Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company’s public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Research Note

Does a sudden shift of testing format from closed-book to open-book change the characteristics of test scores on summative final exams?

M. Ken Cor a, Dion R. Brocks b,∗

a Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta, Edmonton, AB, Canada
b Faculty of Pharmacy and Pharmaceutical Sciences, 2-142 H Katz Group Centre for Pharmacy and Health Research, University of Alberta, Edmonton, AB T6G 2E1, Canada

ARTICLE INFO

Introduction: In response to the COVID-19 pandemic, most universities in North America transitioned to online instruction and assessment in March 2020. Undergraduate pharmacy students in years one to three of two four-year entry-to-practice programs at a university in Canada were administered open-book examinations to complete their didactic winter-term courses in pharmaceutical sciences; behavioural, social, and administrative sciences; and pharmacotherapeutics. The impacts of the switch to open-book examinations on final exam characteristics are examined.

Methods: The ratios and correlations of final exam and midterm grades in 2020, where finals were open-book, and in 2019, where finals were closed-book, were calculated and compared.

Results: In 2020, the ratio of final exam to midterm exam scores for five out of seven courses were significantly larger than they were in 2019. Alternatively, for all but one course, the correlations between midterm and final examination grades showed no significant difference from 2019 to 2020.

Conclusions: Compared to 2019 when finals were administered in a closed-book format, a sudden shift to an open-book format for final exams in 2020 appears to be associated with the final exams becoming easier relative to midterms. However, when considering how final and midterm exam grades correlate year over year, in all but one class, there was no significant difference. These findings suggest that changing exams to be open-book may change how they can be used to inform criterion-referenced or absolute grading decisions but not norm-referenced or rank-based decisions.

© 2021 Elsevier Inc. All rights reserved.

Keywords:
Open-book
Closed-book
Examinations
Performance
COVID-19

Abstract

Introduction: In response to the COVID-19 pandemic, most universities in North America transitioned to online instruction and assessment in March 2020. Undergraduate pharmacy students in years one to three of two four-year entry-to-practice programs at a university in Canada were administered open-book examinations to complete their didactic winter-term courses in pharmaceutical sciences; behavioural, social, and administrative sciences; and pharmacotherapeutics. The impacts of the switch to open-book examinations on final exam characteristics are examined.

Methods: The ratios and correlations of final exam and midterm grades in 2020, where finals were open-book, and in 2019, where finals were closed-book, were calculated and compared.

Results: In 2020, the ratio of final exam to midterm exam scores for five out of seven courses were significantly larger than they were in 2019. Alternatively, for all but one course, the correlations between midterm and final examination grades showed no significant difference from 2019 to 2020.

Conclusions: Compared to 2019 when finals were administered in a closed-book format, a sudden shift to an open-book format for final exams in 2020 appears to be associated with the final exams becoming easier relative to midterms. However, when considering how final and midterm exam grades correlate year over year, in all but one class, there was no significant difference. These findings suggest that changing exams to be open-book may change how they can be used to inform criterion-referenced or absolute grading decisions but not norm-referenced or rank-based decisions.

© 2021 Elsevier Inc. All rights reserved.
normal letter grading system would be abandoned for the term if the final examinations had not yet been administered.

This decision was made in large part not only because of the stay-at-home orders, but also because with such short notice (the switch happened with just one and one-half weeks left in the term), it was virtually impossible to ensure that examinations could be appropriately proctored and held as closed-book examinations. As a result, the faculty of pharmacy and pharmaceutical sciences decided to administer open-book examinations.

These modifications coupled with the direction from the Faculty for instructors to run their exams without making changes to their original assessment plans created a natural quasi-experimental opportunity to investigate how a change in exam format impacted the psychometric properties of final exams. In this study we compare final exam score characteristics including mean score relative to closed-book midterms, correlations with midterms, and reliability in both 2019 and 2020 - when final exams were converted to open-book delivery.

Research on open- vs. closed-book exams

Two recent systematic reviews published in nursing and medical education journals summarize the available evidence comparing open- vs. closed-book exams on a host of outcomes. In regard to exam preparation, current evidence is mostly equivocal with many studies reporting that students spend a similar amount of time preparing for open- vs. closed-book exams with only a small number of studies showing that students spend less time preparing for open-book exams. When it comes to test anxiety and stress, studies consistently reveal a student preference for open-book exams with many finding students' experiencing less stress in an open-book format. Open-book examinations are likely intrinsically less stressful for many students because they focus less on the need to memorize facts.

While evidence about student performance is available, several factors may confound comparisons of open- and closed-book format, such as the type of questions that are contained in the exams. In these studies, open-book exams may be designed differently than closed-book exams to explicitly test application of knowledge with limited to no testing of memorization or basic comprehension. As a result, findings consistently reveal little difference in overall student performance across the formats.

Some studies have explored how the psychometric properties of exams compare across the two formats when questions are kept essentially the same. One of the most rigorous studies showed that when 196 university students in an anatomy and physiology course took a midterm composed of higher-level application and synthesis type multiple choice questions in a closed-book format, and then retook the same test a week later in an open-book format, their mean scores were statistically unchanged. The authors concluded that closed- and open-book formats were equivalent in their ability to measure students when assessing higher-order thinking.

Proponents of open-book examinations argue that they promote "deep" or "higher-order" learning. For example, in one recent paper, it was found that the use of open-book examinations that assessed higher-order thinking yielded advantages in terms of learning outcomes and in impromptu recall of information. Similarly, a larger scale study examining the impact of open-book examinations used with problem-based learning involving over 1600 students in more than 1 year of a program found signs of better long-term retention compared to traditional curriculum with closed-book exams. These and other authors have proposed that open-book examinations have potential utility when properly designed and tailored for certain types of materials. However, others have produced conflicting results as closed-book exams were found to promote more deep-learning strategies than open-book or were found to promote equal levels of deep-learning approaches. In summary, findings about the ways in which converting an exam to an open-book format may change the statistical properties of an exam or affect learning outcomes appear to be mixed.

The literature review reveals identifiable gaps that our study is well suited to address. First, we have the opportunity to look at the question across multiple courses at the same time, all of which have a built-in comparator from the previous year where the same exams were administered as closed-book. This is a study design that takes advantage of a natural experiment and has not been used before to investigate differences in closed- and open-book exams. Finally, research on the difference in closed- and open-book exams have almost exclusively focused on comparing overall mean performance. Our research considers impacts of changing to an open-book format on correlations to midterms and exam reliability. These outcomes provide another important source of evidence to evaluate how closed- vs. open-book formats impact exam properties.

Methods

An exploratory ex-post-facto observational design was used to investigate how a change to open-book exam format impacted final exam score characteristics. Data from 2019, where both midterm and final exams were administered in an in-person, closed-book format, were analyzed with data from 2020 where the final exams were administered with a remote, open-book format. The use of this data for the purposes of this study was approved by the University of Human Ethics Research Office.

Current program description and courses being analyzed

The Faculty is currently in the second year of implementation of a new curriculum and phasing out of the previous bachelor of pharmacy (BSc Pharm) curriculum. As a result, there were 2 years of comparable data from two consecutive cohorts of first-year students who had taken the same new doctor of pharmacy (PharmD) courses and two consecutive cohorts of third-year students who had taken the same BScPharm courses. Data for the current cohort of second-year PharmD students was not considered because these courses had only been delivered once and there was no basis for comparison of whether moving to open-book for final exams has resulted in any differences. While the new PharmD program has an additional focus on pharmacy practice skills and an additional
16 weeks of experiential learning, both programs employed the same approach to midterm and final assessments for the didactic, non-experiential courses in the program. This meant that we could compare exam characteristics of: (1) ratios of final exam to midterm exams grades for seven courses, four from Year 1 of the new PharmD program and three from Year 3 of the BScPharm courses; (2) correlations between midterms and final exam scores for the same seven courses in 2019 and 2020; and (3) Cronbach alpha reliability for final exams in these same seven courses in 2019 and 2020.

**Exam comparability across years**

In an attempt to reduce the amount of stress experienced by students and to ensure students were not thrown off by a change in the type of questions they were tested on in previous examinations, instructors were requested to proceed with their exams as they would have if they were administered in a closed-book format. In particular, instructors were asked to refrain from changing the format of questions to add more application-based or short-answer questions that are typically employed with open-book exams. As a result, instructors used their regular practice of copying the exam from the previous year and making revisions so that at least 80% of the exam was exactly the same when compared to the previous year's exam. In this way, the exams are assumed to be minimally different year over year so that changing to an open-book format is not confounded with changes in question type or the content being assessed.

**Description of data**

Grades from seven of the winter terms 2019 and 2020 midterm and final examination grades were compiled for students enrolled in courses from Years 1 and 3 of the PharmD and BScPharm programs, respectively. The grades were all converted to percentages to allow for comparability.

In both the closed- and open-book format versions of the exams in 2019 and 2020, exams were timed and constructed of primarily multiple-choice questions with a few short-answer questions. In addition, in both years and for both the open- and closed-book formats, a computer-based examination platform developed by the faculty of medicine and dentistry at the University of Alberta was used to administer all midterm and final exams. The 2019 exams and all winter term midterm exams were proctored in-person. All other winter term 2020 final examinations were conducted unproctored and open-book. In these exams, students were permitted to use any resource they wanted including their course notes posted on the course webpages or other resources on the web.

The admit class size for each course and year of the program is 131 students and no course included students who were repeating the course. Students were informed that distance examinations were expected to be completed individually and that any breach of this would be subject to academic discipline. All suspected or known cases of this occurrence led us to remove those student scores from the analysis.

Not all courses had data that was suitable for the analysis. For example, the third-year curriculum of the older program had courses that lasted less than a full academic winter term (January to end of April), and some of those courses were complete before the emergency situation affected courses. In another third-year course, both the midterm and the final examination were affected by the changes imposed by the emergency and, as a result, were not included. Finally, the pharmacy practice skills courses are administered with an open-book format regularly so there was no change in format to be investigated.

**Statistical analysis**

Within each course, the mean and SD of the examination scores were estimated. We also calculated the average ratios of final to midterm exam scores and the correlations between midterm and final exam scores. The total number of statistical tests needed to conduct comparisons across the 2 years for the seven courses and the two statistics was 28. Therefore, a simple Bonferroni adjusted P value < 0.001785 was used to identify statistical significance. The 99% CI were used rather than 95% CI to compare the ratios and correlations to reduce the chance of a type I error. Cronbach's alpha internal consistency reliabilities were also calculated for the final exams for descriptive purposes but no formal statistical tests were run with these.

**Results**

A total of seven Year 1 and Year 3 pharmacy courses were evaluated. Course enrolment ranged from 126 to 138 students. Table 1 shows the descriptive statistics of interest for the midterm and final exams of each of the seven courses considered for the study. In 2019, five of seven courses had a lower average score on the final exam compared to the midterm. All courses showed a significant positive correlation between the midterm and final exam that ranged from 0.39 to 0.66, indicating low to moderate levels of association. Cronbach's alpha for final exams ranged from 0.57 to 0.88 indicating moderate to high levels of reliability.

In the 2020 courses, where final exams were converted to open book, the opposite trend was seen with mean scores. For all courses, the average final exam scores were estimated as higher compared to the average midterm exam scores. Table 1 shows the descriptive statistics of interest for the midterm and final exams of each of the seven courses considered for the study. Table 1 shows the descriptive statistics of interest for the midterm and final exams of each of the seven courses considered for the study. In 2019, five of seven courses had a lower average score on the final exam compared to the midterm. All courses showed a significant positive correlation between the midterm and final exam that ranged from 0.39 to 0.66, indicating low to moderate levels of association. Cronbach's alpha for final exams ranged from 0.57 to 0.88 indicating moderate to high levels of reliability.

In order to test if final exam scores relative to their associated midterm scores in 2020 were different than they were in 2019, we calculated the average ratios of the final to midterm grades for each course in each year and calculated their associated CI (Fig. 1).
The lack of overlap in the 99% CI indicates a statistically significant difference. The results reveal that for five of the seven courses, the ratios of final exam to midterm grades were significantly higher in 2020 compared to 2019.

The same type of analysis was done for the correlations between midterm and final exams in 2019 and 2020 (Fig. 2). The 99% CI were seen to overlap for all courses but one, suggesting that they behaved similarly in how they related to other variables year over year.

Discussion

For seven courses in our entry to practice program, exams were designed in the same way by the same primary instructors in 2020 as compared to 2019. The only major difference in exam administration was the open-book nature of final exams in 2020. The ratios of final exam to midterm scores revealed that students performed better relative to midterms in 2020 than in 2019. These findings support the logical expectation that, all things being equal, the open-book format makes exams easier for students to find the correct answers to some of the questions. This would seem intuitive given that instructors in our program did not change their assessment questions markedly given the immediacy and change in course instruction and assessments posed by the public health directives. Nevertheless, in the absence of a health care emergency, others who have looked at closed- vs. open-book examination formats of the same types of questions and materials have also reported higher examination scores with open-book formats. An implication of this finding is that, when used for absolute grading decisions like pass/fail course grading, where a pre-set pass grade is stipulated, open-book exams that are unchanged in other ways, increase the chances of students passing a course.

Alternatively, our findings reveal that correlations between midterms and finals were highly similar year over year (Fig. 2) in all but one pharmaceutical science course. These results suggest that, from a statistical perspective, open-book final exam scores in 2020 related similarly to midterms as they did in 2019. Combining this with the observation that the reliabilities of the exams have stayed strong, an implication of this result could be that final exams in 2020 have a similar ability to rank order students as they did in

Table 1
Midterm and final exam mean percent scores, correlations, and final exam reliabilities across 2019 and 2020.

| Topic | 2019 | 2020 |
|-------|------|------|
|       | Midterm (mean ± SD) | Final (mean ± SD) | Correlation, (r) | Final exam reliability, (a) | Midterm (mean ± SD) | Final (mean ± SD) | Correlation, (r) | Final exam reliability, (a) |
| BASE-y1 | 79 ± 8.5 | 69.6 ± 7.8 | 0.39 | 0.68 | 74.1 ± 8.7 | 77.8 ± 7.3 | 0.39 | 0.72 |
| BASE-y3 | 73.7 ± 6.7 | 75.2 ± 6.8 | 0.38 | 0.57 | 75.4 ± 7.8 | 79.7 ± 7.7 | 0.36 | 0.67 |
| PS1-y1 | 83.6 ± 10.1 | 71.9 ± 11.0 | 0.66 | 0.87 | 87.6 ± 9.3 | 88.3 ± 5.8 | 0.17 | 0.79 |
| PS2-y1 | 78.4 ± 13 | 74.3 ± 15.7 | 0.65 | 0.88 | 86.3 ± 9.9 | 86.5 ± 7.1 | 0.39 | 0.91 |
| PS3-y1 | 66.6 ± 9.3 | 62.5 ± 10.4 | 0.55 | 0.82 | 66.4 ± 10.9 | 77.9 ± 10.5 | 0.31 | 0.86 |
| PT1-y3 | 80.1 ± 7.4 | 81.7 ± 7.8 | 0.61 | 0.65 | 80.6 ± 8.6 | 88.6 ± 6.5 | 0.42 | 0.71 |
| PT2-y3 | 85.6 ± 6.3 | 80 ± 7.9 | 0.55 | 0.76 | 81.3 ± 6.9 | 87 ± 6.4 | 0.43 | 0.72 |

BASE = behavioural and social administrative sciences; PS = pharmaceutical and basic sciences; PT = pharmacotherapy; yr = year.
a These exams were administered open book.

The lack of overlap in the 99% CI indicates a statistically significant difference. The results reveal that for five of the seven courses, the ratios of final exam to midterm grades were significantly higher in 2020 compared to 2019.

The same type of analysis was done for the correlations between midterm and final exams in 2019 and 2020 (Fig. 2). The 99% CI were seen to overlap for all courses but one, suggesting that they behaved similarly in how they related to other variables year over year.

Discussion

For seven courses in our entry to practice program, exams were designed in the same way by the same primary instructors in 2020 as compared to 2019. The only major difference in exam administration was the open-book nature of final exams in 2020. The ratios of final exam to midterm scores revealed that students performed better relative to midterms in 2020 than in 2019. These findings support the logical expectation that, all things being equal, the open-book format makes exams easier for students to find the correct answers to some of the questions. This would seem intuitive given that instructors in our program did not change their assessment questions markedly given the immediacy and change in course instruction and assessments posed by the public health directives. Nevertheless, in the absence of a health care emergency, others who have looked at closed- vs. open-book examination formats of the same types of questions and materials have also reported higher examination scores with open-book formats. An implication of this finding is that, when used for absolute grading decisions like pass/fail course grading, where a pre-set pass grade is stipulated, open-book exams that are unchanged in other ways, increase the chances of students passing a course.

Alternatively, our findings reveal that correlations between midterms and finals were highly similar year over year (Fig. 2) in all but one pharmaceutical science course. These results suggest that, from a statistical perspective, open-book final exam scores in 2020 related similarly to midterms as they did in 2019. Combining this with the observation that the reliabilities of the exams have stayed strong, an implication of this result could be that final exams in 2020 have a similar ability to rank order students as they did in

Fig. 1. Average ratios and associated CI of the final to midterm grades for each course in 2019 and 2020.
Extending this to decision making, final exam grades might be able to be used to make similar types of norm-referenced decisions as they were in 2019. In fact, in our program, letter grades are assigned based on the distribution of performance. The distribution of grades assigned to students in these courses at the end of 2020 was largely unchanged compared to 2019 (i.e. there were the same amount of As, Bs, Cs in 2020 as there were in 2019). Hence, despite being easier due to the open-book nature of the exams, the evidence suggests that the grades from exams that are essentially unchanged in design with the exception of being open-book could be used similarly to rank students as if they were administered in a closed-book format.

In one class the correlation dropped significantly between the midterm and final between 2019 and 2020. That class also experienced the highest increase in average grade in the final examination between 2019 and 2020. One possible explanation is that moving to open-book using the same basic examination changed the nature of the exam so much that the final scores no longer measured the same construct as was measured in 2019. The examinations used in that class were particularly focused on memorization of facts rather than higher-level application or synthesis of information. This would naturally lead to a flattening of the slope between midterm and final examinations and a reduction in the ability to identify those students with higher levels of knowledge going into the examination from those with lower knowledge levels, who could look up the answers in the final examination.

Limitations

This study is observational and, as a result, may have been influenced by confounding variables. The design used midterms as a common point of reference for evaluating how performance on finals differed from 2019 to 2020 and can be considered a strength. However, the differences or lack thereof that we report could be the result of some uncontrolled variables. For example, while we have applied the assumption that our exams are essentially equivalent in their design from 2019 to 2020, we could not strictly control for this because course instructors have the ability and right to determine the nature of their examination questions. In our program while students are able to see their examinations after they are held, they are secure in the sense that they are retained by the instructor. Thus, instructors reported that they change only a small number of questions year over year which means some of our results could be attributable to the likely small differences in exam construction. Similarly, we were in a pandemic situation when students took these exams. Fortunately, over 90% of the term had been delivered before courses were switched to remote delivery and so impacts of how material was taught, in an overall sense, is assumed to be minimal. And while all courses used for this analysis were taught by the same instructors in 2019 as in 2020, year over year differences in how courses were taught cannot be explicitly controlled for due to practicalities.

The generalizability of the results of this study should transfer to contexts with similar types of programs and students across North America. The programs in this study are based on the national accreditation standards which are similar in most North American jurisdictions. For example, the admissions requirements include 2 years of undergraduate prerequisites in chemistry, biological sciences, mathematics, English, physiology, and other courses that are common to most programs in North America.

Conclusions

The shift to open-book unproctored examinations that were designed similarly to the way they were in 2019 at the end of the 2020 winter term had the general effects of making the exams easier but leaving the correlation between the midterm and final scores mostly unchanged. Cronbach's alpha on the final exams were also similar year over year. Taken together, the findings suggest that exams likely
become easier but retain their relative meaning when simply changing them to an open-book format. These results support the conversion from closed- to open-book exams in contexts where grades are used for norm-referenced decisions but not criterion-reference absolute decisions. The data suggested that open-book exams can generally retain their ability to identify students with higher levels of understanding of the material, but this could be diminished if the examinations are overly focused on rote recall of information. Question design may be an important focus for optimizing the use of open-book examinations.

Author's statement

Most papers that have examined the relative outcomes of performance between open and closed book examinations have focused on a single class or type of class, mostly in medical schools. This paper describes how a shift from closed book to open format examinations changes the outcomes of grades in not just one, but a series of classes. Its novelty is mostly rooted in the breadth of the shift because it examines a program-wide breadth of outcomes necessitated by Covid-19. We cannot find any other papers that have examined the assessment measures across an entire program of courses.

Disclosure

None.

Declaration of competing interest

None.

References

1. Durning SJ, Dong T, Ratcliffe T, et al. Comparing open-book and closed-book examinations: a systematic review. Acad Med 2016;91(4):583-99. https://doi.org/10.1097/ ACM.0000000000000977.
2. Johanns B, Dinkens A, Moore J. A systematic review comparing open-book and closed-book examinations: evaluating effects on development of critical thinking skills. Nurse Educ Pract 2017;27:89-94. https://doi.org/10.1016/j.nepr.2017.08.018.
3. Akresh-Gonzales J. Open-book vs. closed-book exams. NEJM Knowledge+ 25 June 2015. Accessed 10 June 2021: https://knowledgeplus.nejm.org/blog/open-book-vs-closed-book-exams/.
4. Brightwell R, Daniel J, Stewart A. Evaluation: is an open-book examination easier? Biosci Educ 2004;3(1):1-10. https://doi.org/10.3108/beej.2004.03000004.
5. Rimmer R, Schwebel J, Schwede A. Open-book versus closed-book tests in university classes: a field experiment. Front Psychol 2019;10:463. https://doi.org/10.3389/fpsyg.2019.00463.
6. Heijne-Penninga M, Kux JBM, Hofman WHA, Muñíjtens AMM, Cohen-Schotanus J. Influence of PBL with open-book tests on knowledge retention measured with progress tests. Adv Health Sci Educ Theory Pract 2013;18(3):465-95. https://doi.org/10.1007/s10459-012-9386-8.
7. Krasne S, Wimmers PF, Relin A, Drake TA. Differential effects of two types of formative assessment in predicting performance of first-year medical students. Adv Health Sci Educ Theory Pract 2006;11(2):155-71. https://doi.org/10.1007/s10459-005-5290-9.
8. Heijne-Penninga M, Kux JBM, Hofman WHA, Cohen-Schotanus J. Influence of open- and closed-book tests on medical students’ learning approaches. Med Educ 2008;42(10):967-74. https://doi.org/10.1111/j.1365-2923.2008.03125.x.
9. Ramamurthy S, Er HM, Nadanrajah VD, Pook PCK. Study on the impact of open and closed book formative examinations on pharmacy students’ performance, perception, and learning approach. Curr Pharm Teach Learn 2016;8(3):364-74. https://doi.org/10.1016/j.jptl.2016.02.017.
10. Jeyaratnam S. Confidence intervals for the correlation coefficient. Stat Probab Lett 1992;15(5):389-93. https://doi.org/10.1016/0167-7152(92)90172-2.
11. de Vet HCW, Mokkink LB, Mosmuller D, Terwee CB. Spearman-Brown prophecy formula and Cronbach’s alpha: different faces of reliability and opportunities for new applications. J Clin Epidemiol 2017;85:45-9. https://doi.org/10.1016/j.jclinepi.2017.01.013.
12. Schumacher CF, Butzin DW, Finberg L, Burg FD. The effect of open-vs. closed-book testing on performance on a multiple-choice examination in pediatrics. Pediatrics 1978;61(2):256-61.