

Interesting Times: Relation between Test Times and Student Performance in Online Courses

Hedayeh Samavati
Indiana University Purdue University Fort Wayne

Carolyn Fabian Stumph
Canterbury School and Indiana University Purdue University Fort Wayne

David A. Dilts
Indiana University Purdue University Fort Wayne

Many instructors are now using online technology for the assessment of student learning. However, there are still relatively few empirical studies of the technology's implications for student learning. In this study, the authors use student performance data from over two hundred students in principles of microeconomics to assess the correlation between student performance and the duration of an online exam. Conclusions were drawn regarding the "optimum" time allotted for multiple-choice, computer-based course examinations. The results should be of interest to all educators who are involved with distance learning programs generally, and those instructors who administer non-proctored online exams, specifically.

INTRODUCTION

Online course offerings in many academic disciplines, in general, and economics in particular, have experienced enormous growth over the last decade. According to one research report (Allen and Seaman, 2006), in the fall 2005, over 2.6 million students in the United States were taking at least one online class. Using the results of two different surveys, Harmon and Lambrinos (2008) reported that in the interval from 1997 to 2000 the number of institutions offering online economics courses increased 400% and the number of economics courses offered increased by 373%. The explosion of number of courses offered through distance learning is not limited to the United States. Course offerings in Canadian, British, Australian, and many other countries' universities follow a similar trend. For example, based on the 2005 Canadian Internet Use Survey, McKeown and Underhill (2008) concluded that, "Just over one-quarter (26%) of those going online for education-related reasons reported doing so for distance education, self-directed learning or to take correspondence courses. The rapid growth of internet-based courses, which has occurred over a relatively short period of time, has left many instructors puzzled regarding various aspects of online teaching. Given the short time since the emergence of online course offerings, there is a paucity of research concerning effective teaching methods and appropriate assessment techniques in an online environment.

An important aspect of online course offerings is the assessment of student learning. One major concern is whether online exams are given in a proctored or a non-proctored setting. Given that in many cases proctored exams are not an option, non-proctored exams have frequently been used. A particular question that arises regarding non-proctored online exams concerns the “optimum time” for the duration of such exams. This is an interesting question because it is generally observed that, in a traditional classroom setting, high-performing students who arrive well prepared for a closed-book exam often finish well ahead of the stop time. However, students who arrive poorly prepared have also been observed to finish early because they have no resources from which to draw to answer the questions. Of course, exceptions exist for both groups of students regarding the time needed to complete the test. One can conjecture that the same dynamic operates when testing is conducted over the internet. The dilemma of the “optimum” exam time is exaggerated in an online environment, however. That is, on the one hand, if too little time is allocated for finishing the test, it can increase student’s anxiety level and thus hamper performance. On the other hand, an instructor may be concerned that allowing too much time for an exam or placing no time limits on an exam at all may foster inappropriate behavior of some sort or another. For example, in the case of non-proctored exams, there remains concerns regarding misconduct where, all other things the same, the longer exam time makes it easier to engage in inappropriate activities. While different cheating schemes are of major concern, other activities such as “quiz-to-learn,” which is not cheating per se, but nevertheless has detrimental effect on student learning and thus is also of concern. “Quiz-to-learn” is defined by Brothen and Wambach (2004) as attempting assessments without good preparation. This practice contradicts a basic tenet of the learning literature which suggests that the best approach for learning is for students to use a “prepare—gather feedback—re-study” strategy (Thomas and Rohwer, 1986).

Testing hypotheses regarding the relation between test score and length of an exam in a traditional face-to-face course environment where examinations are paper-based are nearly impossible due to lack of suitable data. However, for online courses which take advantage of computer-based testing, the relation between exam duration and performance is, fortunately, a testable hypothesis. This is because a computer-based testing environment facilitates the collection of precise data regarding the duration of the test and performance.

It is the purpose of this study to investigate the relationship between student performance and the time spent to complete the test. Online learners were separated into high-performing and low-performing groups. Length of time spent taking timed online examinations was then analyzed in order to discover whether a significant difference existed in the time used for these groups of students. Conversely, students were divided into fast and slow finishers according to the amount of time used to complete a test. Performance of these groups in terms of test scores was then examined. Conclusions are drawn regarding the “optimum” time allotted for multiple-choice, computer-based course examinations. The results should be of interest to all educators who are involved with distance learning programs generally, and those instructors who administer non-proctored online exams, specifically.

LITERATURE REVIEW

Investigating the relation between the time-limit for online quizzes and student learning, Brothen and Wambach (2004) compared performance of students in an undergraduate psychology course. The students were divided into two groups: one group took their online quizzes with no time-limit imposed and the other had a 15-minute time limit on such quizzes. The quizzes in the study were used to prepare students for in-class examinations. One reason to be concerned with the time-limit for online quizzes was the “quiz-to-learn” behavior. In other words, according to Brothen and Wambach, students may use the quiz time to look up answers in an online glossary or textbook just to complete the quiz. Their findings indicated those students who took online quizzes with no time-limit did poorer on exams. The authors concluded that time-limit may have reduced usage of the “quiz-to-learn” strategy and convinced students of the need to learn in order to be able to answer questions on their own. Therefore, they recommended that, “instructors set *reasonable* time-limits on computerized quizzes” (Brothen, and Wambach, p. 64,

emphasis added). Once again, it is important to discern the “reasonable” amount of time for an online quiz or test.

Online quiz time-limit has been of concern to other scholars. Daniel and Broida (2004) found that when supplementing instruction with quizzes, if the time-limit for a 10-item exam was reduced from 15 minutes to 7 minutes, then the incidence of students looking up answers during the quiz using a glossary declined. Other undesirable student behaviors that fall under the category of cheating which occur when students have no time-limits on quizzes or have more time than they really need to take the quiz were also of concern to the authors. In their research, Daniel and Broida (2004) asked their students to anonymously self report this sort of behavior. Technologically savvy students reported a variety of strategies for cheating including printing and sharing of quizzes and working in groups. These authors found that assigning questions from a larger test bank and decreasing the amount of time allowed for quizzes were effective at discouraging this behavior.

When it comes to assessment, of course, there are those academics whose chief concern has been the possibility of cheating in an online environment. As a result, the proctored exam is determined to be the best approach (Edling 2000; Rovi 2001; Deal 2002; Harmon, and Lambrinos 2008). In support of this conclusion, there is an increasing body of literature that points to the presence of some level of cheating for the un-proctored online exams (e.g., Harmon and Lambrinos 2008). Other research, however, found that the majority of students prefer online exams and quizzes to the paper ones (Barkley 2002). These findings will undoubtedly give rise to further research concerning the costs and benefits of proctored versus non-proctored exams and many universities may well conclude that all online exams should be proctored. But the costs of doing so in terms of time, inconvenience and expense, while potentially manageable, are still significant. Therefore, efforts to improve the experience and outcomes of non-proctored exams for faculty and students are certainly worthwhile.

While there is evidence in the literature that shorter time periods for quizzes and exams may improve learning outcomes, there is still a concern that short exams will frustrate students and could be deleterious for the achievement of learning objectives. Therefore, the question described above regarding how much time will allow the students enough time to fully demonstrate their knowledge while discouraging the undesirable behaviors is examined here.

DATA AND METHOD

Data were collected from an online section of principles of microeconomics course in the fall of 2008 at a single public four year university in the Midwest. There were 78 students registered in the course who, as part of requirements of the course, had to take three online exams. These online exams constituted 75 percent of course grade. In-class exams at the institution are traditionally seventy-five minutes in length for these courses. The instructor had designed the distance learning course so that the online students also had seventy-five minutes to complete their exam. All exams were multiple choice although the number of questions for each exam varied somewhat-- exams consisted of fifty-five to fifty-seven questions. Students were allowed only one attempt for each exam and online exams were administered through the university’s course management system, Blackboard Vista 4. The two variables of interest were the student’s score and the time it took the student to finish his/her exam. While collection of this data is difficult to obtain in a traditional course setting, the submission times for the online exams are recorded to the second (by Blackboard in this case). There were 217 observations because some students missed their tests. The exams were non-proctored so the students were free to use any resource such as textbooks, lecture notes, or any other material they chose. In addition, while seventy-five minutes was the maximum allowable time, the students could finish and submit their exam any time prior to that. Table 1 presents the descriptive statistics for the two variables, exam score, or Grade, and Time, which is the time used to complete and submit the test.

TABLE 1
DESCRIPTIVE STATISTICS FOR GRADE AND TIME

	Introduction to Microeconomics	
	Grade	Time
Mean	74.14	57.44
Median	75.5	62
Mode	75.5	75
Q1	65	49
Q3	85.48	71
Minimum	31.5	3
Maximum	100	75
Sample size	217	217

The mode shows that many students took the entire seventy-five minutes to complete the exams. The mean (57.4 minutes) and the median (62 minutes) finish times are smaller than the mode of 75 minutes which provides support for the observation that the distribution of finish time is negatively skewed. The results show that the majority of students used over an hour to complete their test.

Given the data and using the variable Grade, students were divided into two groups—high-scorers and low-scorers. Next, using variable Time, another two groups of students, fast-finishers and slow-finishers, were identified. Following the approach used by Ferguson et.al (2002), high-scorers were defined to be those students who scored above the median grade and low-scorers were those whose grades were below the median grade (75.5% in this case). On the other hand, fast-finishers were defined as the first 25% of students who finished the exam. For example, this group would consist of students who finished the exam in less than forty-nine minutes (the first quartile for Time). While slow finishers were those who were the last 25% of students to finish their test. In other words, this was the group of students who took seventy-one minutes (third quartile for Time) or longer to complete the exam.

To ascertain whether students used their allotted time of 75 minutes in an “appropriate” manner while taking an unproctored online exam, we test for the evidence of a relation between grade and the time students used to complete the test. Both correlation analysis and t-tests were used to do the statistical analysis of the data. Using correlation analysis, we seek the answer to the following questions. 1) Is there any relation between Time and Grade for all students? 2) Do high-scorers tend to finish early or do they take more time to complete their test? 3) How about the low-scorers? Using the t-test procedure, we test to discover whether there exists a significant difference between the average grades of fast-finishers and slow-finishers. Similarly, a t-test will be used to examine if there is a difference between average amount of time used to complete the test by high-scorers and low-scorers. Statistical findings are reported in the next section.

RESULTS

The results of correlation analysis for all students, the high and low-scorer groups, as well as for fast and slow-finishers are reported in Table 2.

Findings reported in Table 2 show that there is a significant relation between grade and the amount of time students took to complete their test. This is true irrespective of how students are grouped together. While this result may not be remarkable in and of itself, it is interesting to note that the results presented in Table 2 show that for all students, low-scorers, as well as for fast-finishers this correlation is positive, but the correlation is negative for the high-scorer as well as the slow-finishers group. This result suggests that for the group of students who spent the least time (49 minutes or less) to complete and submit their test, scored higher, the more time they used. This makes sense because even when a student is well-prepared and is generally a fast reader and thinker, he/she needs to take the time to analyze in order to

correctly answer the questions. In other words, the finding shows that rushing through an exam does not payoff even for the best students. Also, those students who were at the bottom of the grade distribution scored higher when they used more time to complete their test. The more interesting result though is the observed negative correlation for the group of high-scorers and slow-finishers. That is, those students at the top 50% of the grade distribution tended to take *less time* to finish their tests. In addition, those students who took the most time to submit their tests --slow finishers--(71minutes and more), *scored lower*. This result is inconsistent with inappropriate behaviors during the allotted time for un-proctored online exams (i.e., prolonging the test-time to pursue potential cheating strategies). This is true because those who did use extra time to finish their tests tended to score lower. Also, the evidence showed that the high scorers tend to take less time to submit their test.

TABLE 2
CORRELATION BETWEEN GRADE AND TIME USED TO COMPLETE A TEST

	Observations	Corr. Coeff	t-statistic
All Students	217	0.135	1.997**
High-Scorers	55	-0.304	-2.32**
Low-Scorers	62	0.473	4.16*
Fast-Finishers	57	0.414	3.37*
Slow-Finishers	56	-0.316	-2.45*

Note: * indicates significant at the 1% level
** means significant at the 2.5% level

In addition to the correlation analysis, the mean finish time for the high and low scorers was calculated and a t-test was performed in order to determine if there was a statistically significant difference between the average finish times of the two groups. Those results are presented in Table 3.

TABLE 3
COMPARING AVERAGE TEST TIME FOR HIGH AND LOW SCORERS: T-TEST RESULTS

	High scorers	Low scorers
Mean	55.71	52.71
Variance	214.03	388.01
Observations	55	62
Pooled variance	306.3141	
t statistic	0.925	
P value one-tail	0.178	
t critical one-tail*	1.290	

* Critical value for t with 115 df and 10% significance level

Table 3 indicates that even though the high scorers took an extra three minutes, on average, to finish their tests, this observed difference in exam completion time was not statistically significant at any reasonable level (p-value for the upper tail test is nearly 18%).

The results for the fast and slow finishers are reported in Table 4.

TABLE 4
COMPARING AVERAGE TEST SCORE FOR FAST AND SLOW FINISHERS: T-TEST RESULTS

	Fast finishers	Slow finishers
Mean	72.12	73.62
Variance	381.22	149.06
Observations	57	56
Pooled variance	266.1869	
t statistic	-0.4876	
P value one-tail	0.3134	
t critical one-tail*	1.290	

* Critical value for t with 111 df and 10% significance level

CONCLUSIONS

The results of the study analyzing the relation between the test-time and test-score in an online course environment provide some answers to the questions posed earlier in this study. The correlation analysis yielded interesting results. It showed that the high scorers' time and score had a negative correlation. This may indicate that the high scorers came prepared to the exam and did not necessarily need the extra time. In addition, for the group of students who took the full time or close to it to complete their exam (slow-finishers), the test score was negatively correlated with the time. This suggests that, in an online environment, merely taking extra time to complete a test does not help improve one's grade. In fact, it tends to result in a lower grade. On the other hand, the low-scorers showed a positive correlation between their scores and test time. This suggests that for the group of students who were not prepared for the test, their score went up the longer they take to finish the exam. However, even though these students, by using extra time, increased their score, they did not succeed to improve their score to the degree where they can avoid ending up at the bottom of the grade distribution. Also, the results of the t-test analyses revealed no significant difference between the average grades of those students who used additional time (slow finishers) and those who used less time (fast finishers) and neither was a significant difference between average time used for high and low scorers. Thus, these findings cast doubt on the conjecture that the unproctored online exams are grounds for cheating. More accurately, the results show that even if the students used extra time in order to engage in any misconduct, the effort did not produce any meaningful advantage to merit much professorial concern.

Also, to improve the online assessment outcomes, our findings support the recommendation that one may consider reducing the time allotted for an online exam to 65% to 75% of that which is allowed for a comparable proctored in-class test. For example, for an exam that would have been allowed seventy-five minutes as a proctored in-class exam, should be allotted a maximum of one hour for a non-proctored online exam. This suggestion is also in line with that used by Daniel and Broida who shortened the quiz time by approximately 47% from fifteen to seven minutes.

REFERENCES

- Allen, I., Seaman, E. & J. (2006). *Making the Grade Online Education in the United States, 2006*. Massachusetts, Sloan Consortium.
- Barkley, A.P. (2002). An Analysis of Online Examinations in College Courses. *Journal of Agricultural and Applied Economics*, 34.3, December, 445-458.

Bangert-Drowns, R.L., Kulik, C.L.C., Kulik, J.A. & Morgan, M. (1991) The instructional effect of feedback in test-like events. *Review of Educational Research*, 213-238.

Brothen, T. & Wambach, C. (2004). The Value of Time Limits on Internet Quizzes. *Teaching of Psychology*, 62-64.

Daniel, D.B., & Broida, J. (2004). Using Web-based Quizzing to Improve Exam Performance: Lessons Learned. *Teaching of Psychology*, 207-208.

Deal, W. F. III (2002). Distance learning: Teaching technology online. *Technology Teacher*, 61, (8), 21-27.

Edling, R. J. (2000) Information technology in the classroom: Experiences and recommendations. *Campus-Wide Information Systems*, 17, (1), 10-15.

Ferguson, K.J., Kreiter, C.D., Peterson, M.W., Rowat, J.A. & Elliott, S.T. (2002) Is That Your Final Answer? Relationship of Changed Answers to Overall Performance on a Computer-Based Medical School Course Examination. *Teaching and Learning in Medicine*, 20-23.

Rovi, A. P. (2001) Online and traditional assessments: What is the difference? *Internet and Higher Education*, 3, (3), 141-51.

Harmon, O.R., & Lambrinos, J. (2008). Are Online Exams an Invitation to Cheat? *Journal of Economic Education*, 116-125.

Thomas, J.W., & Rohwer, W.D. (1986). Academic studying: The role of learning strategies. *Educational Psychologist*, 19-41.

McKeown, L. & Underhill, C. (2008). Learning online: Factors associated with use of the Internet for education purposes. Science, Innovation and Electronic Information Division Statistics Canada <http://www.statcan.gc.ca/pub/81-004-x/2007004/10375-eng.htm> accessed March 11-2011.