International Journal of Teaching and Learning in Higher Education 2015, Volume 27, Number 2, 221-226
http://www.isetl.org/ijtlhe/ ISSN 1812-9129

Proctored and Unproctored Test Performance

Sara Brallier and Linda Palm
Coastal Carolina University

This study examined test performance as a function of test format (proctored versus unproctored) and course type (traditional versus distance). The participants were 246 undergraduate students who completed introductory sociology courses during four semesters at a southeastern university. During each semester, the same instructor taught a traditional lecture section and a distance section of the course. Students in both course types took unproctored online tests in two semesters while students in both course types took proctored classroom paper-and-pencil tests in the other two semesters. Students scored significantly higher on the unproctored online tests than on the proctored classroom tests. There was no significant difference in test performance between students enrolled in distance courses and those enrolled in lecture courses. Additionally, no significant interaction was found between test format and course type. Implications of these results for the design and structure of online and hybrid courses are discussed.

Assessment is an integral component of teaching and learning (Rovai, 2000; Rowe, 2004; Serwatka, 2003). Tests are a common form of assessment used in distance and on-campus courses to assess student learning. Proctored, closed-book, pencil-and-paper tests are the predominant method of testing in academia, despite the absence of empirical evidence regarding the advantages of this testing format (Williams, 2004; Williams & Wong, 2009). Amid skepticism and concern, an increasing number of instructors are using online testing in both on-campus and distance courses (Khare & Lam, 2008; Rowe, 2004).

Online testing offers a number of advantages to instructors and students. Online testing is useful in distance courses in which the students are geographically dispersed, and it is not feasible for students to come to campus to take tests (Kinney, 2001). When used to assess student learning in lecture classes, online testing reduces instructional time dedicated to testing (Tao & Li, 2012). For both traditional and distance classes, the flexibility of online testing allows instructors to schedule frequent course tests and quizzes (Bonham, 2006; DeSouza & Fleming, 2003; Graham, Mogel, Brallier, & Palm, 2008). Frequent testing encourages students to keep current on their course reading and studying and increases content mastery (Hadsell, 2009; Smith, 2007). Using computerized scoring, online testing can provide objective, immediate feedback to instructors and students (Bonham, 2006; Hamilton & Shoen, 2005, Tao & Li, 2012). Immediate feedback is useful for student learning because students can remember how and why the mistakes were made while the material is still recent in their minds (Khare & Lam, 2008). This combination of ease of frequent test/quiz administration and immediate feedback allows students the opportunity to monitor their comprehension of course materials (Harmon & Lambrinos, 2008). There is also an advantage for instructors who want to examine aggregate test data or conduct test item analyses since test data are directly entered into electronic databases which allows for easy analysis (Bonham, 2006; Hamilton & Shoen, 2005). Additionally, several studies have reported that students find online testing less stressful and prefer online tests to written ones (Bonham, 2006; Khare & Lam, 2008).

Tao and Li (2012) note that instructors have two options for online testing: (a) students can take an online test on a computer in a proctored setting, or (b) students can take an unsupervised online take-home test in a setting and on a computer of their own choosing. Since the unsupervised online test environment is generally not regulated by course instructors, most instructors allow these tests to be open-book and open-notes (Shultz, Shultz, & Gallogly, 2007; Tao & Li, 2012), and one must assume that students will be using all of the resources at their disposal (DeSouza & Fleming, 2003; Kinney, 2001; Osika, 2006). A primary worry expressed over unsupervised online tests is that test results will be inflated if students are allowed to consult course materials (Rovai, 2000). Several researchers report similar test performance for students taking proctored online tests and students taking proctored tests in the classroom (Anakwe, 2008; Bonham, 2006; MacCann, 2006). There have been mixed findings regarding student performance on unproctored online tests and proctored in-class tests. Frein (2011) examined test scores on multiple-choice tests of military cadets enrolled in an introductory psychology course. A comparison of performance on proctored paper-and-pencil in-class tests, proctored online tests, and unproctored online tests revealed no significant difference as a function of test format. Frein speculated that this result may

have been influenced by the strict honor code enforced at his institution. Schultz, Schultz, and Gallogly (2007) compared performance of students on proctored paper-and pencil exams and unproctored online exams in marketing, management and accounting classes and reported significantly higher performance on the unproctored exams. However, the size of the effect of test format was relatively small. Carstairs and Myors (2009) assessed performance of two cohorts of students in an upper-level psychology course as a function of test format during two semesters. The first cohort completed three proctored paper-and-pencil in-class multiple-choice tests during the first semester while the second cohort completed a take-home paper-and-pencil multiple-choice test, an unproctored online multiple-choice test and a proctored paper-and-pencil in-class multiple-choice test. Students in the two cohorts performed similarly under proctored testing conditions; however, students in the second cohort scored significantly higher on the unproctored take-home and online tests. Evaluation of effect size revealed that test format had a large effect on test performance.

We believe that further assessment of performance on unproctored online tests and proctored in-class tests is important because this difference in the method of testing reflects one of the most fundamental differences between online and traditional face-to-face lecture courses. The purpose of this study was to examine test performance as a function of course type (traditional lecture course versus distance course) and test format (unproctored online exams and proctored in-class exams). In each of four semesters, students enrolled in either a distance or a traditional lecture section of introductory sociology. In two of the semesters, all students in both course types took proctored closed-book exams in a classroom. In the other two semesters, students in both course types took unproctored online course tests. Based on meta-analytic studies reporting no significant difference in academic performance in lecture-based and distance courses (Bernard et al., 2004; Zhao, Lei, Lai, & Tan, 2005), we predicted that test scores would not differ as a function of course type. Recognizing that students would be able to access course materials while completing the unproctored online tests, we predicted that performance on these tests would be higher than performance on the proctored in-class tests. Subsequent to finding an effect of test format on test performance, an additional purpose of the study was to assess the magnitude of the effect and to examine whether course grades varied as a function of test format.

Method

Participants

The participants in this study were 246 undergraduate students who completed Introductory Sociology (SOC 101) during four consecutive fall semesters. The sample included 98 men and 148 women. The mean age of the students was 21.14 years (SD = 6.66). The sample contained 108 freshmen, 69 sophomores, 35 juniors, and 34 seniors. The racial distribution was 196 Caucasian students and 50 African-American/Hispanic/Other students. Eighty-seven students completed a distance course, and 159 completed a lecture course. One hundred thirty students completed online exams, and 116 students completed classroom exams.

Materials

Students in the distance and lecture courses were assigned the same introductory sociology textbook (Henslin, 2005). Students accessed online materials through WebCT/Blackboard. Students in the lecture courses were given instructor-prepared notes and handouts in print form, and students in the distance courses obtained these materials via WebCT/Blackboard.

The director of the university's Office of Institutional Research, Assessment and Analysis supplied the researchers with an Excel file containing the following demographic and academic information for students registered in the introductory sociology courses: age, class rank, gender, race, high school GPA, verbal SAT score, quantitative SAT score, and cumulative college GPA. The provision of the data was done in accordance with the university's privacy policies.

Procedure

Data for this study were collected in introductory sociology courses taught at a midsized state-supported southeastern university during four semesters. Introductory sociology is a required course for sociology majors; for students in other majors, this course may be used to fulfill a core curriculum requirement or serve as an elective course. In each of the four fall semesters, the same instructor offered two sections of the course, one as an online distance course and one as a traditional classroom lecture course. Assignment of students to the distance or lecture courses was not random; students selected one of the formats when they registered for the course.

The distance and lecture courses were designed to be as consistent as possible. Both sections of the course were taught by the same instructor, and the same textbook was assigned. Students in both types of course were provided with the same materials, completed the same written assignments and discussion activities, and took the same tests. The grading system was standardized across the two course types with test scores contributing 32%, written assignments contributing 43%, and discussion activities contributing 25% to the final course grade.

At the beginning of the semester, students in both course types were asked to complete a 50-question online pretest assessing their baseline knowledge of sociological concepts. In all four semesters, the instructor met with the lecture-based class for fifty minutes three times a week. The only time the instructor met with the distance class was for an introductory on-campus meeting during the first week of the semester to introduce herself and provide a WebCT/Blackboard tutorial. During each semester, three non-cumulative 50-item multiple-choice unit tests and one cumulative 60-item multiple choice final test were administered. Students could earn a total of 210 points on these four tests. In two semesters, the instructor administered all tests via WebCT/Blackboard in both the distance and the lecture courses. Students took the unproctored, open book/open notes online tests from the remote location of their choice (e.g., home, computer lab at the university). Students were allotted 50 minutes to take the online tests. Students had access to the tests for a 24-hour period. All students were given the same questions, but the questions were presented in a random order. In the other two semesters, students in both the distance and lecture courses took proctored, closed book/closed notes paper-and-pencil tests in a classroom. In both the distance and lecture course types, the online and classroom tests contained the same content, the tests were given at the same time during the semester, and students had the same time limit of 50 minutes for completing each test. The number of students completing each test format and each course type is shown in Table 1.

Results

A test score was calculated for each student by computing the percent of total points earned on the four course tests. A 2 x 2 between-subjects analysis of variance was used to examine test scores as a function of test format (unproctored online tests versus proctored in-class tests) and course type (distance course versus lecture course). Test format had a significant effect on test scores, $F(1, 242) = 17.41$, $p < .001$, $\eta^2_p = .07$. Students who took unproctored online tests scored significantly higher ($M = 74.66$, $SD = 10.87$) than students who took proctored in-class tests ($M = 68.65$, $SD = 12.12$). No significant main effect was found for course type, $F(1, 242) = 3.45$, $p = .07$, $\eta^2_p = .01$. The mean test score for students who completed distance courses was 70.08 ($SD = 12.75$), and the mean for students who completed lecture courses was 72.78 ($SD = 11.24$). Additionally, no significant interaction was found between test format and course type, $F(1, 242) = 3.27$, $p = .07$, $\eta^2_p = .01$. The mean test scores as a function of test format and course type are shown in Table 2.

Test format covaried with semester of enrollment; therefore, it was possible that the higher test scores of online test-takers compared to classroom test-takers could reflect differences in academic characteristics of students enrolled in semesters when online tests were administered and students enrolled in semesters when in-class tests were administered. To evaluate this possibility, independent *t* tests were used to compare the two groups of students on five academic measures: high school GPA, verbal SAT scores, quantitative SAT scores, cumulative college GPA, and percentage scores on the sociology pretest. As shown in Table 3, no significant difference was found between the mean scores of the two groups of students on any of these academic measures.

The average test scores of students who completed unproctored online tests were 6% higher than those of students who completed proctored in-class tests. This difference represented a medium effect size with test format accounting for 7% of the variance in test scores and raised a concern that course grades for students who took online tests might be inflated relative to the grades of students who took in-class tests. The course grade distribution as a function of test format is shown in Table 4. A chi square test for independence revealed no significant relationship between test format and course grades, $\chi^2 = 3.47$, $p = .48$.

Discussion

Our predictions concerning course type and test format were supported. Test performance of students enrolled in distance versus lecture courses was comparable. This is consistent with numerous studies which have found that student learning outcomes in well-designed distance courses are similar to those in traditional lecture courses (Bernard et al., 2004; DiRienzo & Lilly, 2014; Rivera & Rice, 2002). The online unproctored test format did result in significantly higher test scores than the proctored classroom test format. However, congruent with the

Table 1
Number of Students Completing Each Test Format and Course Type by Semester

Semester	Test Format	Distance Course *n*	Lecture Course *n*	Semester *N*
Semester 1	Online	24	38	62
Semester 2	In-class	19	36	55
Semester 3	In-class	21	40	61
Semester 4	Online	23	45	68

Table 2
Mean Test Scores as a Function of Test Format and Course Type

	Test Format					
	Online Tests			In-class Tests		
Course Type	*M*	*SD*	*n*	*M*	*SD*	*n*
Distance Course	74.50	10.20	47	64.88	13.58	40
Lecture Course	74.75	11.28	83	70.63	10.86	76

Table 3
Comparison of Academic Characteristics of Students Taking Online and In-class Tests

	Test Format							
	Online Tests			In-class Tests				
Characteristic	*n*	*M*	*SD*	*n*	*M*	*SD*	*t*	*p*
High School GPA	116	3.31	0.52	96	3.36	0.49	-0.72	.48
SAT Verbal	91	517.58	77.46	78	505.90	66.69	1.04	.30
SAT Quantitative	91	519.45	64.00	78	520.00	65.70	-0.06	.96
College GPA	75	3.04	0.61	63	3.04	0.58	0.00	1.00
Sociology Pretest	124	52.72	12.28	112	49.91	12.70	1.73	.09

Table 4
Course Grade Distribution as a Function of Test Format

	Course Grade									
	A		B		C		D		F	
Test Format	n	%	n	%	n	%	n	%	n	%
Online Tests	47	36.2	54	41.5	19	14.6	6	4.6	4	3.1
In-class Tests	43	37.1	42	36.2	14	12.1	10	8.6	7	6.0

findings of Schultz et al. (2007) this difference was surprisingly modest. Students taking the online tests scored, on average, 6% higher. Moreover, students who took the online tests did not earn higher course grades in introductory sociology. This suggests that online tests are a viable option in online and hybrid courses as students who had access to books and other resources did not have test scores dramatically higher compared to students who completed the tests in class without access to resources.

Many instructors view testing as a learning activity and are willing to allow students to access course materials while taking online tests. Agarwal et al. (2008) suggest that students find open-book tests less stressful, and open-book tests may encourage students to practice higher level thinking skills like problem-solving and reasoning. When

allowing students to use course materials, online test scores should account for a relatively small percentage of the students' course grade (Harmon & Lambrinos, 2008; Osika, 2006). A more substantial percentage could come from more comprehensive assessment tools, such as essays, projects and portfolios (Rovai, 2000).

Several strategies can be used to limit students' reliance on course materials or sharing answers when taking online tests. First, tests should include questions that require students to process or apply information rather than exhibiting memorization of simple facts. Second, limiting the test completion time decreases students' abilities to use resources and requires that they have done advanced preparation (Harmon & Lambrinos, 2008; Kinney, 2001; Rovai, 2000). Additionally, the test should be administered to all students at the same time to prevent students who take the test first from sharing questions with others (Rowe, 2004). Finally, using a database of questions from which tests are randomly constructed limits students' ability to share answers (Harmon & Lambrinos, 2008; Rovai, 2000; Rowe, 2004; Tao & Li, 2012).

References

Agarwal, P. K., Karpicke, J. D., Kang, S. H. K., Roediger, H. L., & McDermott, K. B. (2008). Examining the testing effect with open- and closed-book tests. *Applied Cognitive Psychology*, *22*(7), 861-876. doi:10.1002/acp.1391

Anakwe, B. (2008). Comparison of student performance in paper-based versus computer-based testing. *Journal of Education for Business*, *84*(1), 13-17.

Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., … Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research*, *74*(3), 379-439. doi: 10.3102/00346543074003379

Bonham, S. (2006). Reliability, compliance and security of web-based pre/post testing. *Proceedings of 2006 Physics Education Research Conference* (pp. 133-136). Syracuse, NY: American Institute of Physics. doi: 10.1063/1.2508710

Carstairs, J., & Myors, B. (2009). Internet testing: A natural experiment reveals test score inflation on a high-stakes, unproctored cognitive test. *Computers in Human Behavior*, *2*(3), 738-742. doi: 10.1016/j.chb.2009.01.011

DeSouza, E., & Fleming, M. (2003). A comparison of in-class and online quizzes on student exam performance. *Journal of Computing in Higher Education*, *14*(2), 121-134.

DiRienzo, C., & Lilly, G. (2014). Online versus face-to-face: Does delivery method matter for undergraduate business school learning? *Business Education & Accreditation*, *6*(1), 1-11.

Frein, S. T. (2011). Comparing in-class and out-of class computer-based tests to traditional paper-and-pencil tests in introductory psychology courses. *Teaching of Psychology, 38*(4), 282-287. doi:10.1177/0098628311421331

Graham, J. M., Mogel, L. A., Brallier, S. A., & Palm, L. J. (2008). Do you online? The advantages and disadvantages of online education. *Bridges*, *2*, 27-36.

Hadsell, L. (2009). The effect of quiz timing on exam performance. *Journal of Education for Business*, *84*(3), 135-141.

Hamilton, D., & Shoen, E. (2005). Same song, second verse: Evaluation and improvement of an established assessment program. In K. Martell and T. Calderon, *Assessment of student learning in business schools: Best practices each step of the way.* (Vol. 1, No. 2, pp. 119-137). Tallahassee, Florida: Association for Higher Research.

Harmon, O. R., & Lambrinos, J. (2008). Are online exams an invitation to cheat? *Journal of Economic Education, 39*, 116-125.

Henslin, J. (2005). Sociology: A down-to-earth approach (7th ed.). Boston, MA: Allyn and Bacon.

Khare, A., & Lam, H. (2008). Assessing student achievement and progress with online examinations: Some pedagogical and technical issues. *International Journal on E-Learning*, *7*, 383-402.

Kinney, N. (2001). A guide to design and testing in online psychology courses. *Psychology Learning and Teaching*, *1*(1), 16-20. doi: 10.2304/plat.2001.1.1.16

MacCann, R. (2006). The equivalence of online and traditional testing for different subpopulations and item types. *British Journal of Educational Technology*, *37*(1), 79-91. doi:10.1111/j.1467-8535.2005.00524.x

Osika, E. R. (2006). Assessing student learning online: It's more than multiple choice. *Online Classroom*, June, 3-7.

Rivera, J. C., & Rice, M. L. (2002). A comparison of student outcomes and satisfaction between traditional & web based course offerings. *Online Journal of Distance Learning Administration*, *5*(3), 1-12.

Rovai, A. P. (2000). Online and traditional assessments: What is the difference? *The Internet and Higher Education*, *3*, 141-151.

Rowe, N. C. (2004). Cheating in online student assessment: Beyond plagiarism. *Online Journal of Distance Learning Administration*, *7*(2), 1-10.

Schultz, M. C., Schultz, J. T., & Gallogly, J. (2007). The management of testing in distance learning environments. *Journal of College Teaching & Learning, 4*(9), 19-26.

Serwatka, J. A. (2003). Assessment in on-line CIS Courses. *Journal of Computer Information Systems*, *44*(1), 16-20.

Smith, G. (2007). How does student performance on formative assessments relate to learning assessed by exam? *Journal of College Science Teaching*, *36*(7), 28-34.

Tao, J., & Li, X. (2012). A case study on computerized take-home testing: Benefits and pitfalls. *International Journal of Technology in Teaching and Learning, 8*(1), 33-43.

Williams, J. B. (2004). Creating authentic assessments: A method for the authoring of open book open web examinations. In R. Atkinson, C. McBeth, D. Jonas-Dwyer, & R. Phillips (Eds.), *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference* (pp. 934-937). Perth, 5-8 December. http://www.ascilite.org.au/conferences/perth04/procs/williams.html

Williams, J. B., & Wong, A. (2009). The efficacy of final examinations: A comparative study of closed-book, invigilated exams and open-book, open-web exams. *British Journal of Educational Technology, 40*(2), 227-236.

Zhao, Y., Lei, J., Yan, B., Lai, C., & Tan, H. S. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. *Teachers College Record*, *107*(8), 1836–1884.

SARA BRALLIER is a professor of sociology at Coastal Carolina University. She earned a PhD in sociology from the University of Buffalo. She teaches courses in research methods, sociology of religion, and social gerontology. Her research interests include identifying teaching and assessment methodologies that promote student learning in traditional lecture-based, hybrid, and online classes.

LINDA PALM is a professor of psychology at Coastal Carolina University who earned a PhD in experimental psychology from the University of South Florida. She teaches courses in statistics, research methods, child and adolescent psychology, history and systems of psychology, and principles of learning. Dr. Palm has worked in program evaluation in the fields of academic under-achievement and child and adolescent mental health. Her research interests include the impact of instructional technologies on academic performance and psychosocial predictors of student success.