Authenticated Testing during Blended Delivery: Impacts on Assessment Scores within an Engineering Undergraduate Core Course

Marino Nader[1], Ronald F. DeMara[2], Baiyun Chen[3] and Adrian Tatulian[2]

[1]Department of Mechanical and Aerospace Engineering, [2]Department of Electrical and Computer Engineering, and [3]Center for Distributed Learning

University of Central Florida, Orlando, FL 32816-2362

Abstract

Sustaining sufficient levels of academic integrity has been a longstanding challenge within various online delivery modalities. The impact of requiring authenticated high-integrity testing was measured in a case study of a required undergraduate mechanical engineering course (*Engineering Analysis: Dynamics*) with 264 students. Herein, proctored quizzes were compared to un-proctored quizzes in a mixed-mode setting. The proctored quizzes took place with a lockdown computer-based assessment (CBA) at the Evaluation and Proficiency Center (EPC), an on-campus testing and tutoring facility that delivers quizzes authored in the Canvas Learning Management System (LMS). Results indicated statistically significant differences between proctored and un-proctored quizzes with the least difference of 41% attributed to the amount of possible cheating that possibly occurred among students when assessments are un-proctored, even for a timed quiz. A Mann-Whitney test indicated that significantly more students agreed that take-home quizzes inflate grades. Data indicated proctoring in lower courses causes higher success in upper classes.

Keywords

Authenticated submission, computer-based assessment, lockdown browser, un-proctored testing.

Introduction

The use of online/blended courses in STEM curricula has increased steadily in recent years due to benefits ranging from convenience to efficacy, as overviewed by K. K. Hollister et al.[1] and D. L. Prince et al.[2]. However, significant challenges associated with various modes of online/blended delivery revolve around student authentication and the confidence in the accuracy of remote assessment. R.W. Yates et al.[3] reported no significant differences between proctored assessment results and un-proctored assessments results. Meanwhile, a controlled study by M. Nader[4] et al. reported very significant differences. Others such as M. M. Lanier[5], A. Fask et al.[6], as well as P. Charlesworth et al.[7] found potential differences as well. These findings vary due to circumstance where assessment may not demand as much proctoring for creative research works. However, contract cheating by hiring someone to do the work, password sharing, and even jointly-conducted work can occur. For example, working on a spreadsheet such as Excel may utilize reduced proctoring[3], although student authentication and lack of collaboration remain open issues. In order to mitigate the complication of academic dishonesty for blended mode delivery, one could allow remote students to complete their quizzes or tests in approved testing centers close to where they reside or work. Moreover, a comprehensive solution could require students to take assessments in

an interwoven testing and tutoring facility, such as the Evaluation and Proficiency Center (EPC), described by T. Tian[8] and R. DeMara, which can also facilitate controlled studies to compare proctored and un-proctored performance, as outlined by M. Nader[4] et al. and V. Beck[9]. Such an approach is undertaken herein. The purpose of this article is not only to emphasize the latter findings via *Engineering Analysis: Dynamics* delivered in a blended mode during Summer 2019 for 264 students, but also to realize the effects of the EPC in the students' success in subsequent courses[10]. Whereas academic integrity is vital for the academic institution's obligation and perhaps even legal exposures[11], controlled studies of the proctoring impacts of engineering assessments can provide useful insights into these essential processes for online/mixed-mode delivery.

Course Overview

The course under investigation herein was designed for a blended format with LearnSmart (LS) of Connect-McGraw Hill[12] as weekly pre-assignments as well as pre-prepared YouTube videos with questions that the students did before their weekly two-hour lectures. During the weekly meetings, the instructor solved 8 to 10 problems to reinforce the students' knowledge they had acquired by the pre-assignments. In this summer course, there were four quizzes and four tests. Each quiz preceded the corresponding test by a week. These quizzes served three purposes: 1) allowed the students to realize their standing in the course before taking more rigorous high stakes tests, 2) encouraged them to better prepare for the tests and, 3) provided a means for comparison between proctored and un-proctored quizzes. All assessments were electronically graded, be it via the Connect-McGraw Hill or via Canvas LMS.

The class was divided into two Groups, A and B. On a given week, when Group A had a proctored quiz in the EPC, Group B had a take home quiz and vice-versa, thus allowing for comparison between proctored and un-proctored quizzes. Overall, the students had two proctored quizzes and two un-proctored quizzes to ensure overall grading fairness to all students. All tests and quizzes were allotted 90 minutes. High-stakes tests were delivered in the EPC testing center.

Assessment Delivery Mechanisms

The students were given a comfortable, anxiety free 3-day window in which they could complete their tests or quizzes in the EPC, which delivered tests from 9:00AM – 9:00PM daily and asynchronously via 120 computer-based workstations, R. DeMara[13] et al. Test monitoring is performed via cameras and in-room proctors during exams. Moreover, the EPC provided a quiet and comfortable setting with Lockdown Browser that restricts all features of computer usage except for test delivery. Once students arrived at the EPC, they placed their belongings in lockers and checked in with their ID. They were immediately handed a piece of scratch paper on which they wrote

Figure 1: Student perception that online exams cause grade inflation.
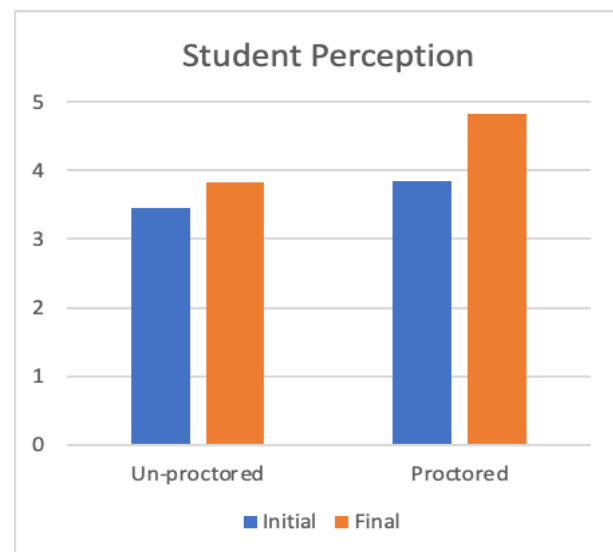
their name, ID numbers and quiz/test number. On this scratch paper, students completed work supplementary to their quizzes or tests. Once started, the students worked on their quiz with a maximum allotted time of 90 minutes during which they could do nothing other than solving their quiz.

Student Surveys

Students were surveyed on their perception of EPC-based testing both at the beginning and the end of the course. Students had five answer choices for each survey item, ranging from "Strongly Disagree" (1) to "Strongly Agree" (5). A Mann-Whitney test performed at the 5% significance level demonstrated increased agreement at the end of the course, among both the proctored and un-proctored cohorts, that online exams taken at home may cause grade inflation. This shift in perception is illustrated in Figure 1.

Results

Canvas LMS was used to randomly disburse problems to students from question banks derived from the *Vector Mechanics for Engineers: Dynamics* textbook by P.J. Cornwell et al.[15] to each quiz. To ensure fairness, each quiz problem was selected from a pool of problems deemed to have comparable difficulty while requiring identical skills. The quizzes' average results for proctored and un-proctored quizzes are depicted in Figure 2. For the same quiz 1, it indicates that when Group B was un-proctored, students in that group scored 42 points higher than Group A. Conversely,

Figure 2: Student scores in proctored and un-proctored modes.
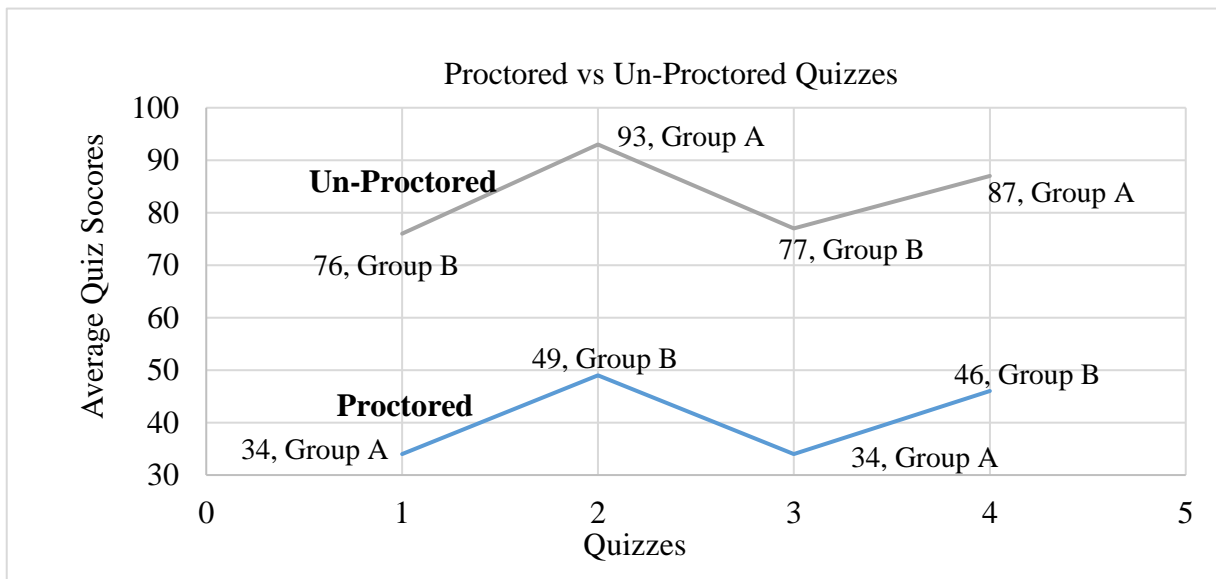
the opposite situation occurred when Group A was un-proctored which obtained 44 points higher, on average, in quiz 2. Figure 2 shows the consistent trend, indicating that when the students conducted a take-home quiz for the same 90-minute duration as that delivered in the proctored environment of the EPC, their grades were very highly inflated, implicating significant changes in

outcomes. The average difference of all four quizzes is 42.5%, which is alarming in the absence of other unnoticeable causes. Independent-sample t-tests were conducted to compare scores between these quizzes. Statistically significant differences in the scores are presented in Table 1.

Table 1: Means & Standard Deviations of Quiz Scores

		Group A	**Group B**
Q 1	Mean	1.68	3.82*
	SD	1.43	1.20
	N	126	124
Q 2	Mean	4.63***	2.43
	SD	0.55	1.21
	N	116	113
Q 3	Mean	1.70	3.85*
	SD	1.49	1.30
	N	115	110
Q 4	Mean	4.34**	2.31
	SD	0.86	1.21
	N	109	102

***p<0.001, **p<0.01, *p<0.05

Intervention Effects

When delivering some parts of instructional content online, it is a legitimate issue to consider the cost versus benefit tradeoffs of proctored testing and secure post-test review such as the environment provided by the EPC in this study. There are many considerations to evaluate with the context of each course. They include various aspects of assessment including the fact that by motivating students to study for their exams, EPC-based delivery sustains the authority and impact of exams. These translate into institutional benefits of credibility and accreditation of the programs offering the courses and hence the university's prestige, and that of the discipline more broadly, as a whole. Another compelling aspect is sustaining competencies expected by engineers as they progress through the course sequence, and eventually graduate upon completing the curriculum. To substantiate the latter, a study was done to verify the percentage of students who obtained DFW in other courses that took the *EGN 3321: Dynamics* course as a pre-requisite. Among these courses that have been investigated are *EML 3262: Kinematics of Mechanisms* and *EML 4225: Introduction to Vibrations and Controls*. Student-resolution longitudinal results indicate that providing the EPC intervention, i.e. from Spring 2018 through Spring 2019 inclusive, the learning outcomes of those students who received the EPC intervention in the prerequisite performed higher in subsequent courses taught by other faculty. In other words, the students who took Dynamics during Spring 2018 – Spring 2019 and were tested in the EPC exhibited higher achievement and rates of retention in the follow-up courses. To clarify, though, the tests or quizzes, before the intervention were proctored in class and after the intervention in the EPC.

Table 2: DFW rate comparison before and after the EPC intervention.

Courses	DFW rate Before the Intervention	Students No.	DFW rate After the Intervention	Students No.	Percentage Improvement
EML 3262	10.9%	46	1.8%	57	83%
EML 4225	12.8%	148	10.7%	215	16%
Average	12.8%	194	8.8%	272	31%

Table 2 shows the average percentage of the sum of D grades, F grades, and W grades (DFW rate) of a few semesters before the intervention as well as the average DFW rate after the intervention. When these two results are compared, it is observed that the DFW rate becomes lowered significantly, at least for EML 3262 and moderately for EML 4225 with an average DFW rate improvement for these two courses by 31%. As listed in Table 2, the students were larger in number after the intervention. If the same number of students were available before the intervention for comparison, the percentage improvement would be higher. This is inversely correlated with a higher DFW rate. Consequently, there is a high confidence that students who took the prerequisite courses' quizzes and tests in the EPC did so in a trusted, secure, and robustly authenticated setting. This cohort succeeded at statistically higher rates in their upper level courses as listed in Table 2. Another possible contributor to this success is the on one-to-one basis tutoring service the EPC provides for each individual student at a large scale.

Discussion and Conclusion

Based on the above results from a controlled study, un-proctored results were routinely much higher scoring. Possible grade inflation was caused by various unauthorized activities by students enabled by online delivery including solution banks, community solution services using Chegg[14,16] or contract cheating consultants, and others. Perhaps this was made possible by the fact that the problems were either directly from the textbook[15] or simplified derived versions. Unrealistic results, therefore, contaminated the assessments and prevented an accurate assessment of students' learning gains. Fortunately, the take-home quizzes contributed to only 10% of the total grade. If unchecked, such degradation of assessment integrity could erode the academic standing of individual students or even entire degree programs at an institution. Risks can become exacerbated as the proportion of the un-proctored assessments and take-home tests increases, especially in undergraduate courses in STEM fields where subsequent coursework builds upon the previous knowledge that the student has or has not acquired.

A further study on how the students performed in later courses that required the Dynamics course as a pre-requisite is being conducted. Nonetheless, the fact that the EPC provides tutoring on a one-to-one basis can facilitate the increased use of additional remedial mechanisms for post-test review to elevate student success. These roles extend beyond the stringent testing integrity the EPC imposes on the students which prevents them from cheating, encourages them to prepare better for successful results and protects the credibility of the university.

Achieving and sustaining a robust testing environment is not without its own costs. The amount of work in preparing large question banks can be significant. The cloning of problems, conceptual and numerical, is key to ensuring fairness of problems in their level of difficulty as an endeavor to maintain the same level of cognitive learning in all students. Overall, the assessment-resolution controlled study of proctoring impacts, along with corroborating longitudinal data, substantiate the need for further investigation of proctoring impacts in online/mixed delivery modalities.

References

1. 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*, vol. 7, No. 1, 271- 294.
2. Prince, D. J, Fulton, R. A, & Garsombke, T. W. (2009) Comparisons Of Proctored Versus Non-Proctored Testing Strategies In Graduate Distance Education Curriculum, *in Journal of College Teaching & Learning*, vol. 6, No. 7, 51-62.
3. Yates, R.W., & Beaudrie, B.(2009) The Impact of Online Assessment on Grades in Community College Distance Education Mathematics Courses, *The American Journal. of Distance Education*, 23:2, 62-70, DOI: 10.1080/08923640902850601
4. Nader, M., DeMara, R. F., Tatulian, A., & Chen, B. (2019). Quantitative impact on learning achievement by engaging high integrity testing using lockdown assessment for online delivery. Proceedings of the ASEE Southeast Section Conference, No. 45. This work was published and presented in March, 2019. http://www.asee-se.org/proceedings/ASEE2019/papers2019/45.pdf
5. Lanier, M. M.(2006). Academic Integrity and Distance Learning, *Journal of Criminal justice Education*, 17:2, 244-261, DOI: 10. 1080/10511250600866166
6. Fask, A., Englander, F., & Wang, Z. (2014). Do online Exams Facilitate Cheating? An Experiment Designed to Separate Possible Cheating from the Effect of the Online Test Taking Environment. *J Acad Ethic*, 12:101–112 DOI 10.1007/s10805-014-9207-1
7. Charlesworth, P., Charlesworth, D.D., & Vician, C. (2006) Students' Perspectives of the influence of Web-Enhanced Coursework on Incidences of Cheating, *Journal of Chemical Education*, vol. 83 No.9.
8. Tian, T., & DeMara, R.F (2018) High-Fidelity Digitized Assessment of Heat Transfer Fundamentals using a Tiered Delivery Strategy," in Proceedings of American Association for Engineering Education Annual Conference (ASEE-18), Salt Lake City, UT, USA.
9. Beck, V. Testing a model to predict online cheating- much ado about nothing, Active Learning in Higher Education 2014, Vol. 15(1) 65 –75.
10. DeMara, R. F. "Technology-Enhanced Learning in CECS Outcomes & Opportunities for Instructional Excellence," in Digital Learning Panel Session, UCF Board of Trustees, Orlando, FL, USA, September 19, 2019.
11. A. Amigud, J. Arnedo-Moreno, T. Daradoumis and A. Guerrero-Roldan (2018). An integrative review of security and integrity strategies in an academic environment: Current understanding and emerging perspectives, Elsevier, Computer & Security, Vol. 76, July 2018, p.50-70.
12. McGraw-Hill Education., website: https://www.mheducation.com, accessed on 8 January 2020.
13. DeMara, R. F., Turgut, D., Nassiff, E. Bacanli, S. Bidoki, N. H. and Xu, J. (2018) "Automated Formation of Peer Learning Cohorts using Computer-Based Assessment Data: A Double-Blind Study within a Software Engineering Course," in Proceedings of American Association for Engineering Education Annual Conference, Salt Lake City, UT, USA, June 24 – 27, 2018.
14. Harmon, O. R., & Lambrinos, J. (2008). Are Online Exams an Invitation to Cheat?. *The Journal of Economic Education*, 39:2, 116-125, DOI: 10.3200/JECE.39.2. 116-125.
15. Phillip J. Cornwell, Ferdinand P. Beer, E. Russell, Jr. Johnston and Brian Self (2015). Vector Mechanics for Engineers: Dynamics, 11th Ed. McGraw-Hill Education, P.O. Box 182605, Columbus, OH 43218, https://www.mheducation.com
16. Chegg Inc., website https://www.chegg.com , accessed on 8 January 2020.

Marino Nader

Marino Nader is an Associate lecturer in the Mechanical and Aerospace Engineering Department at the University of Central Florida and has been working on digitizing courses and exams, creating different course modalities. Dr. Nader obtained his B.Eng., M.Eng. and Ph.D. from McGill University. His Ph.D. was done in conjunction with the Canadian Space Agency where he spent two years doing research and experiments. Upon completion of his Ph.D. he began working in the Aerospace Industry where he spent over 10 years as a Stress Analyst/Consultant. Most of his work is in digital learning. At present, he enjoys working on Distributed Electric Propulsion (DEP) with his students, designing, analyzing, constructing and flying Unmanned Aerial Vehicles.

Ronald F. DeMara

Ronald F. DeMara is a Professor of Electrical and Computer Engineering and the Digital Learning Faculty Felllow at UCF where he has been a faculty member since 1992. His educational research interests focus on classroom and laboratory instructional technology, and the digitization of STEM assessments. He has completed approximately 300 technical and educational publications, 48 funded research projects, and 50 graduate students with Ph.D. dissertation or MS thesis. He serves as the founding Director of the Evaluation and Proficiency Center (EPC) at UCF, and received the Effective Practice Award from the Online Learning Consortium and the Joseph M. Biedenbach Outstanding Engineering Educator Award from IEEE.

Baiyun Chen

Baiyun Chen is a Senior Instructional Designer at the Center for Distributed Learning at UCF. She designs and delivers faculty professional development programs and teaches graduate courses on Instructional Systems Design. Her research interests focus on using instructional strategies in online and blended teaching and learning, professional development for teaching online, and application of emerging technologies in education. She has published 23 peer-reviewed journal articles and book chapters and delivered more than 60 presentations at international and local conferences and event and served as the Co-Managing Editor of the Teaching Online Pedagogical Repository.

Adrian Tatulian

Adrian Tatulian is currently pursuing a PhD in Computer Engineering at the University of Central Florida. His research interests include low-power VLSI design, spintronic-based architectures, and STEM education. Tatulian has three years of STEM classroom teaching experience at the high school level.