EDUCATION
RESEARCH SERVICE

AWARD WINNING ORIGINAL ARTICLE

Effect of unproctored versus proctored examinations on student performance and long-term retention of knowledge

Niu Zhang, MD, MS, James Larose, MBBS, MMed, and Megan Franklin, DC

ABSTRACT

Objective: To compare unproctored and proctored online exams among chiropractic students.
Methods: Pre-existing data of 234 students across 4 consecutive endocrinology classes were analyzed for this study. The course was comprised of 3 lectures (50 minutes per lecture) each week. Student performance was evaluated by midterm exam and summative exam (S1). The students from 3 classes were asked to take a voluntary second summative exam (S2) approximately 7 months after the S1. Since this study was partially conducted during the COVID pandemic, some classes took the midterm and the S1 proctored in the classroom while others took them unproctored from a remote location.
Results: The mean midterm exam ($p < .001$) and S1 scores ($p = .01$) for the unproctored group (93.6 ± 7.0 and 88.8 ± 8.2) were significantly higher than the proctored group (88.1 ± 8.2 and 83.9 ± 11.2). The mean time taken by students was much greater for the unproctored exams than for the proctored exams (midterm: 40.7 ± 10.2 versus 16.7 ± 7.0, $p < .001$; S1: 47.0 ± 8.7 versus 21.5 ± 9.0, $p < .001$). By contrast, the mean unproctored S2 scores were lower than the proctored group (60.2 ± 14.7 versus 88.1 ± 8.2, $p < .001$). A linear regression test showed that the final exam was a statistically significant predictor of the recall exam ($p < .01$, $R^2 = 28.3\%$).
Conclusion: The findings suggest that student performance is significantly altered by test format.

Key Indexing Terms: Distance Education; Chiropractic; Education; COVID-19 Pandemic

J Chiropr Educ 2024;38(2):114–119 DOI 10.7899/JCE-23-16

INTRODUCTION

Tests are a typical practice used to gauge student progress. The traditional testing format in higher education is a proctored, closed-book, pencil-and-paper exam.[1] The key argument in favor of this kind of test format appears to be that it reduces the likelihood that it will be tainted by unethical student activity.[1,2] With the emergence of newer technologies, online exams are being used more frequently by instructors in on-campus and online courses.[3,4] The advantages of online testing are increased grading accuracy, shorter grading times, and the ability to provide students rapid feedback.[5]

Tao and colleagues[6] point out that instructors have 2 options for online testing: (a) students can take an online test in a proctored setting, or (b) students can take an unproctored online take-home test setting at a computer of their choice. They indicated that the biggest advantage of computerized take-home tests is the convenience to instructors and students. However, aberrant student behaviors such as cheating still present a substantial challenge for instructors who implement take-home testing. Reports on the importance of proctoring in the administration of online exams among different majors have been mixed. In a course, Introduction to the Computer in Business, Hollister and Berenson[7] found no significant difference in students' academic performance results between proctored and unproctored online exams. Also, they did not find evidence of cheating behavior in the unproctored section. Frein[8] compared test scores between proctored paper-and-pencil in-class tests, proctored online tests, and unproctored online tests in an introductory psychology course and found no significant differences. However, differences between proctored and unproctored tests have been reported by other researchers. Schultz and colleagues[9] compared student performance in proctored paper-and-pencil tests with that of unproctored online exams in marketing, management, and accounting subjects and observed that the unproctored exams resulted in noticeably better test scores. A considerable performance gap between proctored and unproctored tests in a psychology course was also noted by

This paper was selected as a 2023 National Board of Chiropractic Examiners Research Award at the Association of Chiropractic Colleges – Research Agenda Conference.
First Published Online September 17 2024

Carstairs and Myors.[10] On a proctored paper and pencil exam, students' average scores were 5 points lower than on an unproctored exam. Rakes et al[11] tracked the performance of 114 master's and doctoral level students who completed an online unproctored exam or a conventional proctored exam. In contradiction to the findings from Schultz's study and Carstairs' study, they found that participants in the traditional proctored test performed significantly better than those who participated in the online unproctored test.

Unproctored online assessments raise concerns about test scores being exaggerated if students are permitted to consult course materials.[12] Two studies reported that unproctored online tests had more instances of cheating than proctored tests.[13,14] Although multiple studies reported that cheating was common in online courses, few studies really assessed cheating behavior. According to Corrigan-Gibbs et al,[15] several researchers found evidence of considerable cheating in online tests, while others did not.[16,17]

Due to the inconsistent outcomes from earlier studies, further evaluation of proctored and unproctored online test performance is warranted. We hypothesized that: (1) Students who took unproctored exams would have better test grades than those who took proctored exams; (2) Students who took unproctored exams would spend more time on exams than those who took proctored exams; and (3) The unproctored and proctored test formats would affect long-term retention over a period of 10 months after initial learning. To determine whether the proctoring format had an impact on students' academic behaviors, this study contrasted students' performance on tests with and without proctoring, including test-taking time, and long retention of presented class information.

METHODS

Prior to implementation, the study protocol was reviewed and approved by the Palmer College of Chiropractic institutional review board (Assurance #: X2021-001).

Student Participants

The academic year on our campus is organized into 4 terms: fall, winter, spring, and summer, according to a quarter system. Four sessions of the endocrinology course were provided in a single academic year (October 2020 to September 2021) with the same instructor who provided the same lecture structure, and course materials.

This was a retrospective study. Preexisting assessment data were obtained for 234 third-quarter students enrolled across the 4 consecutive presentations of the same endocrinology course. These 4 classes were divided into 3 testing groups, including unproctored, proctored (2 classes), and mixed based on the format of the midterm and final exams (Table 1). Students taking unproctored exams were informed that tests were to be taken without written or verbal references during the test; however, there was no formal way for instructors to verify this. For data analysis, unproctored midterm exam score data were combined across 2 classes (n = 86) for comparison with proctored midterm exam data combined across 2 classes (n = 148). Unproctored final exam score data were obtained from a single class (n = 26) while proctored final exam score data were combined across 3 classes (n = 208).

Table 1 - Exam Format by Testing Group

Testing Group	No. Class	No. Student	Examination Format: Midterm Exam	Examination Format: Final Exam
Unproctored	1	26	Unproctored	Unproctored
Proctored	2	148	Proctored	Proctored
Mixed	1	60	Unproctored	Proctored

The online testing platform, Brightspace (D2L Corporation), was used to administer each exam. The midterm exam and comprehensive final exam were administered in weeks 5 and 11 of each academic term, respectively. Each exam consists of 50 multiple-choice, single-best-answer questions. Students were allowed 50 minutes to complete each exam. The proctored or unproctored exams were chosen based on the condition of whether the exams were taken on campus or remote due to the COVID-19 pandemic. There was only 1 class, termed a mixed class, whose midterm was unproctored and final exam was proctored in the same academic term.

In the sixth quarter, about 7 months after the final exam, all students were requested to take a voluntary exam called the "recall exam." In all classes, the recall exams were the same and unproctored. Without supervision, these students completed their recall exam. The mixed class' students skipped the recall exam. The students' grades in the "recall exam" were not part of their final grades of that academic year.

Data Analysis

Results were analyzed using SPSS version 22 (IBM Corporation). The independent *t* test was used to compare the impact of proctored/unproctored exams on grades for midterm exams and for final exams, as well as for time taken during these exams. In a secondary analysis, the difference between each testing group's midterm and final exam scores and test duration was analyzed by the paired *t* test (for each test format: unproctored, proctored, and mixed). A percentage scale with a minimum score of 0% and a maximum score of 100% was used to standardize all exam results. As a predictor for the recall exam scores, we conducted a multiple linear regression analysis on the final exam score. The examination format as an independent variable was entered into the regression. Multiple linear regression requires at least two independent variables, which can be nominal, ordinal, or interval/ratio level variables. For all quantitative measures, effect sizes were estimated, and 95% CIs were reported. Statistical significance was evaluated at $p = .05$.

RESULTS

Demographic Information

The total number of students included in this study was 234 (97 women, 137 men). The majority of students in all 4 classes were Caucasians (68%), less than 30 years old (91%), with undergraduate degrees (96%). Very few have graduate degrees.

Table 2 - Means and Standard Deviations (SD) for Exam Scores

Exam Type	Midterm Exam Mean (SD)	Final Exam Mean (SD)
Unproctored	93.6 (6.7)	88.8 (8.2)
Proctored	88.1 (8.2)	83.9 (11.2)
95% CI	−7.5 – −3.4	−9.3 – −.4
p Value	< .01	.01

Effect of Test Format on Midterm Exam Score and Final Exam Score

Independent *t* tests showed the midterm exam mean scores were significantly impacted by the exam format, $t\ (_{232}) = 5.25$, being higher in the unproctored group compared to the proctored group ($p < .01$). Similar results were also observed in the final exams in that unproctored students received higher grades compared to those in the proctored group, according to an independent *t* test; $t\ (_{232}) = 2.13$, $p = .01$ (Table 2).

Additionally, a paired *t* test revealed that the mean scores of the final exam in proctored and mixed classrooms were significantly lower than the corresponding midterm examinations ($p < .01$) in both cases. When compared to the midterm exam, the mixed class's final exam mean score fell by 16% (Fig. 1).

Effect of Test Format on Time Spent on Exams

When we compared the time it took students to complete their midterm exams in the unproctored group (n = 86) and the proctored group (n = 148), we observed that the unproctored group's mean time ($M = 40.7$, $SD = 10.2$) was 24 minutes (144%) longer than the proctored group's ($M = 16.7$, $SD = 7.0$); $t\ (_{132}) = 19.4$, $p < .01$. Similar results were obtained when we compared the amount of time students spent on their final exams. Students in the unproctored group ($M = 47.0$, $SD = 8.8$) spent considerably more time on final exams than those in the proctored group ($M = 21.5$, $SD = 9.0$); $t\ (_{37}) = 17.5$, $p < .01$.

A paired-samples *t* test was performed to compare the length of time spent in midterm exam and in final exam in each group (mixed, unproctored, and proctored). In the proctored group, the results indicated that the final exam time ($M = 20.6$, $SD = 8.4$) was significantly longer than the midterm exam time ($M = 16.7$, $SD = 7.0$); $t\ (_{147}) = 8.8$, $p < .01$. By contrast, in the mixed group, the final exam time ($M = 23.7$, $SD = 10.0$) was significantly shorter than the midterm exam time ($M = 38.8$, $SD = 10.4$); $t\ (_{59}) = -9.7$, $p < .01$ (Fig. 2).

Effect of Test Format on Long-Term Retention

Twenty (77%) students from the unproctored class and 76 (49%) students from the proctored classes completed the recall exam.

Independent *t* tests showed a significant difference between the unproctored ($M = 60.2$, $SD = 14.7$) and proctored $M = 88.1$, $SD = 8.2$) groups recall test scores ($t\ (_{94}) = -4.1$, $p < .01$). The final exam was shown to be a statistically significant predictor of the recall exam, accounting for 28% of total recall exam scores ($p < .01$, $R^2 = 28.3\%$, ß = .53, step 1 of the regression model). The addition of the examination format, into the regression model significantly increased predictive power ($p < .01$, $R^2 = 39.8\%$, ß = .68, step 2 of the regression model).

DISCUSSION

It has been debated in the past whether proctored or unproctored tests could affect students' academic performance. According to the results of our study, students who took tests without a proctor had higher test scores, longer exam times, and lower repeated test scores 7 months after the final exam. Comparing the 3 categories to the students who used proctored tests, all 3 showed statistical significance.

The conclusion that students in the unproctored group scored much better than students in the proctored group in terms of grades agrees with other earlier research.[7,18,19] The score difference between unproctored versus proctored exams differed by 5.46% on the midterm and 4.84% on the final exam. These differences were also found within each class. In other words, the average score decreased by 16 points when

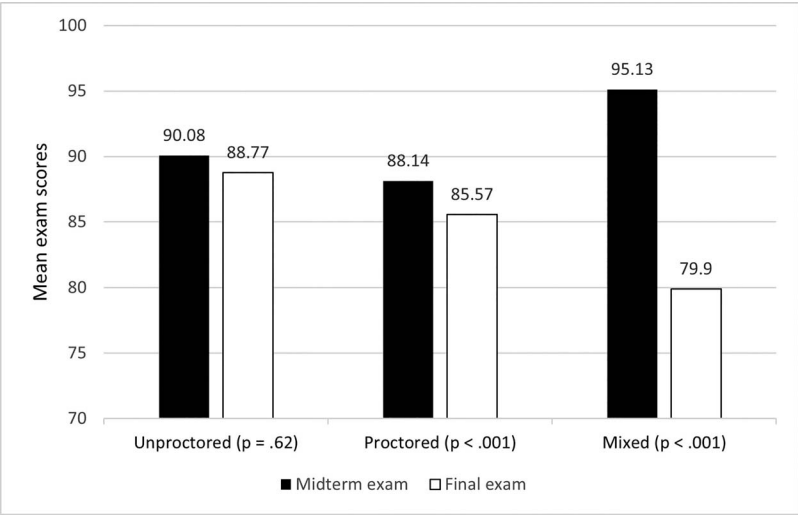

Figure 1 - Comparison of mean midterm and final exam scores by testing group (paired *t* test).

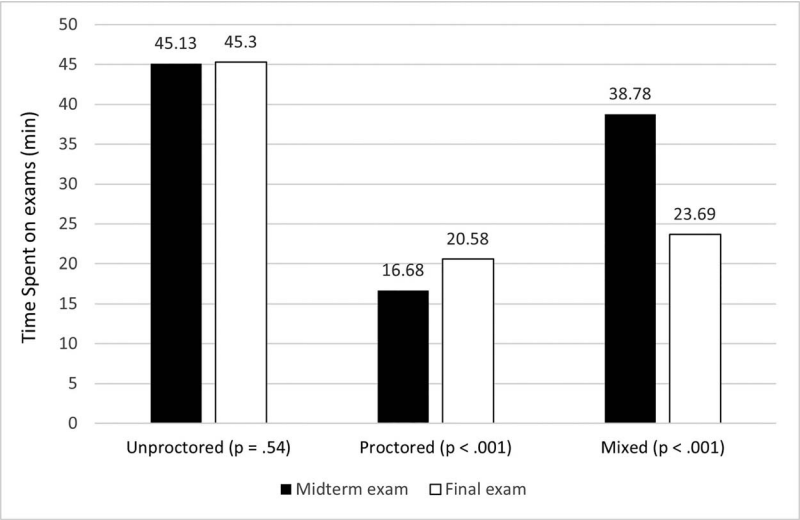

Figure 2 - Comparison of times students spent on exams by testing group (paired *t* test).

the same students switched from an unproctored midterm to a proctored final. Furthermore, students spent significantly more time on unproctored exams compared to proctored ones on average. This supports our hypothesis.

In the current study, the students with unproctored exams were informed that tests were to be taken by themselves without using any resources or discussion with peers. But there were not any monitoring methods, such as proctoring software, used during the exams. On average, the time spent on taking the midterm exam and the final exam was different. In all unproctored exams the average time spent on exams is significantly longer than on proctored exams. It is likely that the better grades and longer exam-taking time are interrelated. Unproctored exams created an opportunity for some students to seek answers from other sources during exams. In other words, they took extra time searching for the answers. Due to such a benefit, the extra time they invested might have improved their grades.

There are several other factors, such as question type, which may explain why students performed better on unproctored exams than on proctored exams. In this study, the question structure was the same for both proctored and unproctored tests. Varble's[20] research indicates that topics on tests fall into 1 of 4 categories: (a) recall, (b) understand, (c) analyze, and (d) apply. Unproctored students may prepare less for recall-type questions because they can check their responses by consulting reference material which also increases the time spent answering questions. According to a prior study, it is easier for students to cheat on these questions than it is for them to understand, analyze, and apply them. All exam questions in the current study were multiple choice, and many of them were recall-type inquiries. The student might benefit much from attempting to look up these questions while taking the test. According to some studies, questions that call for certain abilities to be used, such as math calculations, would not show a significant grade difference between proctored and unproctored tests.[7,21]

According to Hilton et al, students who believed they were being monitored by online proctors felt they could not behave inappropriately throughout the exam, but those who believed they were not being watched collaborated with their peers for some questions.[22] Numerous studies have found that there was a probability for academic dishonesty and a perception that cheating occurred more frequently in online classes.[23–26] Researchers discovered evidence of massive test-taking fraud online.[15] Although this study did not evaluate cheating behavior, given the statistically significant disparities in exam length and scores between the 2 proctoring conditions, we could not completely rule out the potential that some students may have

Of course, the better grades for unproctored students are not always attributed to extra help and longer time, it could also be due to the testing environment. Some students prefer an isolated, quiet, and less distracting environment when taking exams. The distance unproctored exams just provide such an environment that may make students feel less anxious and in turn boost their performance.[27] In a prior study, we discovered that 166 third-quarter chiropractic students (85%) reported having moderate to high test anxiety.[28] Additionally, we noticed a statistically significant inverse relationship between test anxiety and performance on written exams. It is possible that unproctored exams create a less stressful atmosphere and lower the students' degree of test anxiety and, as a result, increase their performance due to a more comfortable exam environment.

The study's final finding was that on the recall exams given 7 months after the final exams, proctored students outscored unproctored students by 20 points, who had performed better on the middle and final exams. This begs the question of whether various students' preparation strategies might affect their ability to remember knowledge over the long run. According to several studies, students who were aware that they would be evaluated via a take-home, unproctored exam tended to study less diligently than they would have if they were evaluated via an in-class, proctored exam.[29–32] It is highly likely that because they were proctored, they had to work harder to prepare, proctored students in our study may have greater long-term information retention for that reason.

Finally, the probability of academic fraud due to an unproctored circumstance should not be discounted. This

conclusion is crucial because it might help to resolve the question of whether proctoring all exams is necessary to ensure academic integrity. The options to preserve academic integrity might be to reduce test duration, use varied question types, use a lockdown browser, use questions on tests that are chosen at random, and prohibit going back.

Study Limitations

There are limitations to this study. First, because it was a retrospective study, there was a disparity in sample sizes between proctored and unproctored groups. Compared to proctored group and mixed group, the sample size for the unproctored group was 82% and 56% smaller, respectively. A sample size that is too small could have an impact on statistics. Second, this was 1 course (endocrinology). This course requires technical terms to be memorized and accurately applied, and where assessment only included multiple choice questions. It is unclear if the effect would be as significant in courses without timed, closed-ended assessments. Finally, the recall exam was not a formal exam. The students might not take it as seriously as formal exams. In addition, the results of the recall exam may not be accurate because it was unproctored. Therefore, the grades could be only used as reference.

CONCLUSION

This study sheds light on the impact of proctoring on students' testing outcome. The findings corroborate our hypothesis that students spend more time and receive higher scores on unproctored exams, which is consistent with prior research. Additionally, it revealed that the unproctored cohort experienced poorer long-term recall. These outcomes were likely explained by differences in exam preparations.

FUNDING AND CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare relevant to this work.

About the Authors

Niu Zhang (corresponding author) is a professor in the Life Science Department at Palmer College of Chiropractic Florida (4777 City Center Parkway, Port Orange, FL, 32129; niu.zhang@palmer.edu). James Larose is a professor in the Life Science Department at Palmer College of Chiropractic Florida (4777 City Center Parkway, Port Orange, FL 32129; Larose_J@palmer.edu). Megan Franklin was an instructor at Palmer College of Chiropractic at the time that this study was conducted. This article was received April 21, 2023; revised August 12, 2023, November 11, 2023, and January 18, 2024; and accepted March 14, 2024.

Author Contributions

Concept development: NZ. Design: NZ. Supervision: NZ. Data collection/processing: NZ, MF, JLR. Analysis/interpretation: NZ. Literature search: NZ. Writing: NZ. Critical review: NZ, JLR.

REFERENCES

1. Williams JB. Creating authentic assessments: a method for the authoring of open book open web examinations. In: Atkinson R, McBeth C, Jonas-Dwyer D, and Phillips R, eds. *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference* (pp. 934–937), Perth, Australia, December 5–8, 2004. https://www.researchgate.net/publication/27472692_Creating_authentic_assessments_A_method_for_the_authoring_of_open_book_open_web_examinations
2. Mallroy J. Adequate testing and evaluation of on-line learners. In: *Proceedings of The Instructional Technology and Education of the Deaf: Supporting Learners*, *K- College*: *An International Symposium*, Rochester, NY, USA, June 25–27, 2001.
3. Khare A, Lam H. Assessing student achievement and progress with online examinations: some pedagogical and technical issues. *Int J E-Learn*. 2008;7:383–402.
4. Rowe NC. Cheating in online student assessment: beyond plagiarism. *Online J Distance Learn Adm.* 2004;7(2).
5. Eplion D, Keefe T. Online exams: strategies to detect cheating and minimize its impact. Paper presented at the *Tenth Annual Instructional Technology Conference*, Murfreesboro, TN; 2005.
6. Tao J, Li X. A case study on computerized take-home testing: benefits and pitfalls. *Int J Technol Learn Teach.* 2012;8(1):33–43.
7. Hollister KK, Berenson ML. Proctored versus unproctored online exams: studying the impact of exam environment on student performance. *Decis Sci J Innov Educ*. 2009;7(1):271–294. doi:10.1111/j.1540-4609.2008.00220.x
8. Frin ST. Comparing in-class and out-of-class computer-based tests to traditional paper-and pencil tests in introductory psychology courses. *Teach Psychol*. 2011; 38(4):282–287. doi:10.1177/0098628311421331
9. Schultz MC, Schultz JT, Gallogly J. The management of testing in distance learning environments. *J Coll Teach Learn*. 2007;4(9):19–26. doi:10.19030/tlc.v4i9.1543
10. Carstairs J, Myors B. Internet testing: a natural experiment reveals test score inflation on a high-stakes, unproctored cognitive test. *Comput Hum Behav Rep*. 2009;25:738–742.
11. Rakes G, Fields V, Cox K. The influence of teachers' technology use on instructional practices. *J Res Technol Educ*. 2006; 38(4):411–426. doi:10.1080/15391523.2006.10782467
12. Rovai AP. Online and traditional assessments: What is the difference? *Internet High Educ*. 2000;3:141–151. doi:10.1016/S1096-7516(01)00028-8
13. Kennedy K, Nowak S, Raghuraman R, Thomas J, Davis SF. Academic dishonesty and distance learning: student and faculty views. *Coll Stud J*. 2000;34(2):309–314.
14. King CG, Guyette RW Jr, Piotrowski C. Online exams and cheating: an empirical analysis of business students' views. *J Educ Online*. 2009;6(1):1–11. doi:10.9743/JEO.2009.1.5
15. Corrigan-Gibbs H, Gupta N, Northcutt C, Cutrell E, Thies W. Deterring cheating in online environments. *ACM Trans Comput Hum Interact*. 2015;22(6):23. doi:10.1145/2810239
16. Ladyshewsky RK. Post-graduate student performance in 'supervised in-class' versus 'unsupervised online' multiple choice tests: implications for cheating and test security. *Assess Eval High Educ*. 2015;40(7):883–897. doi:10.1080/02602938.2014.956683
17. Beck V. Testing a model to predict online cheating: Much ado about nothing. *Act Learn High Educ*. 2014;15(1):65–75. doi:10.1177/1469787413514646

18. Richardson R, North M. Strengthening the trust in online courses: a common sense approach. *J Comput Sci Coll*. 2013; 28(5):266–272.
19. Stack S. The impact of exam environments on student test scores in online courses. *J Crim Justice Educ*. 2015;26(3):273–282. doi:10.1080/10511253.2015.1012173
20. Varble D. Reducing cheating opportunities in online test. *Atl Mark J*. 2014;3(3):Article 9.
21. Howard D. Comparison of exam scores and time taken on exams between proctored on campus and unproctored online students. *Online Learn*. 2020;24(4):204–228. doi:10.24059/olj.v24i4.2148
22. Hylton K, Levy Y, Dringus LP. Utilizing webcam-based proctoring to deter misconduct in online exams. *Comput Educ*. 2016; 92–93:53–63. doi:10.1016/j.compedu.2015.10.002
23. Jones IS, Blankenship D, Hollier G. Am i cheating? An analysis of online students' perceptions of their behaviors and attitudes. *Psychol Res*. 2013;3(5):261–269. doi:10.17265/2159-5542/2013.05.003
24. Moten J, Fitterer A, Brazier E, Leonard J, Brown A. Examining online college cyber cheating methods and prevention measures. *Electron J e-Learn*. 2013;11:139–146.
25. Grijalva T, Nowell C, Kerkvliet J. Academic honesty and online courses. *Coll Stud J*. 2006;40(1):180–185.
26. Raines DA, Ricci P, Brown SL, Eggenberger T, Hindle T, Schiff M. Cheating in online courses: the student definition. *J Eff Teach*. 2011;11(2):80–89.
27. Bengtsson L. Take-home exams in higher education: a systematic review. *Educ Sci*. 2019;9:267.
28. Zhang N, Henderson CN. Test anxiety and academic performance in chiropractic students. *J Chiropr Educ*. 2014;28(1): 2–8. doi:10.7899/JCE-13-20
29. Hall L. Take-home tests: educational fast food for the new Millennium? *J Manag Organ*. 2001;7:50–57. doi:10.1017/S1833367200005216
30. Marsh R. A comparison of take-home versus in-class exams. *J Educ Res*. 1984;78:111–113. doi:10.1080/00220671.1984.10885583
31. Haynie J. Effects of take-home tests and study questions on retention learning in technology education. *J Technol Learn*. 2003;14:6–18. doi:10.21061/jte.v14i2.a.1
32. Moore R, Jensen P. Do open-book exams impede long-term learning in introductory biology courses? *J Coll Sci Teach*. 2007;36:46–49. https://www.jstor.org/stable/42992504