Effectiveness of Web-based Quizzes in Mathematics Courses

Lisa Bromberg
Michael Seminelli

This paper was completed and submitted in partial fulfillment of the Master Teacher Program, a 2–
year faculty professional development program conducted by the Center for Faculty Excellence, United States
Military Academy, West Point, NY 2017.

Abstract

As technology continues to establish a more prominent role in the classroom, educators are turning to web-based assessment tools to enhance the instruction of mathematics. The ease of grading and ability to provide the student quick feedback makes this style of assessment appealing to the instructor. This study investigates the effect of web-based versus paper-based quizzes on student performance in a freshman mathematics course. A sample of 150 students from nine sections were randomly assigned an assessment format: paper, web-based through WebAssign, or the option to select either form. Students also provided qualitative data by responding to surveys at the beginning and end of the semester answering questions on their preference for quiz format: on paper or through WebAssign. The study collected student data from nine assessments and the results show section scores on quizzes, broken down by quiz format of paper, WebAssign, or option.

1 Introduction

1.1 Motivation

Students are familiar with a “paperless” classroom by the use of digital textbooks to online homework assignments. We often assume that it is better to practice under the same conditions that you will take an assessment, so are we hurting our students performance by giving them exams on paper? It is better to return graded quizzes and discuss answers as soon as possible, but does this actually make a difference? Furthermore, we are led to believe in modern society that teaching and testing with technology helps student retention, which may not necessarily be true. This research aims to study differences, if they exist, in student performance based on the medium quizzes are administered.

Beyond the question of deciding how to assess learning throughout the semester, there’s the question of feedback. We traditionally do this by administering quizzes, tests, or problem sets to identify a Cadets course material understanding. We handout the testing material, in the traditional sense of hard copy or paper copy, the students take it, then they pass it back to the instructor to grade later. Often the grading can take a day or two, which is valuable time that a student might reinforce incorrect concepts that they will have to recommit to memory. This is a waste of time for both the instructor and the student that technology could eliminate by providing real-time grading with the added benefit of identifying those students that are quietly struggling with concepts but are afraid to ask.
There are a number of motivations for employing web-based homework in large university courses: providing students with immediate feedback, using the latest technology for instruction, reducing the grading burden on faculty, being able to grade all student work. However, there are potential drawbacks to using web-based homework: a lack of detailed feedback to students, the danger of multiple submissions encouraging lazy habits, and further impersonalization of replacing a human grader with a computer. By and large, web-based homework systems deliver fairly standard materials and exercises, changing the delivery medium but leaving the pedagogical content much unchanged.

1.2 Background

In the Core Mathematics program at USMA, we emphasize the conceptual level, with the goal for cadets to recognize relationships, similarities and differences to internalize the mathematical conceptual framework. With this goal, we often try to discover the best practices to check our cadets level of conceptual retention and understanding.

For approximately 900 cadets, MA103: Mathematical Modeling and Introduction to Calculus is the first course in the Core Mathematics Program at USMA (≈ 250 students are placed in the Advanced Math Program, and ≈ 100 students are placed in MA100 before entering the Standard Core Math Program) [2]. Material in this course was organized into four blocks: Modeling with Discrete Dynamical Systems, Introduction to Vectors and Matrix Algebra, Modeling with Matrices and Systems of Recursion Equations, and Introduction to Calculus and Modeling with Continuous Functions. Coursewide graded events, including Written Partial Reviews (WPRs), a Term End Exam (TEE), Problem Sets, a Fundamental Concepts Exam, a Briefing, and WebAssign, comprised 90% of the course grade. The remaining 10% were left to the individual instructor.

As part of daily class preparation, students are required to complete the assigned reading and attempt WebAssign problems prior to coming to class. Each WebAssign assignment is due at the beginning of the lesson for which it is assigned. Students must complete an e-acknowledgement of assistance received while completing WebAssign problems.

One of the biggest challenges from the instructor standpoint at USMA is competing for cadets’ time. Their time is valuable, and so even with the best of intentions, cadets will often only complete work if points are assigned. In general, the rule of thumb is that a student should spend around 2 hours outside of class for every hour spent in class. For MA103, which meets 4 hours a week, this is a lot of time; as a course, daily homework problems were administered via WebAssign, a web-based homework program. Students obtain a license with their textbook and WebAssign provides students with access to a digital copy of the text. WebAssign has a repository of textbook questions precoded for use in building assignments. Instructors can code their own questions and add them to assignments. Additionally, instructors control the number of attempts, randomization in values, due dates, and penalty when constructing assignments. WebAssign is a tool to help students understand the material better and for instructors to understand student progress with the material.
2 Literature

Opinions on assessment, and why and how to assess student learning, vary greatly. McKeachie [8] asserts that assessment is not simply an end-of-course exercise to determine student grades. Instead, they can be learning experiences for students. In addition to helping assign grades, it can communicate goals to students so they can learn more effectively, and identify misunderstandings to help the teacher teach better. Because grades are determined by a great degree by test scores, tests are the most common form of assessment. In designing tests, McKeachie is a proponent of tests which require students to think. However, sometimes you just want to know if students are getting the basics down before moving on. For this level of assessment, McKeachie argues in-class quizzes are just the thing [8, Chapter 7].

The idea of utilizing technology for assessments is certainly not new. McKeachie dedicates an entire chapter to it [8, Chapter 17]. According to Teaching Tips [8], instructional technology can serve a number of useful functions, including: providing new opportunities for enhancing student learning that would otherwise be impossible or very difficult, and addressing specific learning goals more effectively. Given the increase in use of web-based programs for homework assignments in large courses, using technology for other assessments is a reasonable leap.

There is pedagogical support for utilizing assessments other than tests, as well technology, to enhance the student learning experience. Putting these ideas together, Bonham, Deardoff, and Beichner [3] conducted a comparison of student performance using web-based versus paper-based homework in a college level physics class. They assessed student performance with web homework and with traditional paper homework in a classroom setting for a large-enrollment introductory physics course. One section received paper-based, hand graded homework, while the other group received the standard web-based homework. The two groups were then compared on conceptual and problem-solving performance. No significant differences in student performance were found that could be attributed to the homework method used.

Though perhaps not immediately applicable to our study, Filer [5] reports on a study designed to assess student learning and engagement in a lecture setting that incorporates the use of an audience response system (ARS), more commonly referred to as a “clicker.” The study sought to determine if the use of clickers in the classroom a) increases student knowledge scores; b) enhances student motivation and comfort in the classroom; c) increases the quality of a lecture as perceived by a student; d) increases student participation. In this study, the use of clickers enhanced the students’ emotional experiences in the classroom by encouraging participation and motivating students to answer questions correctly. Students were required to engage physically in the environment and encouraged to process information independently and commit to an answer. While mean quiz scores were not affected by the use of clickers, the benefits of heightened levels of attention ultimately enhanced the student’s learning experience [5].

One of the consequences of using web-based evaluation is immediate feedback. There is a rift on whether immediate feedback or delayed feedback is beneficial. Kulhavy [6] and Webb [10] argue that delayed feedback is better because it students more likely to argue the correctness of his/her incorrect answer which will interfere with an instructor’s attempts to correct the student. However, Kulik and Kulik [7] argue that their data supports immedi-
ate feedback was most beneficial in student retention. In a large enrollment introductory chemistry course, Cole and Todd [4] investigated whether the use of a web-based homework system could help meet the needs of the diverse population of the course. Among the considerations of their study was the desire to provide immediate feedback to promote student learning and retention.

Many people have researched how technology effects the classroom environment but fail to identify if the timely feedback during class assists in a Cadets perceived learning or actual improvements in mathematics. There is no such thing as good teaching without good feedback [8]. Research on feedback indicates that students do value written comments on their work [9]. However, they express concerns when this feedback is illegible, ambiguous, too abstract, too general or vague, and too cryptic. In addition to the comments on the content of written feedback, numerous studies show that student receive feedback too late to be helpful. With the proliferation of technology in the classroom we attempt to identify if its use is beneficial or a hindrance to the conceptual retention and understanding of Mathematics. We plan to identify if there is a quantitative difference, with respect to mathematical concepts, between immediate or delayed feedback.

3 Methodology

In the fall semester, 900 freshman students (most of the freshman class) take MA103: Mathematical Modeling and Introduction to Calculus as their first mathematics course in the USMA core curriculum. MA103 emphasizes applied mathematics through modeling where students develop effective strategies to solve complex and often illdefined problems. The course exercises a wide array of mathematical concepts while nurturing creativity, critical thinking, and learning through activities performed in disciplinary and interdisciplinary settings. The course introduces calculus using continuous and discrete mathematics while analyzing dynamic change in applied problems [1, 2].

The study was conducted across nine sections of the MA103 course taught by three instructors, involving 145 students. As part of the course design, each student would take nine instructor developed quizzes during class time. Each quiz was worth 20 points and contributed to nine percent of the students course grade. Class sections were randomly assigned to one of three testing groups: control group, treatment group, and the option group. The control group was administered the quiz printed on paper. The treatment group was administer the quiz digitally through WebAssign by using their computers. Students in the option group were given the option at the start of the assessment to take the quiz on paper or digitally. The randomized quiz schedule in Figure 1 was used so each section experienced three quizzes from each of the testing groups over the semester.

Quiz development was shared by each of the instructors in the study, who were responsible for creating three quizzes apiece. The study design required all students to take the same quiz, so the questions and format for the paper quiz were coded into WebAssign to ensure both assessments were exactly the same. Figure 2 illustrates Quiz 6 on network flow in both the paper form (left) and coded into WebAssign (right). Although WebAssign has the capability to randomize the values in the problems, this feature was not used to since the quiz is an in class assessment and to ensure consistency with the paper version.
Figure 1: Assignment of sections to testing groups for each quiz.

Figure 2: The paper and WebAssign versions of Quiz 6.
The randomized quiz schedule was implemented to eliminate potential bias from either the instructors and sections. The test administration was standardized for the study so all quizzes were administered on the same day for 15 minutes at the start of the class period. Students were instructed to bring their computers to class for every quiz regardless of the testing group and were notified of the quiz method shortly before the quiz was administered. The classroom instructor computer was made available to any student who had difficulty connecting to the internet while the quiz was administered. Each instructor was responsible for grading their students’ quizzes but an assessment rubric was used by all to ensure standardization in grading.

4 Results

All student quiz scores were converted to a percentage to facilitate analysis of both the paper and WebAssign study-wide averages for each of the nine quizzes. Table 1 summarizes the average scores by quiz. A two sample T-test was conducted on the difference of the mean quiz scores between the paper and WebAssign assessment methods with a significance level of 0.05. Our hypothesis for this study was that the testing medium (paper or web-based) had no impact on student assessment performance. If one method resulted in significantly greater quiz scores we would reject our hypothesis and conclude that the testing medium impacted student performance.

<table>
<thead>
<tr>
<th>Quiz</th>
<th>Paper Average Score (%)</th>
<th>WebAssign Average Score (%)</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quiz 1- Modeling Process</td>
<td>78.9</td>
<td>70.7</td>
<td>0.0073**</td>
</tr>
<tr>
<td>Quiz 2- Discrete Modeling</td>
<td>93.7</td>
<td>94.1</td>
<td>0.8241</td>
</tr>
<tr>
<td>Quiz 3- Vectors</td>
<td>79.8</td>
<td>79.1</td>
<td>0.8666</td>
</tr>
<tr>
<td>Quiz 4- Cross Product</td>
<td>80.8</td>
<td>76.4</td>
<td>0.2676</td>
</tr>
<tr>
<td>Quiz 5- Matrix Algebra</td>
<td>84.1</td>
<td>80.3</td>
<td>0.328</td>
</tr>
<tr>
<td>Quiz 6- Network Flow</td>
<td>85.3</td>
<td>——</td>
<td>——</td>
</tr>
<tr>
<td>Quiz 7- Rates of Change</td>
<td>88.7</td>
<td>78.7</td>
<td>0.0161**</td>
</tr>
<tr>
<td>Quiz 8- Model Evaluation</td>
<td>83.0</td>
<td>87.6</td>
<td>0.2844</td>
</tr>
<tr>
<td>Quiz 9- Review of Equations</td>
<td>93.4</td>
<td>90.9</td>
<td>0.3713</td>
</tr>
</tbody>
</table>

The paired hypothesis tests for quizzes 1 and 7 were the only assessments that returned statistically significant p-values at the 95% confidence level. The analysis for those two quizzes suggest students perform better on paper assessments compared to the web-based quizzes. We believe this result is due to the construct of the individual assessments instead of an actual difference in the two mediums. Quiz 1 was fill in the blank on the steps of the math modeling process. The WebAssign version mirrored the paper assessment however the automatic grading provided by WebAssign did not account for misspelling of correct answers. To ensure consistency with grading, the instructors went through each WebAssign quiz and adjudicated the score. Quiz 7 tested the limit definition of the derivative where students entered several of steps to solve the derivative, so the assessment was similar to a fill in the...
blank style quiz. Although the students have experience entering algebraic expressions into WebAssign from their daily homework assignments, the quiz administration doesn’t provide them the same safety net of multiple submissions to allow them to correct typos.

Analysis of all quiz scores for the three treatments can be observed in Figure 3. The box plots for the three quiz types are nearly identical suggesting there is no difference in student quiz performance based on the method a quiz is administered. Additionally, the first and third quartiles are aligned which illustrates the range of the middle 50% of student scores is the same regardless of quiz type. Lastly, the range of quiz scores in Figure 3 is also similar for the three treatments, demonstrating no difference in student performance.

A challenge observed during this study was the instructors inability to administer partial credit on the web-based quizzes for student work of mathematical merit which ultimately resulted in the wrong answer. This type of feedback illustrates where a student made a mistake while awarding points to the work that was successfully executed. This helps the students focus their studies when preparing for larger assessments and provides positive reinforcement for correctly applying some of the concepts tested. It is easy to assign partial credit when you have a student’s written work. WebAssign only provides the instructor the student’s answer to each quiz question. The WebAssign software has an option to allow students to submit work, however the students in the study didn’t use this feature in their homework and we didn’t want to put them at a disadvantage by spending quiz time to type in their work.

As instructors, we felt obligated to grade our assessments in the same manner as their major course wide events, following a holistic rubric. The instructor who created the quiz also created a grading rubric for everyone to use. The rubric assigned partial credit for common
conceptual mistakes. Instructors would review the answers provided by their students and based on the solution they entered, the appropriate partial credit would be awarded. This is not an ideal or perfect situation since the instructor had to use judgment to determine if a concept was executed correctly by the student base solely on the response. Grading the paper assessments identified common incorrect solutions which made assessing partial credit for the WebAssign responses easier. Overall, to maintain the integrity of the study, we minimized the partial credit assigned for either assessment to only situations where the mistakes are clearly identifiable on both assessment types.

To ensure the quizzes we developed were both fair and a reflection of the course wide assessments, we compared student quiz performance to their performance on the hour long exams. Figure 4 is a plot of each student’s average quiz score compared to their average performance on the four course wide exams. A plot along the diagonal indicates the student achieved similar results on the major assessments as they did on in-class quizzes, while plots above or below the diagonal suggest better performance on the exams or on the quizzes respectively. The data is clustered around the diagonal which suggests the quizzes developed for this study were of the same difficulty and graded in the same way as the exams given to all 900 students in the course.

The qualitative feedback we received from the students was overwhelmingly in favor of paper quizzes. When they were presented with the option, over 80% of the students in a given section elected to take their quiz on paper. The students said they already have to write out their work on paper to solve the question and were afraid that entering the values into WebAssign used too much of their time, limiting their ability to complete the quiz. Each section, however, had a couple of students who stuck with the WebAssign quizzes because they wanted the instant feedback. One student said that he used WebAssign homework to

Figure 4: Student quiz averages plotted against their averages on course wide exams.
prepare for the quizzes and preferred web-based assessments since they were in the same format as his study method. While one assessment method will never please everyone, the students’ familiarity and comfort level pushes most of them towards paper assessments.

5 Conclusions

Overall, there is no difference in student performance on quizzes given on paper or on a computer. In particular, the middle 50% of student scores is the same regardless of method of administration. Quizzes 1 and 7 were the only assessments on which students performed better with the paper version as opposed to the WebAssign version. Unlike the other quizzes, these two had fill in the blank style questions: on Quiz 1, students were asked to enter the steps of the Mathematical Modeling Process; on Quiz 7, students filled in the steps to finding a derivative using the limit definition. With the former, WebAssign was case-sensitive and did not account for spelling errors. For the latter, students are unaccustomed to typing algebraic expressions with a single attempt. For spelling errors, the instructors were able to adjudicate scores. For typos with algebraic expressions, this becomes more difficult, as now it is more a question of a conceptual error versus a more basic one, and partial credit becomes more murky. Because of this, question design in web-based quizzes needs to consider potential issues when grading student responses.

In conjunction with question design, a challenge of this study was assigning partial credit. Since one of the goals of in-class quizzes is to reinforce learning and give feedback between major graded events like exams, awarding points for some correct mathematical work provides positive reinforcement to the student. While WebAssign has a feature to allow students to submit their work, this was not used in our study. Since the instructors did want to grade according to the holistic model used throughout the course, compromises had to be made in awarding partial credit on paper quizzes, and assumptions of the type of error had to be made on WebAssign quizzes.

Overall, student quiz scores plotted against student performance on course wide exams indicate that the quizzes in this study were of the same difficulty and graded similarly as exams given to all students in the course.

6 Future Work

It would be interesting to repeat this study focusing on alternative methods for assigning partial credit. WebAssign has a built-in feature which allows students to submit their work. One issue here becomes the time it takes to administer a quiz: it takes students significantly longer to type a mathematical expression into WebAssign than to write out their work, and it would take much longer for them to include their work. The other downside to this feature is that an instructor would still have to look over the students’ work to assign partial credit. It would be interesting to see if using this feature canceled out the goal of immediate feedback to students.

Another idea for future work is to determine if the grades assigned using a multiple attempt format are the same as those graded using partial/follows credit. Rather than...
having an instructor evaluate work submitted, perhaps allow students using WebAssign several attempts at each question. Similar to homework submissions, students have multiple attempts. If after 3 submissions, the student is still incorrect, they get no credit for that question, whereas if they submit the correct answer on attempt 2, they get full credit. Educators incorporating web-based assessments will need to explore creative methods for awarding partial credit to ensure the focus remains on the methodology and not just the answer.

References


