The Impact on Learning Outcomes using Three-Attempt Tests in an Engineering Undergraduate Core Course: Dynamics

Marino Nader[1], and Ronald F. DeMara[2]

[1]Department of Mechanical and Aerospace Engineering, and

[2]Department of Electrical and Computer Engineering

University of Central Florida, Orlando, FL 32816-2362

Abstract

During the COVID-19 pandemic, the virtual M-mode *Dynamics* class of 236 students was delivered weekly via zoom. Herein, three tests were given using *Proctor Hub, Respondus monitor and LockDown Browser*, for the students' assessments. To preserve high-integrity testing, for every attempt, the students were permitted to use only one screen, they showed their faces to the camera and what they were always writing. For every test, they were allotted 90 minutes per attempt and a week to finish all three attempts. Randomized questions were delivered from large pools of problem banks. Grading was facilitated immediately upon students' submission of their tests through Canvas Learning Management System (LMS) and computer-based assessment (CBA). The results indicated an average score improvement from the first of three attempts as compared to the final attempts exceeded 30%. Additionally, when asked for their anonymous perceptions regarding the use of a three-attempting assessment approach, more than 94% of the responding students either strongly agreed or agreed that this testing style helped them learn more effectively than a single-attempt assessment.

Keywords

Respondus monitor, LockDown Browser, computer-based assessment, three-attempt tests, Zoom.

Introduction

Many junior and senior-level university students enrolled in engineering majors continue to actively refine their skills to study and prepare themselves for tests. Fortunately, there are a range of approaches to tackle each and every different quiz and exam. Meanwhile, there are assignments guides that help on solving homework such as Chegg[1] and others, which inadvertently make the students think they learned enough when they actually did not. This illusive impression deters the students from preparing themselves appropriately and sincerely for the tests ahead. In wonderment, students are surprised and confused when they get low grades. For the sake of practice and learning, Connect-McGraw Hill allows multiple attempts homework, K. K. Archer[2]. In general, students don't take the assignments as seriously as they do their tests. On the other hand, instructors may not provide an opportunity of multiple-attempt tests whereas each test is typically considered to comprise a portion of the students' assessment, assuming the students are ready for the tests and expected to do well.

Unfortunately, during the COVID-19 pandemic many students were faced with additional studying, testing and remediation challenges. Y. Terada[3] studied the motivation-levels during the pandemic and concluded that motivation during the pandemic era has affected their tests negatively, compounded by the reality that many do not know how to prepare adequately under less challenging circumstances. The question is why don't we take a similar approach in testing them as we do with the assignments? One answer may be a result of impacts to the integrity of the examination and the institution as a whole to avoid cheating, which M. M. Lanier[4], A. Fask et al.[5], as well as P. Charlesworth et al.[6], discussed to avoid grade inflation and prevent any indignant reputation that could be brought to the degree program. G. Herman[7] has examined a second chance testing impact on students' learning outcomes and found an overall positive improvement.

Herein, we emphasize both integrity and students' success via multiple-attempt testing as a method of learning and bringing the students to higher level of knowledge, since they take their tests more seriously, and since the method of repetition works on their favor for their education. Multiple-attempt testing seems to be a good medium to harbor and foster learning, via repetition, retention as per Arora et al.[8] hope and satisfaction. It is noteworthy to mention that digital learning described by M. Nader et al.[9] facilitated this operation since the tests grading were conducted automatically via Canvas LMS and CBA.

Course Summary

The subject course of this paper is junior-level Dynamics, which is a required Mechanical Engineering and Aerospace Engineering course. It was previously designed as a blended virtual and live mixed-mode delivered course, M. Nader et al.[10]. When it was delivered during Spring 2021, in which students did weekly pre-assignments, such as LearnSmart (LS) – Connect - McGraw Hill and a video HW before they attended the 1.5-hour weekly virtual lectures. All the videos were created by the author, deposited into YouTube, and were all linked to Canvas LMS. The virtual lectures were focused mostly on solving a few problems similar to those in their after-class assignments. After attending the live lectures, the students become more ready for the more rigorous Connect assignments, whose problems were chosen from the course textbook by P.J. Cornwell et al.[11] These assignments were due in a couple of days before their e-tests started, which served as a review for their tests. An hour after the assignments were due, the solutions were out as feedback for the students to learn from their mistakes in preparation for the tests. This was a gradual knowledge building and scaffolding as per L. A. Fish[12], between the pre-assignments, via the lectures and through the rigorous assignments in preparation for the tests. Using CBA, all of the pre-assignments, assignments and tests were graded. In this course, there were three tests offered which were to be finished within a week, while allowing three attempts each and 90 minutes per test, while retaining the highest score.

Why Three attempts?

Often students may be apprehensive or feel stressed when required to sit for a test for the first time due to their lack of not knowing what to expect, unaware of the examination style and therefore, lack confidence, which in turn raises their level of anxiety. Many students insisted that the tests grades did not reflect their true knowledge of the course because the tests were like a miss or a hit,

but the three-attempt tests mitigated this issue. At least, it is better to repeat the tests than the entire course. As per the survey, about 98% of the students surveyed responded that they agreed that they felt more comfortable taking the first attempt, knowing they had more chances to do better in later attempts. Perhaps later tests tended to make them less worrisome as they became more used to the assessment style and flow. As for the tests, they were all conducted remotely using Lockdown browser, Respondus Monitor, and Proctor Hub. Each attempt was allotted 90 minutes. Students were told to allow 2 ~ 3 feet between the computer screen and what they were writing so as to show their work as it is happening. This was possible by pushing their computers' screens forward by 2 ~ 3 feet while deflecting the camera to show their faces and the desk where they were conducting their calculations, otherwise, their tests were invalid. The camera footage was later spot-checked for identification of concerns of integrity. This high-level of strictness, which was imposed during delivery, was undertaken to elevate towards a high-integrity testing environment.

Students were allowed one week to finish three attempts for each of a total of three tests (T1, T2 and T3). On average, 10 problems were randomly selected per attempt from a large set of question banks (over 250) creating minimum chances of repetition in sequential attempts, thus minimizing the chances of cheating, C.J. Lee[13]. Yet, if a problem repeated itself in any one of the attempts, the numbers within the problem did not repeat due to their high variation range within the problem statements. Nonetheless, students felt encouraged to learn what students missed, or what students felt uncomfortable with during the tests, just in case a similar question would show up in their next attempt with different numbers. As such, students were also allowed to ask the instructor and the Teaching Assistants (TA) about problems they saw in any of the attempts before their next attempt. The instructor/TAs explained these problems without solving them. It was the students' responsibility to solve them. This feedback was an important step in the students' learning with tease, enticement and engagement, fostering a high-learning ambiance. To examine the students thoroughly different types of problems styles were utilized described by T. Tian & R. F. DeMara[14], ranging from various concepts and simple calculations not only utilizing multiple choice questions, Marsh et al.[15], but True/False, and numerical calculations. To avoid bias grading, each problem was pulled from a question bank of a similar difficulty, Nader et al.[10], i.e. for each student, the first question would come from one particular bank with easy problems, the second would be pulled say from a moderately difficult bank and so on. Thus, ensuring fairness to all students. Out of the three attempts, the highest score was then recorded and retained for that test with no consideration for partial credit.

Student Surveys

Out of the 236 students about 200 participated in the survey and over 93% of those 200 students agreed that overall, the three-attempt testing method is an effective method of learning. For further information, the reader is referred to Appendix A.

Results

The tests results are depicted in Figure 1. It is evident that for each test (i.e. curves indicated for T1, T2, and T3) the average score for each attempt increases progressively. For example, the average class score of the first attempt for T1 is 47% that becomes elevated to 56% in the second

attempt, while the mean score attains 62% in the third attempt. Similarly, performance gains are noted for the other two tests, namely, T2 and T3. However, T3 seems to reach a plateau where the students could not improve as significantly as in T1 and T2. In essence, it is clear that the score improvement from the first attempt to the last in T1 translates to 31% improvement and that in T2 50%, as well it is 21% in T3 (see Table 1).

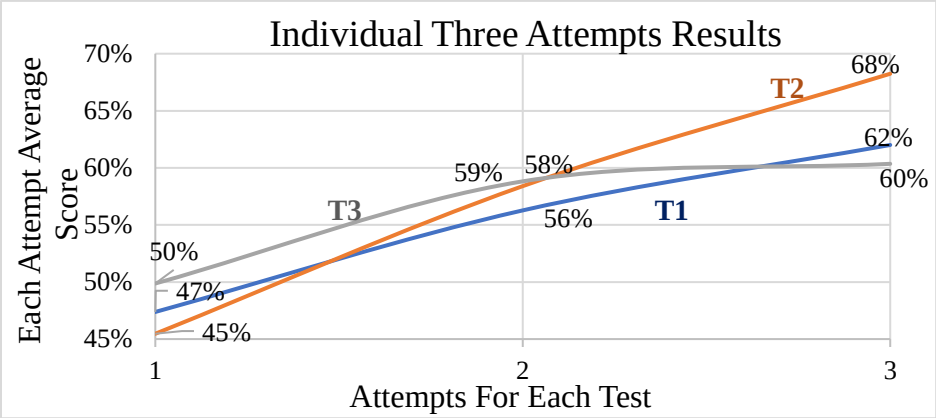

Figure 1: Individual Tests with the Corresponding Three Attempts Results.

Table 1: Class' Averages Using Three-Attempt Tests

Test	Attempt 1	Attempt 2	Attempt 3	Percentage Improvement
T1	47%	56%	62%	**31%**
T2	45%	58%	68%	**50%**
T3	50%	59%	60%	**21%**
Class average	**48%**	**58%**	**64%**	**34%**

Moreover, if we consider that the students were allowed only one attempt per test the class average would be 48% as in Table 1. Yet, with the three attempts the class average improves by 34% for all three tests, indicating a significant students' success, represented graphically in Figure 2.

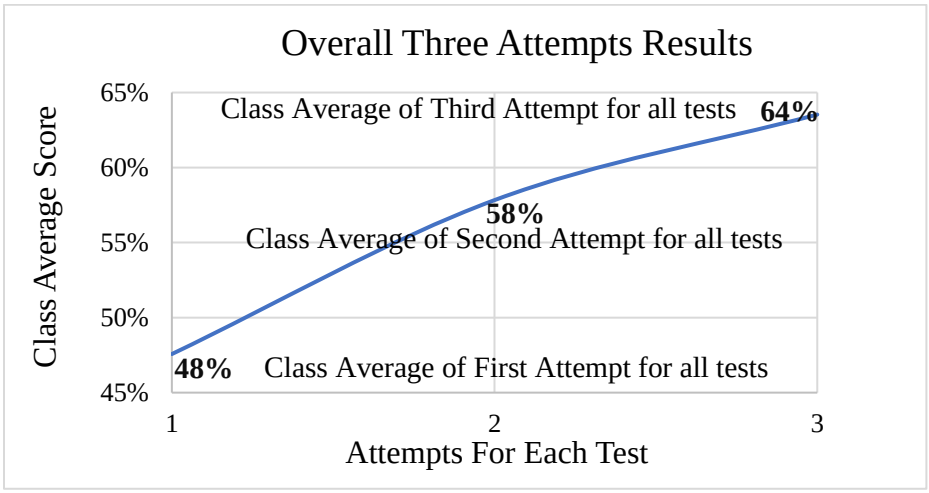

Figure 2: Class Average for each Attempt

Discussion and Conclusion

It is intriguing to investigate the role that allowing repeat attempts of assessments can have on students' learning outcomes. When automated with Computer-Based Delivery, such can be considered with minimal additional human effort. In general, it seems as students repeat the test, then higher learning outcomes are attained *assuming that students sufficiently prepare* themselves before each and every attempt. There is no clear evidence that this assumption is completely true for all students. However, at least it can be said that these repetitions enhance the students' knowledge, expand their understanding and improve their grades. The overall grade improvement is significantly high, i.e. 34% for this course, indicating that the three-attempt testing method, which at some pro-active level has a mechanism of enticing learning by repetition, offers a viable method to embrace. A consortium of researchers can further investigate, refine, and validate this assumption using additional data obtained from various courses.

When surveying students agreed on the fact that the three-attempt tests made them more comfortable, fostered a learning environment, learned more, and gave the students a higher satisfaction. The three-attempt tests also increased the students' grade and prepared them better for higher courses based on their surveyed perceptions.

Appendix A

Statement	Strongly Agreed	Agreed
Helped me take the test with less stress, knowing I have other chances.	85%	13%
Allowed me to go back to learn the material better before my next attempt, given a week-time for all 3 attempts.	82%	15%
It gave me the opportunity to know where I stand, before my next attempt.	80%	15%
It gave me the chance to recognize how much more I should learn before my next attempt.	80%	16%
It gave me the opportunity to focus on my weakness in order to do better.	78%	15%
In a way, I prefer the 3-test attempts for I rather repeat the test than the course.	91%	8%
It gave me the chance to repeat the test instead of the repeating the entire course.	85%	12%
The fact that I could go back and ask about a problem I saw in the test to learn it before my next attempt advanced my knowledge of the subject, despite the fact that it would or would not show up in my next attempt.	78%	18%
It assisted me to grow in knowledge.	76%	19%
It allowed me to do better in the course.	78%	16%
It created a learning-based environment.	75%	20%
I learned a great deal using this method, irrespective of my grade.	73%	21%
In comparison to the one attempt, the 3-test attempts is not a hit or a miss.	61%	24%
In comparison to the one attempt, the 3-test attempts assess the student's knowledge correctly, given the few chances during a week.	69%	25%
It is useless because no matter how much I tried, I still got the same grade.	1%	3%
It is useless because it opened up for cheating without learning.	0%	2%
I prefer the 3-test attempts it to help students learn the material better.	80%	18%
All-in-all the 3-test attempts is a more fun learning style.	68%	22%
All-in-all the 3-test attempts is a more enjoyable learning style.	74%	21%
All-in-all the 3-test attempts is my worst experience, since I did not learn much more.	0.5%	1%
All-in-all the 3-test attempts is my worst experience, since I did not improve my grade much.	0.5%	3%
In the future, I hope to see more courses offered with 3-test attempts during a full week.	81%	12%

References

1. Chegg Inc., website https://www.chegg.com , accessed on November 7, 2021.
2. Archer, K. K. (2018). Do Multiple Homework Attempts Increase Student Learning? A Quantitative Study. The American Economist. 63(2):056943451877479, DOI:10.1177/0569434518774790. Colangelo College of Business, Grand Canyon University, 3300 W Camelback Road, Phoenix, AZ 85061-1097, USA.
3. Terada, Y. (2020). Covid-19's Impact on Students' Academic and Mental Well-Being. *Edutopia*, website: https://www.edutopia.org, accessed on November 7, 2021.
4. Lanier, M. M.(2006). Academic Integrity and Distance Learning, *Journal of Criminal justice Education*, 17:2, 244-261, DOI: 10. 1080/10511250600866166
5. 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
6. 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.
7. Herman, G. L., Cai, Z., Bretl, T., Zilles, C., & West, M. (2020, August). Comparison of Grade Replacement and Weighted Averages for Second-Chance Exams. *In Proceedings of the 2020 ACM Conference on International Computing Education Research* (pp. 56-66).
8. Arora, M. L., Rho, Y. Jin, & Masson, C. (2013). Longitudinal study of online statics homework as a method to improve learning. *Journal of STEM Education: Innovations and Research*, v14 No.1, p.36-44.
9. 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
10. Nader, M., DeMara, R. F., Tatulian, A., & Chen, B. (2021). Authenticated Testing during Blended Delivery: Impacts on Assessment Scores within an Engineering Undergraduate Core Course. Proceedings of the ASEE Southeast Section Conference, No. 53. This work was published and presented in March, 2021. https://sites.asee.org/se/wp-content/uploads/sites/56/2021/04/2021ASEESE53.pdf
11. 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
12. Fish, L. A. (2015). Undergraduate students computer-managed homework versus in-class performance for different testing formats. *Business Education Innovation Journal*, 7, 5-14.
13. Lee, C. J. (2018). Automated Randomization of Test Problems for Cheating Prevention. *World Journal of Research and Review (WJRR)*. ISSN:2455-3956, V.3, Issue-2, Feb. 2018, p.10-15.
14. 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.
15. Marsh, E. J., Roediger III, H. L., Bjork, R. A. & Bjork, E. L. (2007). The memorial consequences of multiple-choice testing. *Psychonomic Bulletin & Review 2007*, V. 14 (2), p. 194-199.

Marino Nader
Marino Nader is a 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. 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 Pegasus Professor in the Department of Electrical and Computer Engineering at the University of Central Florida (UCF), where he has been a full-time faculty member since 1993. His educational research interests focus on classroom instructional technologies and the digitization of STEM assessments. He is Principal Investigator (PI) of NSF IUSE-HSI grant on Building Undergraduate Capacity utilizing Culturally-Relevant Instruction with Micro-Credentialing and previously PI of the NSF Workshop on Digitally-Mediated Team. He has completed over 300 technical and educational publications, 50 funded projects as PI/Co-I, and 25 Ph.D. graduates. He serves as the founding Director of the Evaluation and Proficiency Center (EPC), was an iSTEM Fellow, and a Digital Learning Faculty Fellow at UCF and received the Joseph M. Biedenbach Outstanding Engineering Educator Award from IEEE.