

ONLINE VERSUS PROCTORED TESTING: HOW DO THEY COMPARE?

Thomas Lee Hench

ABSTRACT

A major challenge in offering a successful online course is the deciding upon the method by which students are tested; namely either online tests or proctored tests in classrooms or specialized learning centers. When faced with this decision, instructors frequently ask the question “Is online testing equivalent to proctored testing occurring in either classrooms or specialized learning centers?”. This paper provides a positive answer to this question by first setting forth two criteria for equivalency. These criteria are then evaluated using quantitative (i.e., how do test scores for online and proctored students compare) and qualitative (i.e., how do testing environments compare) data resulting from both online and proctored testing situations. In particular, the first criterion, evaluated via a statistical analysis of examination scores, yields the conclusion that student scores on online examinations are found to be quantitatively equivalent to the scores of those students taking proctored tests. The second criterion, when examined by an analysis of student preferences obtained via a survey of test takers, leads to the conclusion that meeting student learning style preferences (such as taking tests at home, the need for a quiet environment, the flexibility to test at anytime of day) results in a qualitative equivalency of testing environments (and may lend further support to the observed quantitative equivalency of online examinations). The paper concludes with a discussion of online testing security and an alternative explanation of the observed quantitative test results in terms of breaches in online security. A proposed test designed to identify when such breaches may have occurred is outlined and a discussion of how the proposed test applies to the results in the paper is presented.

KEYWORDS

Teaching and learning in science, distance learning, online testing

INTRODUCTION

When offering an online distance learning course, one important question to answer is where to test; either online or in a proctored testing center or classroom. Online testing offers several advantages such as a potentially more comfortable and familiar testing environment and enhanced flexibility as to where and when the testing occurs and can reach those students unable to be on campus (Williams, 2000). These advantages may benefit students in providing a more closer match to individual learning and testing styles (Admin, 2002). However, inherent in online testing are disadvantages associated with issues of security, issues that proctored testing centers minimize while also providing a more structured test-taking environment (White, 2000). The possibility of cheating on online exams casts serious doubts upon the accurate measurement of student performance. Given the different natures of and issues between online and proctored testing, the answer to the question of “where” to test depends first upon determining how equivalent online testing is to proctored testing in terms of student performance and the physical testing environment. The minimum condition for usage is that online testing be at least equivalent to proctored testing. If equivalent, online testing may and should be utilized and may possibly effect a positive educational change. If, on the other hand, it is determined to be not equivalent, online testing should not be considered an option. Therefore, to determine the equivalency between online and proctored testing, evaluation criteria must be established and applied. After equivalency is established, security issues may then be addressed.

EQUIVALENCY CRITERIA

To determine the equivalency of online to proctored testing, this paper proposes the following two criteria –

Equivalency Criterion 1: Student performance in the online testing environment be quantitatively equal to student performance in proctored testing environment.

Equivalency Criterion 2: Student preferences for the online testing environment be both quantitatively and qualitatively equal to student preferences for the proctored testing environment.

In particular, these criteria should be evaluated separately for each course in which online testing is a potential option, for the nature of some courses may preclude the use of online testing. However, if for a specific course both criteria are met, then online testing may be deemed to be a viable option. This paper describes the evaluation of these equivalency criteria for an undergraduate astronomy course taught by the author at the author’s institution. In addition, the course is intended for non-science majors and the sections in this study consisted of two groups of students, one group taking the course completely online and the other group taking the course in the hybrid format (Hench, 2003). Both groups were given a two-hour time period to complete each exam and the course itself extended over a period of sixteen weeks.

EVALUATION OF THE CRITERIA

For the aforementioned course, equivalency criterion 1 is evaluated by a statistical analysis (t-test, $p = 0.05$) of student examination scores taken in both the online and proctored testing environments. In this analysis, the scores of each group are compared to the historical average score of examinations taken in a proctored testing environment (Table 1). More specifically, proctored examinations are Exams 1, 3, and 5 (non-shaded) and online examinations Exams 2 and 4 (shaded). All scores shown are out of a maximum of 50.

Table 1. Comparison of online examination scores with the historical mean.

	Sample Mean	Standard Deviation	Number	Historical Mean	t value	Critical value	Significance compared to the Historical Mean
Hybrid/Exam 1	36.2	10.1	18	38.4	-0.90	2.11	Equal to
Hybrid/Exam 2	39.0	7.6	20	37.6	0.80	2.09	Equal to
Hybrid/Exam 3	33.0	10.0	14	39.3	-2.27	2.16	less than
Hybrid/Exam 4	40.0	8.7	12	36.5	1.33	2.20	greater than
Hybrid/Exam 5	33.3	4.9	14	40.0	-4.93	2.16	less than
Internet/Exam 1	35.5	10.7	23	38.4	-1.27	2.07	less than
Internet/Exam 2	43.3	4.4	22	37.6	5.94	2.08	greater than
Internet/Exam 3	37.4	9.9	19	39.3	-0.81	2.10	Equal to
Internet/Exam 4	42.1	7.2	21	36.5	3.48	2.09	greater than
Internet/Exam 5	39.9	6.1	19	40.0	-0.07	2.10	Equal to

Firstly, from examination of Table 1, it can be noted that the proctored examination scores for both groups are either equal to or marginally less than the historical mean. This indicated that the students comprising the two groups used in the study were comparable to students who have taken the course in the past. Secondly and of importance to this study, the online testing scores are found to be either equal to or greater than the historical mean for tests taken in the proctored environment, a result consistent with other studies (White, 2000). Thus, equivalency criterion 1 is met and in some cases exceeded. The second equivalency criterion is evaluated by the administration and analysis of post-course surveys asking students to evaluate preferences and experiences with both the online and proctored testing environments. The survey questions and results are shown in Table 2.

Table 2. Post-course survey questions and response results.

	Not very Important	2	Important	3	4	Very Important	5
1. How important to you was the option provided by online exams to take an exam at a location other than at school? (N=30)				4.0			
2. How important to you was the option provided by online exams to take an exam at any time during the day? (N=30)				4.7			
3. How important to you was the option provided by online exams to take an exam in the surroundings of your choice (at home, at work, a den, a study, a library, etc.)? (N=30)				4.4			
4. How important to you was the ability provided by online exams to give you immediate feedback on how well you did on all or part of an exam? (N=30)				4.3			
5. How important to you was the opportunity provided by online exams to minimize the time between studying for the exam and taking the exam? (N=30)				4.1			
6. How important to you was your comfort level with computers in taking an online exam? (N=30)				4.3			
7. When taking an online exam, you could not ask your instructor any questions. How important to you is the ability to ask an your instructor questions while taking an exam? (N=30)				2.3			
8. Do you prefer exams given online or exams given in a classroom or exams given in a learning center? (N=30)	Online 80%		Classroom 7.0%			Learning Center 13%	
9. Do you prefer computer-based tests or traditional "paper and pencil tests"? (N=27)	Computer-based Exams 85%					Traditional Exams 15%	
10. When compared to exams given in a classroom or learning center, do online exams require a shorter time to complete than exams given in a classroom/learning center, about the same time as exams given in a classroom/learning center, a longer time than exams given in a classroom/learning center? (N=30)	Shorter time 10%		Same time 77%			Longer time 13%	
11. When compared to exams given in a classroom or learning center, do you feel that you did worse on exams given online than on exams given in a classroom/learning center, about the same as exams given in a classroom/learning center, better than exams given in a classroom/learning center? (N=30)	Worse 10%		Same 57%			Better 33%	
12. When compared to exams given in a classroom or learning center, do you feel that cheating is less common on exams given online compared to exams given in a classroom/learning center, about the same as in exams given in a classroom/learning center, more common than exams given in a classroom/learning center. (N=29)	Less common 14%		Same 55%			More common 31%	

In addition, the length of time required to complete both the proctored and online examinations was also recorded and is shown in Table 3. Again, the non-shaded areas represent the proctored environment and shaded areas indicate the online test-taking environment. It should be noted that no student used the allowed total two-hour maximum testing time.

Table 3. Average time for completion of proctored and online examinations.

	Average Time (min)	Standard Deviation (min)
Hybrid/Exam 1	31.8	12.7
Hybrid/Exam 2	52.2	23.7
Hybrid/Exam 3	26.5	10.6
Hybrid/Exam 4	54.4	32.0
Hybrid/Exam 5	33.4	17.7
Internet/Exam 1	33.9	16.5
Internet/Exam 2	53.0	19.7
Internet/Exam 3	34.9	15.1
Internet/Exam 4	47.9	17.9
Internet/Exam 5	36.2	14.7

As indicated by the responses to questions 1 through 5, the flexibility offered by the online testing option is evident and reinforces this as a potential advantage. The comfort level of students with computers is also found to be an important factor in choosing online testing (Question 6), while having the instructor or present during the examination is given a low level of importance (Question 7). Thus, for this group of thirty students, online testing represents a viable means for evaluating their progress in the course.

Regarding student preferences for either proctored or online testing, survey results show a strong preference for online testing and the online testing environment. As indicated by the response to survey question 8, eighty percent (80%) prefer the online environment as compared to twenty percent (20%) for the proctored environment. In question 9, eighty five percent (85%) of all students indicate a preference for computer-based examinations with fifteen percent (15%) expressing a preference for traditional “paper and pencil” examinations as would be offered in the proctored environment. Quantitatively, therefore, the responses for questions 1 to 9 indicate that students prefer the online testing environment when given the choice between an online or a proctored testing environment.

Qualitatively, written comments of students preferring online testing (“... much more convenient, ... more relaxing atmosphere, ... room to think and concentrate better, ... don’t feel like hawk eyes are on you, ... it is a computer generation, ... at ease, no pressure, ... less stressful, ... flexibility, ... typing is more comfortable, ... complete silence for concentration”) indicate that this environment provided, in general, a more comfortable and less stressful testing environment. The comments of those students showing a preference for proctored testing focus mainly on computer issues (“... problems with computers, ... temptation of cheating, ... I hate computers, ... As long as server is up and operating correctly”), in addition to an absence of a comfortable testing environment (“... *at home* there was a lot of distraction”), a need for interaction (“... *be able to* go over exam when all are finished”), or a simple preference (“... more comfortable *with traditional exams*”). When both quantitative and qualitative results are viewed together, equivalency criterion 2 is deemed as met.

Thus, for this course, the online testing environment is found to be equivalent to (and in some instances preferably to) proctored testing environment and therefore provides an acceptable form of testing student progress in the course.

DISCUSSION

In addition to determining the equivalency of online and proctored testing, the previous results indicate both a clear preference for the online testing environment and better examination scores in three of the four online examinations given. These two outcomes may be alternately explained in terms of matching learning style and test-taking preferences and/or lapses in online security. Each of these possibilities is now discussed.

From the quantitative and qualitative survey results for questions 1 to 9 shown in Table 2, the inference may be drawn that the flexibility of the online testing environment better suited students by providing them with a more relaxed testing environment. The better match of test-taking environment and test-taker preferences thus resulted in better test taking and higher scores. This explanation is further supported by the response to survey question 10 and the data illustrated in Table 3. Survey responses show that seventy seven percent (77%) of students felt that they required the same time for completing both online and proctored examinations. A comparison of test-taking times presented in Table 3 shows online test-takers used significant more time in completing examinations. Thus, the perception of taking less time for examination completion while actually using more time suggests a more comfortable and less stressful test-taking environment. In addition, as shown in the responses to survey question 11, one third (33%) of students felt that they performed better on online examinations when compared with proctored situations. This perception of better performance is consistent with the suggestion of a more compatible testing environment and supported by the quantitative examination scores shown in Table 1. Hence, the better scores observed for online examinations may be explained by the more conducive test-taking environment provided by this mode of evaluation.

The student responses to survey question 12 offer another possible explanation for the better performance of online students. As seen, eighty six percent (86%) of the students felt that cheating is the same or more common for online testing when compared to proctored testing. This is consistent with another study found by the author, although the examinations in that study were of the take-home (open book) variety (Williams, 2000). Indeed, thirty one percent (31%), indicated that cheating was more common online, with the lack of monitoring and human nature the most common explanations of why cheating occurs. The results from most of the survey questions in Table 2 may be re-interpreted as indicating the possibility of cheating on online examinations and hence the better performance shown in Table 1. In addition, the additional time taken for examination completion illustrated in Table 3 may also indicate that students used the extra time to look for answers in notes or in the text. Furthermore, the identity of the test-taker may be in question. Hence, the author recognizes that, realistically, some level cheating did occur in both online and proctored environments.

The presence and magnitude of cheating may be suggested by making two assumptions. Firstly, it is assumed that the benefits of the online test-taking environment previously discussed are real. The magnitude of the benefit may be estimated by the difference between the average online exam scores and the average proctored scores for all the students taking part in the study. This difference may be noted as ΔE and is shown in Table 4, along with the standard deviations. Hence, the overall benefit to the students in testing online would be in the range of ΔE plus or minus the online exam score standard deviation, as indicated in Table 4. Secondly, the assumption is made that some of the examination score improvement results from the time spent on completing the examination. Students whose preference is the online testing environment would spend more time taking their examinations and receive higher scores, while those whose preference is the proctored environment would spend less time taking online examinations and realize smaller or possibly negative results. Thus, the time difference between online completion time average and the proctored completion average may also serve as an indicator for student performance. This difference is designated Δt and, along with the standard deviation, would result in the range of completion times as shown in the table.

Table 4. Examination score and time differences for online and proctored environments.

Online Exam Score Average	Standard Deviation	Proctored Exam Score Average	Standard Deviation	$\Delta E = (\text{Online} - \text{Proctored Exam Averages})$	$\Delta E \pm \text{Standard Deviation}$
41	7.0	36	9.1	5.0	+12, -2
Online Completion Time Average (min)	Standard Deviation (min)	Proctored Completion Time Average (min)	Standard Deviation (min)	$\Delta t = (\text{Online} - \text{Proctored})$ (min)	$\Delta t \pm \text{Standard Deviation}$ (min)
52	22	34	13	18	+40, -4

The benefit of testing online for any individual student (ΔE_i) may be expected to fall within the range of score differences given in Table 4; namely -2 to +12. In a similar way, the completion time difference for any individual student (Δt_i) is also expected to be within the range shown in the table; namely, -4 to +40 minutes. Taken together, these ranges constitute a region of acceptable values of examination score improvements and completion times. Any value of ΔE_i or Δt_i falling outside these ranges may be considered as suspect. Figure 1 shows a plot of individual score and time completion differences for 16 students for which both online and proctored exam and time data was available. The shaded region in the figure illustrates the region of acceptable values.

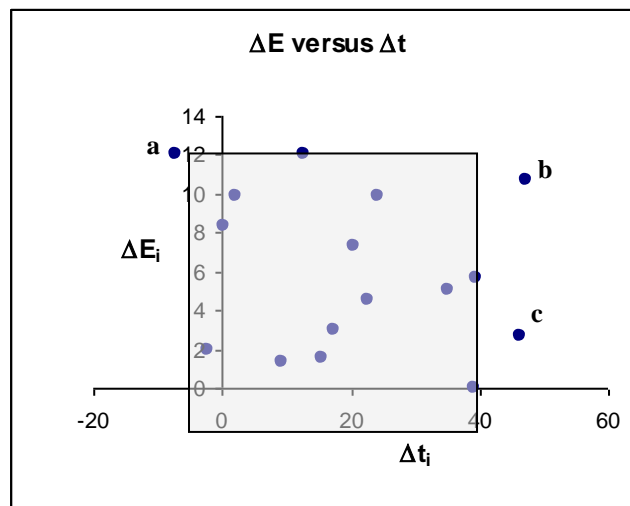


Figure 1. Plot of examination score differences versus completion time differences.

As may be observed, three points (**a**, **b**, and **c**) lay outside the acceptable region of values. In particular, points **a** and **b** show a large exam differences and may indicate that cheating has occurred. For the point **a**, the negative time difference may be an indication that the student was assisted in taking the exam, thus finishing more quickly and with a better score. For point **b**, the large score difference may be the result of a student using additional time to look through a textbook or other materials to identify the answers. The same may be said for point **c**, although in this case the resulting improvement in the examination score was not as great. It must be emphasized that the analysis performed here does not unequivocally result in a determination of whether cheating has occurred. It merely suggests a method for interpreting large examination score and completion time differences. In practice, a completely secure online testing environment is mostly likely no more achievable than the typical proctored environment and would required great costs to implement (Williams, 2000).

CONCLUSIONS

By evaluating the established criteria, the online testing used to evaluate student performance in the author's Introduction to Astronomy course was found to be equivalent to the proctored examinations. The results of the quantitative and qualitative data suggests that the online testing environment is more suitable to the learning and testing styles of the students and thus resulted in equivalent or, in some cases, improved examination scores. In particular, the longer time spent by students on online tests was offered as evidence of this greater suitability. The occurrence of cheating on online examinations could not be ruled out and an analysis of the online and proctored scores and completion times suggests that cheating possibly did occur in the course.

FUTURE PLANS

To refine the conclusion that longer completion times resulted in better test scores, the author plans to repeat the study for shorter maximum allowed testing times. A further analysis of potential cheating will also be performed for these shorter testing times to determine if potential incidences of cheating were affected. In addition, the type of examination to be given in the online environment will be studied to determine the optimum method of assessing student performance while maintaining a secure testing environment.

ACKNOWLEDGMENT

The author would like to extend his sincerest appreciation to Dr. Jeanne Buckley, past Vice President of Information Technology at Delaware County Community College for her constant support and encouragement in introducing technology into instruction.

REFERENCES

Admin, N. (2002). Online Assessment-Friend or Foe!, *Achieve Online Australia*. This article is available online at: <http://www.achieveonline.com.au/article/articleview/272/1/8/>

Hench, T. (2003). A Model for Combining Computer-based Distance Learning with In-class Instruction, *Proceedings of the 6th International Conference on Computer-Based Learning in Science*, Nicosia, Cyprus, July 5 – 10, 2003.

White, J. (2000). Online testing: The dog sat on my keyboard, *Proceedings of the 13th Annual International Conference on Technology in College Mathematics*. This paper is available online at: <http://www.archives.math.utk.edu/ICTCM/EP-13/html/paper.html>

Williams, J. (2000). Flexible Assessment for Flexible Delivery: Online Examinations that Beat the Cheats, *Proceedings of Moving Online I: A conference to explore the challenges for workplace, colleges, and universities*. This paper is available online at: <http://science.uniserve.edu.au/newsletter/vol18/williams.html>

Thomas Lee Hench
Delaware County Community College
1901 South Media Line Road
Media, Pennsylvania, 19063
USA
thench@dccc.edu