Student Performance in Online Accounting Tests: In-class vs. Take-home

Mohamed Gomaa
California State Polytechnic University - Pomona
Fei Kang
California State Polytechnic University - Pomona
Hong Pak*
California State Polytechnic University - Pomona

ABSTRACT

Online and hybrid course offerings have increased over the years, with a large percentage of students taking advantage of the flexibility these offerings provide. Online testing is a major component of these types of courses, but it is also increasingly being used in traditional face-to-face courses. The purpose of this study is to compare student performance in unproctored open book online exams with that in proctored open book online exams for an undergraduate accounting course. To address our research question, students in an intermediate accounting course took two midterm and one final exams by either traditional paper-based format in class, unproctored online format at home, or proctored online format in class. Our results show that the average student performance in unproctored online exams is higher than that of online exams proctored in class.

Keywords: Online Accounting Test, Performance, Accounting Education

JEL: M41, M49

*corresponding author contact: hspak@cpp.edu

1. INTRODUCTION

The number of universities offering online and hybrid course options has been steadily increasing over the years. A recent survey conducted by the Babson Survey Research Group (2017) has found that in 2016 distance education enrollment has increased for the fourteenth straight year. Their results show that over 30 percent of higher education students are taking at least one distance education course. Students provide many reasons for taking online courses. The vast majority cite flexible scheduling, and convenience, as their main reasons (Daymont et al. 2011; Cochran et al. 2016).

One component of online and hybrid courses is online testing. Additionally, many faculty members have been adding online testing to their face-to-face courses (Rowe 2004; Khare and Lam 2008; Stowell and Bennett 2010; Brallier et al. 2015). There are several advantages to online exam administration over traditional paper and pencil exams, including the efficiency of administration, lower cost, instant feedback to students, scheduling flexibility, and time savings through automatic recording of grades (DeSouza and Fleming 2003; Bonham 2006; Graham et al. 2008; Harmon and Lambrinos 2008; Khare and Lam 2008; Stowell and Bennett 2010).

Findings regarding the equivalence and validity of online exams as compared to in-class face-to-face exams have been mixed. The purpose of this study is to compare student performance in unproctored open book online exams with performance in proctored open book online exams for an undergraduate accounting course. To answer this research question, we conducted analysis on students' exams in an intermediate accounting course at a large public university. Participants in our study took three exams, which were either traditional paper-format or online conducted. The online exams incorporated the features available in the software to reduce the possibility of cheating. Our results suggest that the average student performance in unproctored online exams is significantly higher than that of online exams proctored in class. We also performed additional analysis comparing student grades by type of questions, by AACSB category, and by question difficulty, providing further evidence that students taking unproctored online exams at home significantly outperformed those taking proctored online exams in class.

The remainder of this paper is organized as follows. Section 2 discusses the literature related to online exams. Section 3 presents our method, followed by Section 4 on our results. Our final section provides concluding remarks.

2. LITERATURE REVIEW

Research comparing exams administered online to traditional in-class exams has produced mixed results. Alexander et al. (2001) conducted an experiment to examine the student performance equivalence for online compared to traditional paper and pencil testing methods. Their results showed that the test scores were not significantly different between the two groups. Hollister and Berenson (2009) found similar results. Their study differed from Alexander et al. (2001) in that their two groups were both online, with one being proctored while the other was unproctored. They found no significant difference between the two groups. Frein (2011) found similar results when comparing three groups; proctored in-class paper-based exams, proctored in-class computer-based exams, and unproctored online exams.

However, other researchers found differences between unproctored online exams and proctored in-class exams. Some researchers found that unproctored online or computer-based exams negatively affected test scores. Lee and Weerakon (2001) examined the extent and effect of 'computer anxiety' by comparing student performance in computer-based and paper-based multiple-choice tests. They found that students performed significantly better in the paper test than

in the computer test. In an earlier paper, Lee et al. (1986) examined if mean scores on computer-based version of an arithmetic reasoning test would be significantly lower than those on the paper-based version. They had 585 participants in their study, and there was no time limit imposed on the participants. They found that the mean score obtained by participants in the computer-based group was significantly lower than that obtained by participants in the paper-based group. Wallace and Clariana (2005) examined the effects of test mode, gender, and race on paper-based versus computer-based multiple-choice exams in upper-level business courses. They had 144 students participating in their study. They found that participants taking the paper-based exams significantly outscored those taking the computer-based exams.

On the other hand, prior studies have documented that online unproctored exams increased test scores compared to proctored in-class exams (DeAngelis 2000; Clariana and Wallace 2002). DeAngelis (2000) examined the equivalence of computer-based and paper-based versions of an exam in an upper-level dental hygiene class. They found that students taking the computer-based exam outperformed students taking the paper-based exam. The difference was statistically significant. Clariana and Wallace (2002) randomly assigned 105 freshman business undergraduates to either computer-based or identical paper-based exams to examine possible test mode effects. Their findings showed that the computer-based test group outperformed the paper-based test group. In a lecture-based introductory sociology course, Brallier and Palm (2015) compared performance on online unproctored exams to proctored classroom exams. On the unproctored online exams average student performance was four percent higher compared to the proctored classroom exams. Brallier and Palm (2015) also compared exam performance as a function of exam mode in an online version of the course. Student performance was ten percent higher on unproctored online exams compared to proctored classroom exams. These differences were statistically significant. Similarly, Carstairs and Myors (2009) studied performance difference in a lecture-based upper-level psychology course. In their study, average student scores were five points higher on an unproctored online exam than on a traditional proctored in-class exam. Schultz et al. (2007) had similar findings in online marketing, management, and accounting classes. In their study, they found that students performed around four points higher on the unproctored exams as compared to proctored in-class exams.

In summary, prior studies have documented mixed findings on student performance with respect to a variety of testing forms. Thus, this leads us to the following research question:

RQ1: Is student performance in unproctored (taken at home) open book online exams significantly different from that in proctored (taken in class) open book online exams for accounting courses?

To address the research question, this study was conducted in the first intermediate accounting classes at a large public university in the United States. The study compares student performance on unproctored open book online exams (computer-based) taken at home with their performance on proctored open book online exams (computer-based) taken in class for an undergraduate accounting course.

3. METHOD

3.1. Participants

Participants of this study are the students who took the first intermediate accounting course of three intermediate accounting courses series (ACC 311, ACC 312, and ACC 313) in Summer

and Fall 2017 taught by the same instructor (who is one of the authors). Two sections of the course were offered for each term. The total number of students enrolled for the course was 110. Of these students, we excluded three students who took the tests at Disability Resource Center (DRC) and nine students who took the course repetitively in Fall 2017 as they failed in the course in Summer 2017. We analyzed the data collected from 98 students for this study.

The participants took two midterms and one final exam. They took the first midterm (traditional paper and pencil) test and the online (McGraw-Hill Connect) final exam in class. For the second midterm test, the instructor had one section take the open-book online test at home (unproctored), and the other section take the open-book online test in class (proctored). Both groups took the test during the class time period. The assignment was made one week before the second midterm exam was taken, to minimize confounding factors (e.g., cheating schemes) that might contribute to the students' test results. Twenty-eight students in Summer 2017 and 19 students in Fall 2017 took the proctored second midterm exam in class, while 29 students in Summer 2017 and 22 students in Fall 2017 took the unproctored same exam at home. Thus, 47 students in total (Group 1) took proctored online 2nd midterm test, whereas 51 students in total (Group 2) did the unproctored same test at home.

3.2. Materials

Students took three exams for the course, including two midterms and one non-cumulative final exam. The first midterm exam covers Chapter 1 of Spiceland et al.'s Intermediate Accounting (8th ed. Summer 2017 and 9th ed. Fall 2017) and Financial Accounting Standards Board's (FASB) Conceptual Framework Statements 8, 6, and 5. The test consists of 20 multiple choices, one short answer, one short essay, and three computational questions. As the questions are asked to assess the students' understanding of basic financial accounting concepts and terminologies, they are dominantly the conceptual and applicational types of questions[1]. We use this exam as our benchmark for our study. The second midterm exam covers Chapters 2 to 4 of Spiceland's Intermediate Accounting. The exam consists of 25 multiple choice and 4 problem questions. These questions are selected from McGraw-Hill's online test bank for the textbook. They consist of the computational, journal entry, and financial statement presentation questions, whose natures are either analytical or applicational. The exam was designed in a way to minimize student cheating (e.g., communications among students or other individuals, and/or help from online resources) by adopting the algorithm feature and by selecting the higher level of difficulty questions (14 medium and 15 difficult questions by the test bank classification for both Summer and Fall 2017). The second exam is used as the testing instrument for this study. The final exam covers Chapters 3 to 6 of Spiceland's Intermediate Accounting and FASB's Conceptual Framework Statement No. 7. The test consists of 50 multiple choice questions whose types are computational, conceptual, or journal entry. The final exam is used as the second benchmark.

While the content of the first midterm exam in Summer 2017 is the same as that of the first midterm exam Fall 2017, the contents of the second midterm and final exams in Summer 2017 are slightly different from those of the exams in Fall 2017 as the new 9th edition of Spiceland's

[1] The first midterm should be graded without bias since the majority (95%) of the test consists of objective questions. Grading bias for one essay question should be minimal because the remaining 5% of the test (i.e. the portion of one essay question) asks the explanatory question as reference to objective conceptual attributes.

Intermediate Accounting was adopted Fall 2017. However, we believe that the difference does not materially affect the test results.[2]

4. RESULTS

4.1. Descriptive Statistics

Table 1 reports the summary statistics of the student grades from Midterm I (Panel A), Midterm II (Panel B), and Final (Panel C), respectively. For Midterm I, both student groups took the same paper-based exam in class, with average grades 62.6% (Group 1) and 55.9% (Group 2), respectively; for midterm II, students in Group 1 took online exam in class with an average 67.4%, while students in Group 2 took online exam at home with an average 72.8%; for the final, both student groups took the same online exam in class, with average grades 77.0% (Group 1) and 74.3% (Group 2), respectively. As shown in the table, the mean values of the student grades are not statistically different across the two groups, except for Midterm I (for which the difference is marginally significant with a two-tailed *P*-value 0.066).

Table 1: Descriptive Statistics of Student Grades

Panel A: Student grades from Midterm I

	Group 1 (Paper in-class)	**Group 2 (Paper in-class)**
Mean	0.626	0.559
Variance	0.030	0.035
Observations	47	51
t Stat	1.860	
P(T<=t) two-tail	0.066	

Panel B: Student grades from Midterm II

	Group 1 (Online in-class)	**Group 2 (Online home)**
Mean	0.674	0.728
Variance	0.037	0.029
Observations	47	51
t Stat	-1.454	
P(T<=t) two-tail	0.149	

[2] We performed t-test to compare the test results between the students taking the exams in Fall vs. Summer 2017 and did not find significant differences across the two groups.

Panel C: Student grades from Final

	Group 1 (Online in-class)	**Group 2 (Online in-class)**
Mean	0.770	0.743
Variance	0.020	0.030
Observations	47	51
t Stat	0.838	
P(T<=t) two-tail	0.404	

Table 2 provides the Pearson correlation matrix between the three exam grades. All of them are significantly correlated with each other at the 0.10 level.

Table 2: Correlation (p-value) Matrix among MI, MII, and Final

	Midterm1	**Midterm2**	**Final**
Midterm I	1.000		
Midterm II	0.202 (0.023)	1.000	
Final	0.321 (0.001)	0.144 (0.085)	1.000

4.2. Comparing test results between online in-class vs. online at home

Panel A of Table 3 presents the ANOVA results of comparing student grades from the online in-class format vs. online at-home format, i.e., Midterm II results of Group 1 vs. Group 2. To mitigate the across-group differences, we used the student grades from Midterm I as the benchmark, and performed the difference-in-difference analysis, i.e., comparing the improvement of the grades from Midterm I to Midterm II (M2-M1) across groups. As reported in Panel A of Table 3, the ANOVA results suggest that the improvement of the average grades of the online exam taken at home (Group 2) is 12.2% higher than that of the online exam taken in class (Group 1). The difference is statistically significant at the 0.01 level (*p*-value = 0.009). The results are similar to the ANOVA tests for each quarter. Our results suggest that the average student performance in unproctored online exams is significantly different from that of online exams proctored in class.

Table 3: ANOVA Results for Online In-class vs. Online Home

Panel A: (Grades improvement M2-M1)

Quarter	Group	Count	Average	Variance	p-value
Both	in-class	47	0.047	0.050	0.009***
	home	51	0.169	0.050	
Fall	in-class	19	-0.025	0.035	0.084*
	home	22	0.080	0.036	
Summer	in-class	28	0.097	0.055	0.027**
	home	29	0.236	0.052	

Panel B: (Grades improvement M2-Final)

Quarter	Group	Count	Average	Variance	p-value
Both	in-class	47	-0.096	0.051	0.076*
	home	51	-0.016	0.047	
Fall	in-class	19	-0.180	0.043	0.575
	home	22	-0.148	0.026	
Summer	in-class	28	-0.039	0.050	0.033**
	home	29	0.085	0.040	

* Significant at the 0.10 level.
** Significant at the 0.05 level.
***Significant at the 0.01 level.

Panel B of Table 3 presents the ANOVA results of comparing student grades for the two groups for the final exam for the online in-class format vs. online at-home format. Similar to the previous test, to mitigate the across-group differences, we used the student grades from the final exam as the benchmark, and performed the difference-in-difference analysis, i.e., comparing the grades improvement from Midterm II to the Final Exam (MII-Final) across groups. As reported in Panel B of Table 3, the ANOVA results suggest that the average grades of online exams taken at home (Group 2) is 8% higher than that of online exam taken in class (Group 1). The difference is statistically significant, though less significant compared to the results using Midterm I as the benchmark, as reported in Panel A. Our results suggest that the statistical difference between the groups might be related with whether the implementation of the unproctored take-home online test is initial or not. This issue suggests an interesting area which could be addressed in the future study.

4.3. Additional Analysis

To provide further evidence on the effect of test format on the student test performance, we performed additional analysis comparing student grades by type of questions, by AACSB category, and by question difficulty. Table 4 summaries the ANOVA results for each category. Panel A shows that students taking online exam at home significantly outperformed those taking online exam in class on both multiple choices and numerical problems (the difference is 12.9% on multiple choices with a *P*-value 0.003, and 33.6% on numerical problems with a *P*-value 0.003); Panel B illustrates that students taking online exam at home received significantly better grades than those taking online exam in class on both analytical and application questions (the difference is 10.3% on analytical questions with a *P*-value 0.038, and 16.5% on application questions with a *P*-value 0.001); lastly, the results from Panel C indicate a similar pattern showing that students taking online exam at home perform significantly better than those taking online exam in class on both medium and hard questions (the difference is 14.1% on medium questions with a *P*-value 0.003, and 11.3% on hard questions with a *P*-value 0.026).

In summary, the additional analysis provides further evidence that students taking online exam at home significantly outperformed those taking online exam in class across all the different types of questions, which is consistent with our results in the main analysis as shown in the previous table. Our findings shed additional insights on the effect of test format on student test performance.

Table 4: Additional Analysis

Panel A: By Type of Question

Type of question	Group	Count	Average	Variance	p-value
Multiple choice	in-class home	47 51	0.183 0.312	0.043 0.047	0.003***
Problems	in-class home	47 51	-0.088 0.248	0.089 0.084	0.060*

Panel B: By AACSB Category

Category by AACSB	Group	Count	Average	Variance	p-value
Analytical	in-class home	47 51	0.059 0.162	0.062 0.054	0.038**
Application	in-class home	47 51	0.033 0.198	0.056 0.067	0.001***

Panel C: By Question Difficulty

Category by Difficulty	Group	Count	Average	Variance	p-value
Medium	in-class home	47 51	0.205 0.346	0.052 0.050	0.003***
Hard	in-class home	47 51	-0.014 0.099	0.062 0.061	0.026**

* Significant at the 0.10 level.
** Significant at the 0.05 level.
***Significant at the 0.01 level.

5. CONCLUSION

Online and hybrid course offerings have increased over the years. A large percentage of students are taking advantage of the flexibility these offerings provide. Online testing is a major component of online and hybrid courses, but its use has also increased in traditional face-to-face courses. Online tests provide faculty and students with several advantages including more efficient exam administration, lower cost, instant feedback to both faculty and students, scheduling flexibility, in addition to time savings through automatic recording of grades. However, one major concern relates to the equivalence and validity of online exams as compared to in-class face-to-face exams. Prior studies have found mixed results. The purpose of this study is to compare student performance on unproctored open book online exams taken at home with performance on proctored open book online exams taken in class for an undergraduate accounting course.

To address our research question, students in an intermediate accounting course participated in our study. Our results show that the average student performance in unproctored online exams is higher than that of online exams proctored in class. The difference is statistically significant. Additional analysis comparing student grades by type of questions, by AACSB category, and by question difficulty provide further evidence that students taking online exam at home significantly outperformed those taking online exam in class.

Findings from our study extend research regarding the equivalence and validity of online exams as compared to in-class face-to-face exams and shed additional insights on the effect of test format on student test performance. The results of this study are particularly important at a time when universities have moved to virtual instructions as a result of COVID-19. Faculty around the world have expressed concern over the validity of their online assessments. With the sudden move to virtual instruction mid-semester, it is expected that faculty may see higher performance from students this semester as compared to the same time period last year due to the difference in the

format of the administered exams. However, our participants were all students in an intermediate accounting course. Further research is needed to see if our results hold in other disciplines.

References

Alexander, M. W., Bartlett, J. E., Truell, A. D., & Ouwenga, K. (2001). Testing in a computer technology course: An investigation of equivalency in performance between online and paper and pencil methods. *Journal of Career and Technical Education*, 18, 69-80.

Babson Survey Research Group. (2017). Higher education reports: Online and distance education. Retrieved May 11, 2018, https://www.onlinelearningsurvey.com/highered.html

Bonham, S. (2007). Reliability, compliance and security of web-based pre/post testing. In *Proceedings of 2006 Physics Education Research Conference*, American Institute of Physics, 133-136.

Brallier, S.A., & Palm, L.J. (2015). Proctored and unproctored test performance in traditional and distance courses." *International Journal of Teaching and Learning in Higher Education*, 27(2), 221-226.

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*, 25, 738-742

Clariana, R., & Wallace, P. (2002). Paper–based versus computer–based assessment: Key factors associated with the test mode effect. *British Journal of Educational Technology*, 33, 593-602.

Cochran, J. D., Baker, H. M., Benson, D., & Rhea, W. (2016). Business student perceptions of online learning: Using focus groups for richer understanding of student perspectives. *Organization Management Journal*, 13(3), 149-166.

Daymont, T., Blau, G., & Campbell, D. (2011). Deciding between traditional and online formats: Exploring the role of learning advantages, flexibility and compensatory adaptation. *Journal of Behavioral and Applied Management*, 11, 156-179.

DeAngelis, S. (2000). Equivalency of computer-based and paper and-pencil testing. *Journal of Allied Health*, 29, 161-164.

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.

Frein, S. (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.

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.

Harmon, O. R., & Lambrinos, J. (2008). Are online exams an invitation to cheat? *Journal of Economic Education*, 39 (2). 116-125.

Hollister, K. K., & Berenson, M. L. (2009). Proctored versus unproctored online exams: Studying the impact of exam environment on student performance. *Decision Sciences Journal of Innovative Education,* 7(1), 271-294.

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..

Lee, G., & Weerakon, P. (2001). The role of computer-aided assessment in health professional education: A comparison of student performance in computer-based and paper-and-pen multiple choice tests. *Medical Teacher*, 23, 152-157.

Lee, J., Moreno, K., & Sympson, J. (1986). The effects of mode of test administration on test performance. *Educational and Psychological Measurement*, 46, 467-474.

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.

Stowell, J. D., & Bennett, D. (2010). Effects of online testing on student exam performance and test anxiety. *Journal of Educational Computing Research,* 42 (2), 161-171.

Wallace, P., & Clariana, R. B. (2005). Test mode familiarity and performance-Gender and race comparisons of test scores among computer-literate students in advanced information systems courses. *Journal of Information Systems Education*, 16, 177-182.