Do proctored online University exams in Covid-19 era affect final grades respect face-to-face exams?

Antoni Alegre-Martínez[1], María Isabel Martínez-Martínez[2], José Luis Alfonso-Sanchez[3]

[1]Biomedical Sciences department, Faculty of Health Sciences, Cardenal Herrera CEU University, Spain, [2]Public Health department, Faculty of Nursing and Podiatry, University of Valencia, Spain, [3]Department of Preventive Medicine, University of Valencia, Spain.

Abstract

The Covid-19 pandemic forced universities to convert their traditional face-to-face exams to online exams with doubts as to whether student cheating or technical difficulties would affect their final grades. After taking three of these exams online, we considered comparing their grades with those of previous years on traditional exams. The average mark of the traditional exams before the pandemic was 6.95 over 10, while the average mark of the three exams carried out in the Covid-19 era is 6.64. The student's t test indicated that there are no significant differences between the two types of exams in the mean (p = 0.408), the median (p = 0.378), the range (p = 0.307), the minimum (p = 0.410) and the maximum (p = 0.072). Taking online exams did not modify the exam grades compared to previous years. There is a lot of variability in similar studies in the literature due to cheating that can be performed in online exams. A proctoring system, good question design, and limited exam time can minimize these differences.

Keywords: *Proctoring; online; exams; cheating; tests; grades.*

1. Introduction

The quarantines and lockdowns decreed by practically all governments during the spring of 2020 due to the Covid-19 pandemic forced universities around the world to quickly transform their teaching, practices, and exams to online training due to the impossibility of performing them face-to-face. One of the most difficult decisions was regarding exams: not holding them in classrooms with the supervision of professors would lead to cheating, fraud and copying between students using the multiple possibilities of information technologies. One of the possible solutions was the oral exams by videoconference, but they ran into many difficulties: impossibility of solving problems that require autonomous work or writing, finding questions of homogeneous difficulty or excessive examination time to evaluate hundreds of students (Haus *et al*., 2020).

Many Universities chose online exams, a technology that had been available for decades. As early as 2001, Alexander's team, realizing that many teachers were moving their exams to online platforms, warned about the need to maintain academic integrity through supervised online exams. (Alexander *et al*., 2001). It is easy with current technology to find the answers to the exam questions, and there are even companies that carry out the exam for the student (Daffin Jr & Jones, 2018). Academic dishonesty includes: conducting the exam by an expert instead of the student, several students taking an exam that should be individual, using prohibited sources such as books or Internet webpages (Reisenwitz, 2020), sheets with annotations, plagiarizing a work or copying exercises (Daffin Jr & Jones, 2018), waiting to take your own exam so you can get answers from other students, making false statements to be able to take the exam again (for example claiming that the connection failed), setting up two computers (one to perform the exam and another to find answers) or even buying the answers on the Internet (Moten Jr *et al*., 2013).

For this reason, numerous supervision software was developed based on the same principles: a camera on the student's computer that records its activity, allowing the professor (live or later) to observe if the student speaks, looks for information in books, or uses the mobile phone. Furthermore, these programs limit the functions of computers: they prevent opening the internet browser, calculator, copy and paste functions, etc. In this way, it is possible to control unethical behavior or cheating.

These programs do not always work well. Milone's team stopped using them due to ongoing technical problems experienced by students, the overall impact on the educational experience, excessive technical requirements, and additional costs for students (for example to buy a webcam), which to their judgment outweighed the benefits of avoiding cheating on exams (Milone *et al*., 2017).

One of the main doubts that online exams raise is whether they will impact students' grades, an issue on which our work will focus.

2. Objectives

Our goal is to check whether taking online exams has had an impact on students' grades. Although general research wishes to find differences between groups (alternative hypothesis), in this concrete case the good news would be that the implementation of online exams has not influenced students' grades, and therefore migration to online exams during the Covid-19 pandemic has not affected the performance of our students.

3. Material and methods

In the first place, we will carry out a retrospective search of the grades of the Anatomy III and Anatomy IV subjects taken by medical students at the Cardenal Herrera CEU University in Alfara del Patriarca (Spain). Both subjects are taught respectively in the first semester (autumn) and second semester (spring) by the same professor, which guarantees homogeneity in teaching and in the difficulty of the questions in the exam. In each subject, a midterm exam and a final exam are carried out and they are assessed on a total of 10 points. We collect the marks from the 2017-2018 academic year until the recent final exam held in January 2021 (the only exception is the Anatomy IV 2020 midterm exam which was canceled). The last three exams were taken online during the Covid-19 pandemic using proctoring tools, and the previous ten exams were taken in the classroom with teacher supervision. We will use the student's t test to compare the means between both groups of exams to check if there have been differences between both groups.

4. Results

Table 1 shows the main descriptive results of each of the 13 exams. The first three are the online exams while the next ten are traditional face-to-face exams. You may notice some small fluctuations between the number of students who take the midterm and final exam, they are exceptions based on several explanations: students who finally decide not to take the final exam, or who enroll the course late, or who cannot take the midterm for whatever justified cause.

At first glance, the average has oscillated in each course between values close to 6 or 7, every year there have been outstanding marks of 9 or even 10 and every year there have been failed students below 5. It is striking that in the last exam in January 2021 the range has decreased a lot, finding all the marks between 4.27 and 9.73.

Table 1. Measures of central tendency that describe the different exams.

	Delivered exams	Average	Median	Standard deviation	Variance	Range	Niminum	Maximum
Final Anatomy III 2021	59	6,98	6,93	1,46	2,14	5,47	4,27	9,73
Midterm Anatomy III 2020	58	6,28	6,71	2,06	4,25	8,40	1,35	9,75
Final Anatomy IV 2020	55	6,65	6,80	1,77	3,13	8,13	1,60	9,73
Final Anatomy III 2020	59	5,93	5,67	2,12	4,51	9,53	0,47	10,00
Midterm Anatomy III 2019	59	6,16	6,30	1,99	3,97	8,10	1,90	10,00
Final Anatomy IV 2019	47	7,14	7,47	2,03	4,12	7,07	2,93	10,00
Midterm Anatomy IV 2019	49	7,71	8,20	1,85	3,44	7,20	2,80	10,00
Final Anatomy III 2019	51	7,14	7,50	1,81	3,26	7,60	2,40	10,00
Midterm Anatomy III 2018	50	6,60	7,30	2,18	4,74	9,70	0,30	10,00
Final Anatomy IV 2018	61	7,03	7,33	1,88	3,54	8,13	1,60	9,73
Midterm Anatomy IV 2018	60	7,88	8,29	2,00	3,98	9,17	0,83	10,00
Final Anatomy III 2018	63	7,17	7,27	1,65	2,71	6,81	2,73	9,54
Midterm Anatomy III 2017	61	6,85	7,10	2,01	4,06	8,20	1,80	10,00
Total online tests	172	6,64	6,86	1,79	3,22	8,40	1,35	9,75
Total traditional tests	560	6,95	7,29	2,03	4,10	9,70	0,30	10,00

Source: self-made

Figure 1 represents in a boxplot the distribution of the exams for a quick easy visualization. At first glance, is not observed a great variability among the three tests of the Covid-19 era.
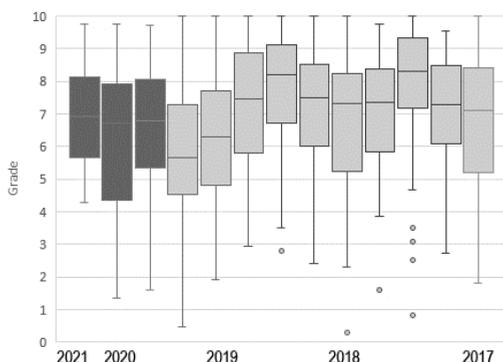

Figure 1. Boxplot of each one of the exams, showing retrospectively the range and median. The first three dark on the left are the on-line tests. Source: self-made.

We can group the grades of all students into two groups: online exams and traditional exams, thus obtaining the last two rows of Table 1 and graphically expressed in Figure 2.
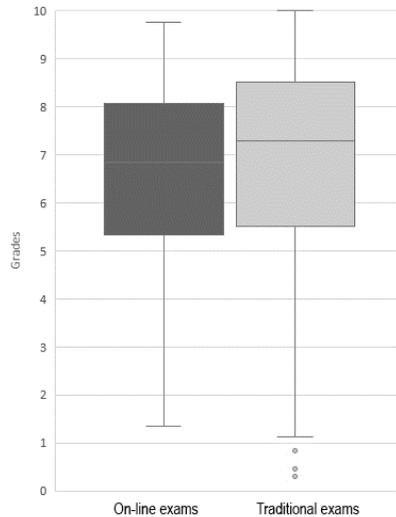

Figure 2. Boxplot of all exams grouped in two categories: On-line and traditional. Source: self-made.

The average mark of the traditional exams carried out before the pandemic was 6.95, while the average mark of the three exams carried out in the Covid-19 era is 6.64. To check if this difference is statistically significant, we have performed the student's t test to compare the different parameters using the categorical classification variable "online exam" or "traditional exam". In all the applied parameters, Levene's test indicated that equal variances were assumed, and the two-sided p-values for each of the parameters were not significant. That is, for 11 degrees of freedom, the mean (t = -0.860, p = 0.408), the median (t = -0.919, p = 0.378), the range (t = -1.072, p = 0.307), the minimum (t = 0.857, p = 0.410) and the maximum (t = -1.993, p = 0.072) do not represent a statistically significant difference between the two groups. This statement allows us to conclude as the main objective of the study, that the fact of conducting online examinations did not modify the exam parameters with respect to previous years.

5. Discussion

One strength of the study is to have compared it with exams given by the same teacher to different groups of students with a homogeneous level of difficulty and the same subject for all. One limitation is having done it on only three exams, so it is likely that as students are getting used to taking exams online, they may have better (or worse) results than with

traditional exams. When reviewing similar experiences in the bibliography we find two types of studies:

5.1. Studies that found differences

Some studies disagree with our findings and do find significant differences. Alessio *et al* indicate that the results with video-proctored exams were significantly lower (17% obtained an A grade) than the unsupervised ones (63% obtained an A grade). In addition, the unsupervised needed more time to complete the exam, which make them suspect that they were searching for the answers. They conclude by emphasizing the importance of using proctoring to combat academic dishonesty (Alessio *et al*., 2017). According to Goedl & Malla (2020), the average grades of students on unsupervised online exams are statistically higher than on proctored online exams, so the two types of exam are not equivalent.

We found the same results in Harmon & Lambrinos (2008), suggesting that differences were due to cheating. Hylton *et al*., (2016) also indicated that the unsupervised students needed more time to complete the exam, in a similar way to Prince, Fulton, & Garsombke, (2009), Richardson & North, (2013), Truszkowski, (2019), Weiner & Hurtz, (2017), and Wellman & Marcinkiewicz, (2004) that obtained same conclusions. Curiously, no one found better grades in the supervised group than in the unsupervised group.

5.2. Studies that found no differences

However, many other studies find results similar to ours, with no differences between groups. Mozes-Carmel & Gold (2009) conducted a large-scale study with 1,800 students, taking advantage of the fact that Nova Southeastern University decided in 2003 that all exams would be held online. They found no statistically significant difference between the average scores of the supervised and unsupervised exams and attributed this to the type of exam questions. Beck (2014). also found no differences between the two groups and attributed it to techniques to reduce cheating: randomizing the order of the questions, showing one question at a time, and not going backward through the questions. Other studies that do not find differences between groups are Howard, (2020), Ladyshewsky, (2015), Lee, (2020), Stack, (2015), Wibowo *et al*., (2016), Woldeab & Brothen, (2019) and Yates & Beaudrie, (2009).

5.3. What can we do to fight against academic dishonesty?

To consolidate the non-difference of results it is necessary to fight against academic dishonesty. Tips found in literature for minimizing cheating on exams include: do not allow going back on questions, show the questions randomly, show the different answers to the question randomly, reduce exam time, design questions based on learned knowledge and skills rather than identifying one answer among five options. If it is possible, include essay questions. Disciplinary sanctions and their consequences also have a dissuasive effect, as

well as including a clause of good repute at the beginning of the examination (Daffin Jr & Jones, 2018). It is also possible to use open book exams (Alessio *et al*., 2017), use comprehension, analysis, and practical application questions instead of simple memory questions (Howard, 2020). It also helps to take a practice test at the beginning of each course (Anderson & Gades, 2017).

6. Conclusion

We conclude that conducting online exams during the Covid-19 pandemic has not had an impact on the grades of students, and no grade differences were found respect those exams performed in a traditional face-to-face way. These conclusions allow us to encourage teachers and students who fear that online exams will alter the grades to have solid data that leads to reject that worry. There is a lot of variability in similar studies in the literature that can be explained by different cheating methods that can be performed in online exams. A proctoring system, good question design, and limited exam time can minimize these differences.

References

Alessio, H. M., Malay, N., Maurer, K., Bailer, A. J., & Rubin, B. (2017). Examining the effect of proctoring on online test scores. *Online Learning, 21*(1), 146-161.

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*(1), 69-80.

Anderson, C., & Gades, P. (2017). Proctoring exams in an online environment.

Beck, V. (2014). Testing a model to predict online cheating—Much ado about nothing. *Active Learning in Higher Education, 15*(1), 65-75.

Daffin Jr, L. W., & Jones, A. A. (2018). Comparing student performance on proctored and non-proctored exams in online psychology courses. *Online Learning, 22*(1), 131-145.

Goedl, P. A., & Malla, G. B. (2020). A study of grade equivalency between proctored and unproctored exams in distance education. *American Journal of Distance Education,* 1-10.

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

Haus, G., Pasquinelli, Y. B., Scaccia, D., & Scarabottolo, N. (2020). Online written exams during covid-19 crisis. *E-Learning 2020,* pp. 79-86.

Howard, D. (2020). Comparison of exam scores and time taken on exams between proctored on-campus and unproctored online students. *Online Learning, 24*(4)

Hylton, K., Levy, Y., & Dringus, L. P. (2016). Utilizing webcam-based proctoring to deter misconduct in online exams. *Computers & Education, 92*, 53-63.

Ladyshewsky, R. K. (2015). Post-graduate student performance in 'supervised in-class' vs.'unsupervised online'multiple choice tests: Implications for cheating and test security. *Assessment & Evaluation in Higher Education, 40*(7), 883-897.

Lee, J. W. (2020). Impact of proctoring environments on student performance: Online vs offline proctored exams. *The Journal of Asian Finance, Economics and Business (JAFEB), 7*(8), 653-660.

Milone, A. S., Cortese, A. M., Balestrieri, R. L., & Pittenger, A. L. (2017). The impact of proctored online exams on the educational experience. *Currents in Pharmacy Teaching and Learning, 9*(1), 108-114.

Moten Jr, J., Fitterer, A., Brazier, E., Leonard, J., & Brown, A. (2013). Examining online college cyber cheating methods and prevention measures. *Electronic Journal of E-Learning, 11*(2), 139-146.

Mozes-Carmel, A., & Gold, S. S. (2009). A comparison of online vs. proctored final exams in online classes. *Journal of Educational Technology, 6*(1), 76-81.

Prince, D. J., Fulton, R. A., & Garsombke, T. W. (2009). Comparisons of proctored versus non-proctored testing strategies in graduate distance education curriculum. *Journal of College Teaching & Learning (TLC), 6*(7).

Reisenwitz, T. H. (2020). Examining the necessity of proctoring online exams. *Journal of Higher Education Theory and Practice, 20*(1), 118-124.

Richardson, R., & North, M. (2013). Strengthening the trust in online courses: A common sense approach. *Journal of Computing Sciences in Colleges, 28*(5), 266-272.

Stack, S. (2015). The impact of exam environments on student test scores in online courses. *Journal of Criminal Justice Education, 26*(3), 273-282.

Truszkowski, D. (2019). Proctored versus non-proctored testing: A study for online classes. *Dissertation of the Doctoral Program of the American College of Education,*

Weiner, J. A., & Hurtz, G. M. (2017). A comparative study of online remote proctored versus onsite proctored high-stakes exams. *Journal of Applied Testing Technology, 18*(1), 13-20.

Wellman, G. S., & Marcinkiewicz, H. (2004). Online learning and time-on-task: Impact of proctored vs. un-proctored testing. *Journal of Asynchronous Learning Networks, 8*(4), 93-104.

Wibowo, S., Grandhi, S., Chugh, R., & Sawir, E. (2016). A pilot study of an electronic exam system at an australian university. *Journal of Educational Technology Systems, 45*(1), 5-33.

Woldeab, D., & Brothen, T. (2019). 21st century assessment: Online proctoring, test anxiety, and student performance.

Yates, R. W., & Beaudrie, B. (2009). The impact of online assessment on grades in community college distance education mathematics courses. *The Amer.Jrnl.of Distance Education, 23*(2), 62-70.