study groups in Facebook™ and post materials at will. It is a fine line to encourage group study, which in this era can be done asynchronously via such websites, and to also promote appropriate behaviors that respect the integrity of each student's work. At UGA, we are still grappling with appropriate methods to plug such security leaks. We vigorously prosecute known offenders, but many more are unknown. This problem existed before online testing; it occurred any time testing took more than one period to

administer. Saving and sharing test information electronically just amplifies the same difficulty. This means that there is an eternal need for new problems, particularly word problems, since a "new" test question this term has been seen by some subset of students next term.
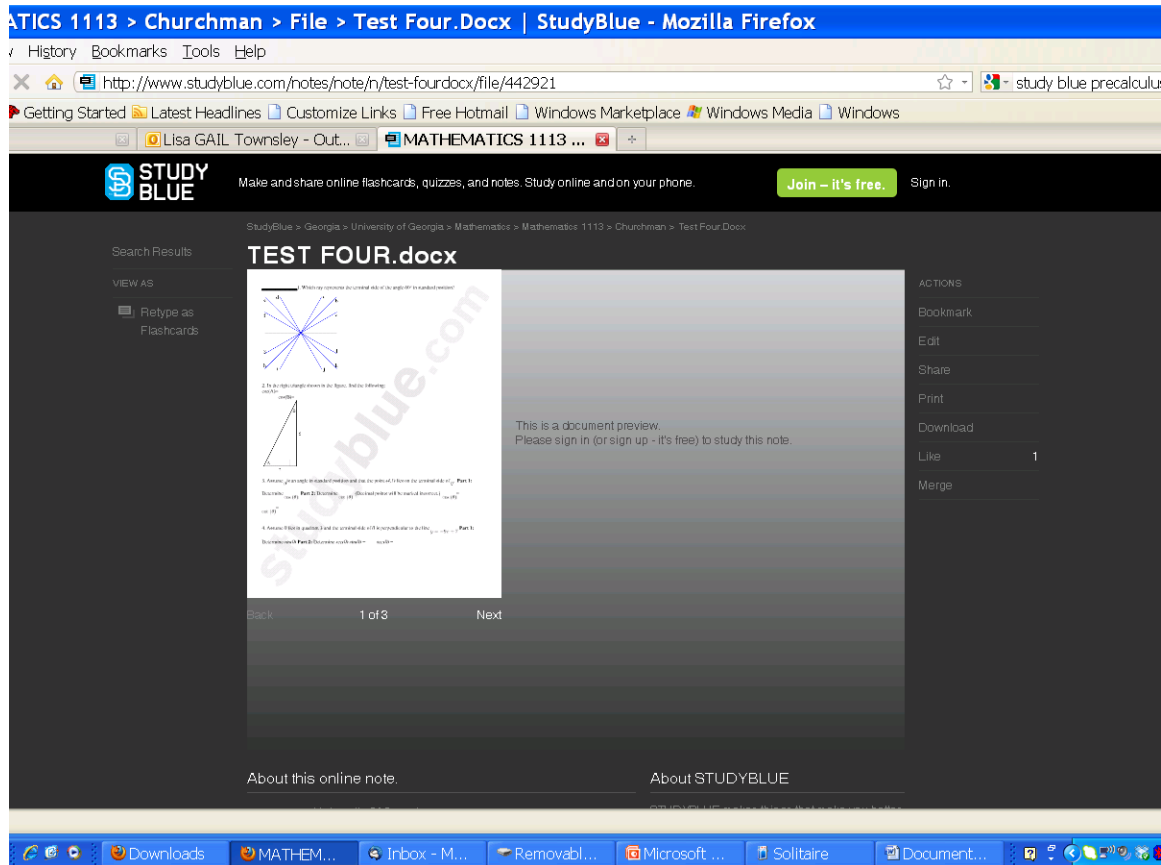

Figure 6. Screen shot of web posting of a UGA test.

WebAssign allows teachers and the course administrator to gather meaningful data about student progress and overall student scoring. In each class, both individual problem and class-wide mean and median statistics are available. By clicking on the student's name in the data, the student work can be viewed. A teacher can see not just the final answer entered by a student on the particular problem drawn from the pool, but all attempts (for example, the student may write a correct formulation, but spelled in upper case $y = 3X + 4$ rather than $y = 3x + 4$). In addition, the teacher can override scores on assignments (add or subtract penalties); together, these features allow the teacher to ameliorate the perceived pain of impartial computer grading. See Figure 7 below for additional class analysis available for each problem. The screen shot shows a summary for test 5 (advanced trigonometry). Students drew one of a pool of 3 exercises. In this class 12 students drew this particular problem with randomized numbers. As the first test problem, it was selected to be easy and the students did well on the problem.

Just as test means and medians are available for each section, a download of all section scores (that is somewhat cumbersome to manipulate in Excel) allows for comparison of individual class scores to overall means and medians. This data is regularly used by the department head, together with the record of when a particular student was downloading and executing his/her homework, to answer queries from "helicopter parents".
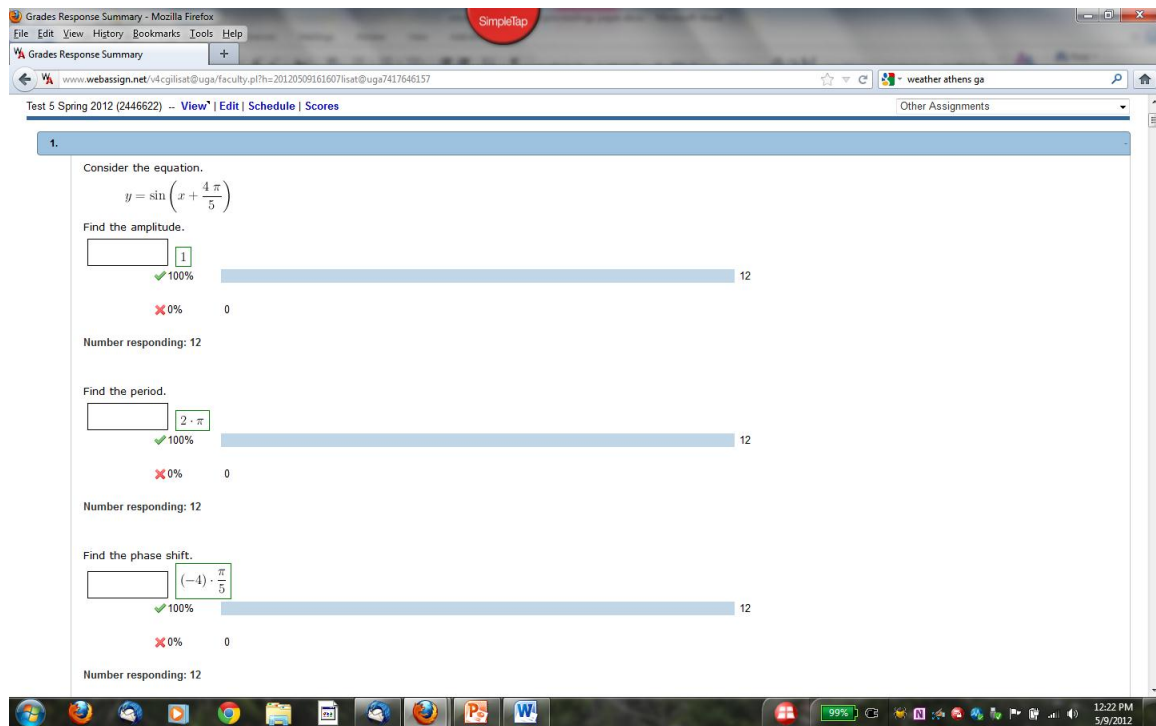

Figure 7. Response Summary for a test 5 problem

UGA mathematics has been using WebAssign as its testing platform for Precalculus for 4 years now, and both the software and our testing design has evolved over time. We are satisfied that we are providing a course of study that can help students learn and demonstrate their understanding better than other mass-testing options. We have refined our testing design to insure accurate assessment of students understanding and optimize student performance. We continue to look for ways to enhance the student learning experience in order to appropriately prepare them for calculus.

References

1. UGA Precalculus website: http://www.math.uga.edu/116/1113home.htm as part of Department of Mathematics website: http://www.math.uga.edu
2. Virtual TI v2.5 beta5. Download from http://www.ticalc.org/archives/files/fileinfo/84/8442.html . The organization ticalc.org is not affiliated with Texas Instruments Incorporated in any way.

3. WebAssign® is a registered service mark of North Carolina State University under license to Advanced Instructional Systems, Inc. © 1997-2003 by North Carolina State University. Portions © 2003-2009 by Advanced Instructional Systems, Inc. Various trademarks held by their respective owners. Website http://www.webassign.net