2006-2238: A COMPARISON OF ON-LINE AND TRADITIONAL TESTING METHODS

Steve York, Virginia Tech

Dr. Steven C York is an assistant professor in the department of Engineering Education at Virginia Tech. He received his BS degree in chemistry from Radford University in 1984 and his PhD in chemical engineering from Virginia Tech in 1999. Dr York has taught courses in engineering problem solving and design, chemical engineering and chemistry. Dr York has also designed and implemented a number of design-build projects and engineering laboratory experiences for first-year engineering students at VA Tech. Dr York is a member of ASEE and the American Chemical Society.

Address: Engineering Education Department, 331-B Randolph Hall, Virginia Tech, Blacksburg, VA, 24061 Phone: 540-231-9783. Email: syork@vt.edu

Tamara Knott, Virginia Tech
Introduction and background

Until the 2001-2002 academic year, the department of Engineering Education (EngE) at VA Tech used no course management software in the administration of its freshman program courses (EngE 1024 and EngE 1114). All class handouts, worksheets, quizzes and tests were distributed to students via paper copies, which was a labor intensive and expensive process for the department. When the course management software Blackboard became available, a few EngE faculty members began using it to post assignments and other pertinent course information. This assured student’s ready access to these materials through their own Blackboard course link and instructors benefited from reduced copying and in-class distribution time as compared to paper copy distribution. Posting course materials in this manner using Blackboard is now the standard means for distribution of handouts, homework assignments and lecture notes within the EngE department and is widely used for the same by departments throughout the VA Tech campus.

While the Blackboard system is successfully used by faculty as a means site for providing course information and assignments to students, many of the other Blackboard functions such as testing, grading forms and discussion groups were deemed too cumbersome by most faculty to merit regular use in large classes and/or multiple sections of the same course. Therefore, the in-class testing capabilities of the software were never explored by EngE.

During the second semester of the first-year engineering course series (EngE 1114) students are taught algorithm development and flowcharting (using Visio); programming (using MatLab); and CAD modeling (using Inventor), in an overall context of engineering design. During the summer of 20051 a small, unpublished study used the assessment tools of Blackboard to conduct multiple choice quizzes and tests and used the drop box function to collect MatLab programs and Inventor CAD files during in-class student assessments. The recent enactment of a student laptop computer requirement by the VA Tech College of Engineering and the advent of wireless internet access across campus have made it possible for students to use their own computers during these in-class assessments. The summer study demonstrated that Blackboard was useful in such utilities as the drop box and was used with limited success for in-class assessment.

In paper testing conducted in previous years students were required to write MatLab programs by hand as a workout problem solution, in addition to a multiple choice section (opscan-paper). The EngE faculty never liked this method of testing programming skills, but no acceptable alternative could be conceived. Similarly, Inventor skills were simply not tested during an in-class exam because of the impossible logistics of generating, collecting and grading the student CAD files.

The results of the preliminary study cited suggest that online testing using Blackboard in the classroom is more accurate and effective in testing student’s MatLab and Inventor skills. The preliminary study also showed that an in-class multiple choice exam using student laptop computers can be accomplished for a small group of students, but it was also found that Blackboard can be cumbersome to use for both students and faculty; Blackboard has numerous compatibility issues with commonly used software; Blackboard can be frustrating to faculty who wish to exercise more control over their test
format and question selections; and testing using laptops meets with a certain amount of student resistance.

The current study involves a total of 300 VT freshman engineering students distributed across 5 sections of EngE 1114 during the fall of 2005. All students that are registered in the course are required to have a functioning laptop computer with wireless internet access. Additionally, all EngE 1114 students are required to have working copies of MatLab and Inventor on their laptops. Every student has a Blackboard link on their VT homepage that connects them directly to their course web-pages.

The current study involves two tests administered in EngE 1114 during fall 2005. To put the tests into context and organize the discussion, the course content and material covered on each test will be presented first, followed by a chronological discussion of test one and test two formats and scores. Next, the results of a student survey on testing practices will be presented and discussed. The paper will conclude with a summary of the faculty impressions of online testing for this type of material and comments concerning usefulness of Blackboard as a tool for conducting online student assessments.

Course structure and learning objectives

EngE 1114 is a two credit introductory engineering course that covers a diverse set of topics. Upon course completion it is hoped that students can successfully:

- Create a flowchart for an algorithm (using Visio)
- Apply basic programming techniques to model an algorithm (using MatLab)
- Create CAD drawings (using Inventor)
- Use the engineering design process
- Apply basic project management tools
- Create semi-rough, by-hand multiview and isometric sketches
- Interpret working drawings

Because many of the students are unfamiliar with any programming language when they enter the course, the flowcharting and programming material is covered at the onset of the course. The focus on flowcharting and programming uses an in-house set of documents that we have used for a number of years with great success. These documents are posted to Blackboard, and many of the classes conducted during the flowchart-programming portion of the course are laptop sessions, where an in-class programming exercise is performed.

The laptop computer and wireless internet requirements make it possible to initiate modify and collect student’s in-class work in ways that have not been possible before. However, faculty must make sure that students develop familiarity with the Blackboard interface and the various submission processes; otherwise valuable in-class testing time is lost by students fumbling to discover the process under testing conditions. This leads to very frustrated students and resentment of the entire online testing process. Faculty can help to assuage this problem by requiring in-class submission of several assignments using Blackboard’s multiple choice questions and image links. Hopefully, this will bring problems to the surface for solution before the pressure of an in-class assessment compounds the problem. Results from the student survey taken at the end of the fall ’05 semester do indicate that student familiarity with the Blackboard testing
interface does improve test scores and lessens student frustrations over what I will broadly describe as “Blackboard/computer problems”.

Discussion of test 1 and test 2

Test 1 format

The first test covered flowcharting and *MatLab* programming. Students were allowed 60 minutes to complete all portions of the test. The test consisted of 15 multiple choice questions and two work-out problems and had the following point distribution:

- Multiple choice= 70%
- Workout 1 (Create a flowchart for a given algorithm)= 20%
- Workout 2 (Write the program for a given flowchart)= 10%

Workout 1 required a hand written solution that was submitted by the students upon leaving the testing room. The multiple choice questions were presented in *Blackboard* and taken using laptops and Workout 2 required the students to write and save *MatLab* code, then submit it to their instructor via the *Blackboard* digital drop box prior to leaving the testing room. All electronic submissions are automatically tagged by *Blackboard* with the submission date and time.

The 15 multiple choice questions for test 1 were randomly ordered for each student and chosen from a pool of 30 questions. At first, several broad categories of questions were created using the pool creation tools and keyword identifiers found in *Blackboard*, however this process was cumbersome and the test canvas in *Blackboard* did not allow questions to be selected or sorted using categories/keywords. *Blackboard*’s lack of flexibility in question selections is frustrating to faculty creating the tests. Eventually, the creation of many pools for a single test was necessary in order to assure that the randomly selected questions issued to each student did in fact cover the range of topics desired and did not overload any single student with too many questions from a single topic area.

Test 1 results and discussion

Being the first time within our department that this large a group (~300) was tested using *Blackboard*, some problems were bound to arise; at least, that is what we figured. To head off some of these problems it was decided that the students would be allowed to make multiple submissions of the multiple choice portion of the test. The *Blackboard* student feedback on questions was limited to the score and students were instructed to complete the multiple choice portion of the test before moving on to the workout problems. Nevertheless there were some problems and glitches, some arising from a lack of familiarity with the *Blackboard* interface, but some problems also in the appearance of images and with *Blackboard* locking-up.

Overall the flowcharting/*MatLab* (test 1) scores seemed representative of scores from previous years, but as some students groaned about the unfamiliar testing format and how it ruined their test scores, the authors decided to see if this was in fact the case. The day following the common time test 1, students were sent an email stating that a required in-class exercise would be conducted in the next class period and they were to bring their
laptops. When the students arrived in class they were given paper copies of 15 randomly selected questions from the test 1 multiple choice pool. Students were instructed to immediately answer these questions, for credit, within 40 min. No further explanation of the second multiple choice offering was made. All students were required to take the partial re-test, regardless of their grade on the electronic submission. Finally, all students were sworn to secrecy concerning the partial re-test so that no unauthorized studying would benefit later classes. The authors consider the two testing conditions for the test and re-test to be a fair comparison of scores that may result from application of the two testing methods. The paper and online testing results are summarized in Table 1 below.

<table>
<thead>
<tr>
<th></th>
<th>Mult. Choice on line</th>
<th>Mult. Choice paper</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average Score</strong></td>
<td>67%, s=11.5</td>
<td>69%, s=8.5</td>
</tr>
</tbody>
</table>

While a few students did dramatically improve or lower their test 1 score when given the re-test on paper, the average score did not see a significant change. Overall the slight increase in average score is consistent with the expected average increase if each student were instead offered another online attempt. The clear conclusion from examining the average scores and the standard deviations (s) is that little, if any, testing bias is introduced by offering the test online versus the more traditional paper/opscan method.

Comparable time limits were place upon both paper and online test offerings. Observation of student completion times and the number of attempts completed indicate that students do not take the online versions more quickly, as might be expected. This was especially true during test 1 as students were less familiar with the interface and the very concept of taking an in-class test online. The average number of test attempts per student more than doubled for test 2, though the number and format of the questions was identical. Therefore, somewhat of a learning curve for online test taking should be expected by faculty wishing to conduct this type of testing. Frequent online quizzes and in-class submissions help to develop student familiarity with Blackboard and lessen the time required to take an assessment.

As previously mentioned, two workout problems were also required from students for test 1. Workout 1 consisted of a hand-written flowchart and is not relevant to this paper. Workout 2 consisted of a flowcharted algorithm that students had to code in MatLab and submit via Blackboard’s digital drop box using the wireless internet connection. The drop box function worked very well except in those few cases where students failed to properly send (not just save) the file. Files were tagged by Blackboard with the date and time submitted and students receive a conformation message when the file is successfully sent.

Using the drop box function to grade the MatLab files from within Blackboard had advantages and disadvantages for both student and faculty. From the faculty standpoint, the inability to search, sort or order the files by anything other than the time index is frustrating. The Explorer “ctr-F” was the only means discovered for searching for specific file names inside the Blackboard environment. However, files may be opened and run in MatLab directly from the drop box window, this made running and reviewing
the workout 2 code submissions quite fast and easy. I would estimate that I typically spent no more than 1.5 min correcting most submissions and that the same grading task performed by hand would have taken three times longer. Code that obviously followed the flowchart structure, ran and gave the right output were able to be rubber stamped “correct” in about 15 s. However, adding corrections and comments to the student’s without paper to write on made feedback difficult.

Students also reaped benefits and gave up some things when taking the electronic test. The largest student benefit is that they may create and de-bug their code on their laptop, not write it by hand as done in the past. Writing programming code by hand is not a representative or realistic model for most programming situations and has never been considered desirable within our department. The advent of laptops and wireless within our classrooms has much improved this testing practice. Not surprisingly, a comparison of these data to previous year’s tests showed that students testing online scored an estimated average of 25% higher on programming workout problems.

The cost to students for this advantage is the lack of feedback received on corrected tests. Items graded within the drop box window are not easily marked, commented on or returned to students. Printing the submissions out is one solution, but this drastically increases the time and paper required to grade. The multiple choice portion of the test allows several levels of feedback, but for randomized tests chosen from multiple test pools the options are limited. The configuration of the multiple choice portion of test 1 did not allow Blackboard to reveal the correct answers, only the students score on the questions and the current class average on the same. The test feedback options were structured to prevent wide scale guessing over a large number of questions in order to discover the correct answers for final submission (students have figured this angle out from previous online quizzes). Consequently students receive rapid scoring of their online tests, but never see the questions they missed or the correct answers. Student found the lack of feedback from the online tests unsatisfying, though it does serve to protect the integrity of the question pools for future testing.

Test 2 format

Test 2 covered a much wider range of topics than test 1. Topics covered were the design process; project management; Inventor skills; sketching and print reading. The test format was the same as test 1, but students were allowed 2 hours to complete the exam because students required more time to complete the sketching workout problems. The point distribution is below:

- Multiple choice = 70%
- Workout 1(create a multiview sketch by hand) = 12%
- Workout 2 (create an isometric sketch by hand) = 12%
- Workout 3 (Create a CAD model using Inventor) = 16%

The multiple choice questions covered project management, design and engineering graphics. The same format as used for test 1 was followed for test 2. Students could submit multiple attempts, but only their score is revealed following submission. It should be noted by faculty and students using Blackboard assessments that Blackboard records only the most recent test submission for grading purposes. Neither the high score nor the average score will be retained for any one student. It is important that students realize this
before starting a Blackboard online test. Students do not want their grade to result from a hurried final submission when a previous attempt may have resulted in an acceptable grade.

The testing mentality and strategy is different for students when taking an online test. Nevertheless, students appeared much more comfortable with the practice on the second go around. Students were offered the option to take a paper/opscan version as the test 2 was about to commence, but surprisingly there were no student takers of this offer. All students ultimately chose to take the online version when confronted with the option.

Workout problems 1 and 2 were hand drawn sketches and so not relevant to the discussion here, but workout problem 3 was an Inventor model that had to be created and submitted by the students to demonstrate prowess with the Inventor software. This was the first time that Inventor problems had been included on any in-class examination, as it was just not feasible before. The Inventor problem (model + dimensions) was presented to the students via paper, but their solution (CAD file) was submitted via the Blackboard drop box.

Test 2 Results and discussion

Students were clearly more comfortable with the idea of online testing by the second test. Students were offered the option of taking a paper test, but despite some previously stated positions by the students, not a single student elected to take the paper version. This was a somewhat surprising turn of events and runs a little contrary to survey responses from the same group.

The average test 2 multiple choice score was 64% with a standard deviation (s) of 10. A comparison of similar paper test scores is not possible since no student elected to take to that version. Based on a comparison of same and similar questions posed in previous years on paper/opscan tests, student performance can be reasonably judged as comparable to student performance on this material in years past.

Grading was again easily and quickly accomplished. Inventor files submitted by students via the digital drop box were readily open in Inventor and easily graded. The same problems mentioned previously with test 1 remain. Apparently, Blackboard does not contain any tools that allow instructors to comment and then return graded work electronically. The lack of instructor feedback is a serious failing in the use of purely electronically graded assessments. The functionality that is truly needed to grade such e-submissions is the ability to use an electronic pen to write on the student’s submission and then save and return the work to the student.
Survey results and Instructor comments

Student survey results

Students were asked 6 questions about their online testing experience and test method preferences and invited to make a recommendation to students that must take a Blackboard, online test. The questions and student responses are given below.

1. I prefer to take the multiple choice portion of the test multiple times on Blackboard versus only once on paper.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>33%</td>
<td>35%</td>
<td>12%</td>
<td>16%</td>
<td>4%</td>
</tr>
</tbody>
</table>

2. I prefer to take the multiple choice portion of the test on Blackboard even if only one submission is allowed.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>6%</td>
<td>18%</td>
<td>11%</td>
<td>28%</td>
<td>36%</td>
</tr>
</tbody>
</table>

3. How many times did you submit the multiple choice portion of Test 1 on Blackboard?

<table>
<thead>
<tr>
<th>Once</th>
<th>at least twice</th>
<th>at least three</th>
<th>at least four</th>
<th>at least 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>27%</td>
<td>26%</td>
<td>16%</td>
<td>15%</td>
<td>17%</td>
</tr>
</tbody>
</table>

4. Which attempt on test 1 yielded the highest score?

<table>
<thead>
<tr>
<th>The first attempt</th>
<th>The last attempt</th>
<th>One of the middle attempts</th>
</tr>
</thead>
<tbody>
<tr>
<td>19%</td>
<td>51%</td>
<td>30%</td>
</tr>
</tbody>
</table>

5. How many times did you submit the multiple choice portion of Test 2 on Blackboard?

<table>
<thead>
<tr>
<th>Once</th>
<th>at least twice</th>
<th>at least three</th>
<th>at least four</th>
<th>at least 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>15%</td>
<td>17%</td>
<td>21%</td>
<td>42%</td>
</tr>
</tbody>
</table>

6. Which attempt on test 2 yielded the highest score?

<table>
<thead>
<tr>
<th>The first attempt</th>
<th>The last attempt</th>
<th>One of the middle attempts</th>
</tr>
</thead>
<tbody>
<tr>
<td>18%</td>
<td>56%</td>
<td>26%</td>
</tr>
</tbody>
</table>

From the student survey it seems apparent that students gained confidence and skill in taking the online tests as they gained experience with the new method. During test 1, 27% of the students managed to complete and submit only a single test attempt. During test 2 all but 5% of students made multiple submissions. This is no doubt due to students mastering the mechanics of taking an on line test and developing their own test-taking strategies. Many students reinforce this conclusion from the comments made at the end of the survey.
The survey also reveals that a significant portion of the students prefer the online testing format, though they strongly prefer that multiple submissions be allowed. However, no student opted for a paper test over the Blackboard version when presented with the choice between the two. This stands in apparent contrast to the survey results, but perhaps the students feel a certain comfort level with using the laptop that they have come to rely upon as a tool.

The final survey question asked students to offer online testing advice to future students, their primary comments are summarized below.

- Use a mouse to navigate the site and resize windows for viewing.
- Bring a fully charged battery and your power cord.
- Set a performance goal and stop when you reach it; don’t be greedy.
- Take the test once quickly, take a second time carefully, and stop unless time is abundant.

As you can see from the comments, the strategy of test taking is most certainly affected when an online test is presented. Therefore it is advised that students be allowed a little extra time to adjust when the transition is first experienced, but ultimately the time required and the student scores appear very comparable between paper and online testing using Blackboard.

Instructor comments and conclusions

The authors still have mixed feeling about using the current version of Blackboard for online testing. The “klunkiness” of Blackboard is often frustrating when creating and when giving a test. Blackboard experiences many screen lock-ups, which are cumbersome to clear and compatibility problems with other common software packages cause many “surprises” on the student monitors. Many students using browsers other than Internet Explorer experience the problem of being locked out by Blackboard, which requires instructor action to clear. When in-class test are given to large numbers of students, the lock-outs can be a very difficult problem to solve under the test conditions.

The worst of Blackboard’s shortcomings is the lack of ability to comment on and return work submitted by students. Grading is easily accomplished, but the feedback loop to the student is not properly closed and this must be addressed. The ability to use Blackboard’s drop box to collect work generated on the student’s laptop (i.e. MatLab and Inventor files) is a big positive step in assessing student performance. While grading these submissions is rather easy, the lack of ability to easily comment and offer students feedback is troubling. Perhaps later versions of Blackboard will seek to resolve these shortcomings. If so, Blackboard may eventually become an indispensable tool for conducting certain student assessments. Until then, however, the authors feel that Blackboard is not yet ready to be used in the manner first attempted in this study; the in-class testing of large numbers of students who are tested using their own laptop computers to complete the test.

1 SC. York, VA Tech unpublished work, 2005
2 Jean Kampe, VA Tech in-house document, Flowcharting, 2005
3 Jean Kampe, VA Tech in-house document, Programming, 2005