RESEARCH

The Impact of Eliminating Backward Navigation on Computerized Exam Scores and Completion Time

Gary L. Cochran, PharmD, Jennifer A. Foster, PharmD, Donald G. Klepser, PhD, Paul P. Dobesh, PharmD, Allison M. Dering-Anderson, PharmD

University of Nebraska Medical Center, Omaha, Nebraska

Corresponding Author: Gary L. Cochran, University of Nebraska Medical Center, 986145 Nebraska Medical Center Omaha, NE 68198-6145. Tel: 712-310-3992. Email: glcochran@unmc.edu

Submitted February 19, 2019; accepted August 19, 2020; ePublished August 2020

Objective. Computer based testing has become a standard at many colleges of pharmacy. Restricting an exam takers ability to revisit questions previously answered (elimination of backward navigation), is an option available in some testing software. Our objective was to determine whether elimination of backward navigation resulted in changes in examination score or time to complete the exam.
Methods. Six exams administered in courses that eliminated backward navigation were compared to exams given the previous year where backwards navigation was allowed. The primary comparison of interest was change in a subset of identical questions included on both exams. Secondary outcomes included change in total exam score and completion time. The exam data used in this research were obtained using ExamSoft testing software.
Results. No statistically significant reductions in exam scores were observed. Average time spent on a question was significantly reduced on 2 of the 6 exams.
Conclusion. No adverse effect on scores or testing time was found across 3 years of the didactic pharmacy curriculum.
Keywords: computer-based testing, computer-based assessment, backward navigation, exam security

INTRODUCTION

Computer-based testing has quickly replaced paper based testing to become the standard in higher education. [1-3] Computer based testing provides a number of options that were not possible with paper exams including: question and answer randomization, missing answer reminders, and elimination of backward navigation. Elimination of backward navigation restricts an exam takers ability to return to questions they have already answered. Students are not permitted to skip a question and return. Additional computer coding of the examination, when activated, will not allow students to leave a question without a response.

Potential benefits to restricting backward navigation include improved exam security and the ability to include a series of questions where the previous question may provide the answer to a subsequent question. Several articles address methods to improve security with computer-based testing, most of which focus on access to the examination material and how the exam is administered. [2-5] Literature that specifically comments on removing the ability to revisit examination questions indicates that removing backwards navigation on an exam is generally unpopular with examinees, but there is limited information regarding the impact of this decision on test efficiency and performance at the graduate or professional school level. [6-9] The added imperative of new on-line and virtual education, created by the 2019 SARS-CoV-2 pandemic, has increased concern about the computerized testing process. Limited information on the outcome of these decisions can frustrate the decision-making process.

A direct student benefit to prohibiting backward navigation is providing an opportunity for students to gain experience with no reverse navigation, as this is a restriction placed on both the North American Pharmacists Licensure Examination (NAPLEX) and the Multistate Pharmacy Jurisprudence Examination (MPJE) which are required for pharmacist licensure in the United States. Thus far, studies have indicated that removing backwards navigation on examinations may impact performance at the elementary and middle-school levels but has not shown significant impact on test performance at the high school and college levels. [10,11] The purpose of this evaluation was to determine whether the elimination of backward navigation led to changes in examination score or time to complete the exam in a four-year, Doctor of Pharmacy program when compared to the same exam from the previous year where backwards navigation was allowed. At the time the decision was made to restrict backward navigation, these two, potential negative impacts were identified by both faculty and students as undesirable. Due to the lack of sufficient literature to determine if these potential

negative impacts occur, this evaluation was developed to monitor both change in score on identical questions and the time to complete the examination.

METHODS

Backward navigation was available for exams in all courses prior to the 2019 academic year. Prior to the beginning of the fall 2019 semester, the decision was made to limit backward navigation for exams in courses originating in the Department of Pharmacy Practice and Science. Six examinations administered early in the semester (two exams in each of the P1, P2, and P3 years) were identified in courses where backward navigation had been eliminated. The pedagogy for the P1 courses were mainly lecture with active learning methods mixed throughout. The P2 and P3 courses used team-based learning with a mix of active learning through case presentations and other simulated patient interaction exercises. Each exam given in the fall of 2019 was compared to a similar exam given in the fall of 2018. Exam content, topic presentation order, and lecturer were largely the same between years. While new lecturers were rare in the courses studied, in all instances, these instructors created new exam questions. Therefore, none of these questions would have been included in our primary analysis of identical questions from year to year.

The primary comparison of interest was the change in mean score for a subset of identical questions given on both 2018 and 2019 exams. Questions were required to have the same question identification number and revision number on the 2018 and 2019 exams to be included in the primary analysis. The examination software used assigns a new revision number for any change in the question or in the answers, including spelling corrections or distractor/foil replacement. Questions with identical stems, but different revision numbers were excluded from evaluation. Mean score for the total exam (new and previously used questions) was also compared. Only multiple-choice and True/False questions were included in the analysis. Essay questions were rarely used on the six exams and excluded from the analysis.

One exam given in the P3 year was reorganized in 2019. The infectious disease section of the Pharmacotherapy course was divided into two exams for the 2018 semester. The same content was divided into three exams for the 2019 year. Because the 2019 exam questions were included in the first two exams given during the 2018 year, a direct comparison of overall examination time and total examination score is not possible. A comparison of change between the 2018 and 2019 exams for identical questions (primary outcome) was performed.

Three student demographic variables were collected: student age at the beginning of their P1 year, student sex, and student grade point average (GPA) at the end of the P1 year.[12,13] These variables were included in a linear regression model to control for differences in student characteristics. P1 GPA represents the cumulative grade point average of all courses completed during the first year of the four-year program. P1 GPA is only available for comparisons between second and third year exams.

Average time students spent answering each question was compared for each exam by dividing the total student exam completion time by the total number of questions. Median, 90%, and maximum completion times were also reported.

All exam data was extracted from ExamSoft. The Student's t-Test was used to compare age, P1 GPA, exam scores and average time spent on exam questions. The Chi-squared test was used to evaluate differences in sex. A p-value of $< .05$ was considered statistically significant. All statistical comparisons were performed using Stata 14.2. The UNMC Institutional Review Board determined the project does not constitute human subject research.

RESULTS

Student age, sex and P1 GPA are reported in Table 1. Demographics were similar between 2018 and 2019 students for each of the P1-P3 years. The primary outcome of interest was change in score for a subset of identical questions included on both the 2018 and 2019 exams. Change in mean score ranged from + 3.3% (+ 0.6 questions) for the P1 Pharmacy and Health care exam to – 2.6% (- 0.5 questions) for the Pharmaceutical Care exam (Table 2). Four of the six mean exam scores were slightly lower after eliminating backward navigation, however none of the changes were statistically significant. Since there was only one identical True/False question on all six exams combined, no analysis could be done based on question type.

Changes to total exam scores were also evaluated. The total exam score for the second nutrition exam given in the P2 year was 4% higher after elimination of backward navigation (78% vs. 82%; p=.02). No significant differences between total exam scores were observed for any other exam (Table 3).

Change in time to complete the examinations was evaluated using four metrics. First, mean time spent on each question was calculated to provide a normalized measure, so exams with different numbers of questions could be compared. Mean time students spent on each question was significantly reduced for the no backward navigation group on two of the five exams compared (Table 4). Median, 90% student completion, and maximum exam time was evaluated to determine if slower exam takers had sufficient time to complete the exam.

Age, gender, and student performance were postulated as potentially confounding variables by students and faculty.[12, 13] P1 GPA was used as a proxy for student performance. Even though no differences in demographic characteristics exist (Table 1), linear regression models were conducted as planned, to control for potential confounding. Regression models were only created for the primary outcome, change in score for identical questions (Table 5). After controlling for student characteristics, changes between the 2018 and 2019 exams either remained the same or moved closer to the null when compared to the unadjusted analysis. No changes in adjusted scores were statistically significant.

DISCUSSION

This evaluation did not find a negative impact on pharmacy student exam performance after eliminating backwards navigation. Across courses in the first three years of the professional program, average student scores on a set of identical questions did not decrease as a result of eliminating backwards navigation. Similarly, the average time spent per question did not increase as a result of eliminating backwards navigation. These findings are important because they represent the two primary unwanted and unintentional consequences of restricting students' ability to revisit test questions.

Students had two hours to complete the P1 exams evaluated in this study. The four remaining exams were three hours in length. We did not see any evidence that restricting backward navigation resulted in longer exam completion times. There are no institutional guidelines defining exam length. Exam length is determined by each course or section coordinator. Assurance that time to complete the exams would not be increased was important to both students and faculty.

When faced with the elimination of backwards navigation, it is reasonable to expect student anxiety and concern, particularly in a rigorous discipline such as pharmacy, where a slight decrease in exam performance can halt a students' academic progress. At many colleges and schools of pharmacy, courses are only offered once per year, which means that a student failure may result in a "lost" year. Even for students at no risk of failing a course, a difference of a few points may result in a lower course grade, which can affect scholarships and residency placements. With such serious stakes, it is easy to understand student anxiety about any process or procedural change that could negatively impact performance. This evaluation was intended to determine if a negative impact was seen following the elimination of reverse or backwards navigation. In part to quell student anxiety about this change and in part to offer reassurance to faculty who shared similar concerns.

These changes are a particular concern when they run counter to how students have been trained on how to test. Many students today have been taught to skim through the entire exam before starting to answer questions or to work through an exam answering the questions they know immediately, while skipping and returning to questions that require more time.[14] This approach is intended to give students confidence and help them manage their time by not lingering on difficult questions to the point that they have to rush through other questions. Since the goal of examinations is to serve as a method for measuring knowledge, not as a method to evaluate testing strategies, it was important that scores or testing duration not reflect negative outcomes when compared to those students who used backwards navigation. A similar study by Caetano and Pawasauskas likewise showed no impact on item or question performance overall but did not examine student exam performance as we did.[11]

The benefits of eliminating backwards navigation from an exam have to be weighed against the impact on student performance along with their concerns and anxiety. With the use of computer-based testing, faculty have opportunities that do not exist in paper-based tests. One advantage of eliminating backwards navigation is the ability to use questions that build off of previous questions, without giving away the answer to the previous question. Also, if navigation is restricted, the instructor does not have to be concerned that the one question could help answer earlier questions on the exam. For example, if a short case described a patient with cellulitis, the first question on that case could be the correct assessment of the patient's medical condition. The student could select from choices of cellulitis, osteomyelitis, and others. The next question could then ask the student to select the most appropriate antibiotic for the patient's cellulitis. With backwards navigation, the student would now be able to determine the prior answer was "cellulitis" by being in the stem, or possibly by the antibiotic options if they were all for cellulitis and none were for osteomyelitis. Without backwards navigation, the instructor can better assess the ability of the student to treat cellulitis and not have the students answer contingent on getting the first question correct. This could also lead to a third question about appropriate monitoring of the antibiotic selected. Without backwards navigation, the student could have thought the case was osteomyelitis, and then selected a quinolone antibiotic, and then would select monitoring for the quinolone antibiotic. Without backwards navigation, the student is not at risk of this "waterfall" of missed questions because they got the first one wrong. The correct monitoring of the correct antibiotic could be assessed, such as a β-lactam antibiotic. In this scenario, the removal of backward navigation not only provides a better assessment of the individual concepts in the

exam, but also avoided having one wrong answer impact subsequent questions where the student may know the correct answer.

Another advantage of eliminating backwards navigation is related to exam integrity. Creating reliable and valid assessments, particularly multiple-choice questions, is a time-consuming process so it is reasonable that precautionary measures should be taken to protect those questions. While students excused to leave the exam room for the restroom or other reasons have always had the opportunity to seek out answers to questions they have seen on the exam, current technology has made it easier to get information about the exam material and more difficult for exam proctors to monitor. Eliminating backwards navigation does not remove the potential for reviewing material if the student leaves the classroom, but it does reduce the number of answers that can be changed down to a single question. Likewise, given the requirements of both Title IX and the Americans with Disabilities Act, restricting a student from leaving the classroom during an exam is not always possible.

Restricting backward navigation may also benefit exam integrity within the classroom. When backward navigation is available, an exam taker may see an answer selected from another student's computer screen and navigate forward or backward to change their answer. With backward navigation eliminated, the student cannot change responses to questions already completed, and cannot skip ahead in the exam to find a question not yet encountered. Enhanced exam integrity is a benefit to students, as well. Students are anxious, not only about their own performance during tests, but also the performance of classmates. With scholarships and grant dollars becoming more limited, confidence that exam scores and grade point averages will not be impacted by academic dishonesty is a relief to the students, as well as to the faculty.

Finally, the ability to mirror the testing used in both the NAPLEX and MPJE offers students multiple opportunities to practice with this restriction. While the primary goal of all pharmacy educators is to train high quality practitioners, the need for these students to successfully pass both the NAPLEX and MPJE are critical to their ability to be licensed.

Our study has a number of limitations that should be mentioned. The current study did not evaluate the impact of eliminating backward navigation on student test anxiety. When the new policy was announced, student leaders did express their concerns about both performance and anxiety. The concerns were particularly high among 2nd and 3rd year pharmacy students, for whom this would be a change in testing procedures from previous years. For the 1st year pharmacy students, this was likely viewed as less of a disruptive change. In an effort to be transparent and allay fears, course coordinators agreed to monitor and communicate exam performance relative to past years. While this likely did not eliminate all student concerns, anecdotal evidence suggests that anxiety and concerns went down over the course of the semester.

This evaluation consists almost exclusively of multiple-choice questions and no essay or short answer questions were evaluated. While graduating class years are generally comparable, and we did account for variation in GPA at the end of the 2nd semester of the program, sex of the student, and student age. There may be other student and/or faculty characteristics not included in the analysis that might impact exam performance. Comradery and study habits were not evaluated, for instance. The impact on the subgroup of students with exam length accommodations (150% of the time of the standard examination) was not analyzed due to the small sample size.

Despite its limitations, our study reports on data not previously available regarding the impact of changing navigation through a computerized exam at the level of a professional, graduate program. As computer-based testing becomes more common and faculty have more decisions to make about the assessment features they use, it is important to consider the potential impacts on student performance. It is also important to consider timing and communication of any assessment changes. Clearly communicating the reasons for changes in testing procedures to students may not eliminate all concerns, but it can help to reduce the some of the confusion, anxiety, and potential animosity. Other approaches include transitioning the change gradually with each incoming first-year class, instead of changing exam procedure across all years of the program at once. Finally, giving students an opportunity to take a low stakes or practice exam that restricts navigation may provide students an opportunity to familiarize themselves with the changes without the pressure of a high-stakes summative assessment.

CONCLUSION

Changes to evaluation processes and procedures should assess the potential unintended consequences on learner performance. In this initial evaluation of the impact of eliminating backwards navigation in a professional pharmacy program, a model was developed for assessing the impact on exam scores and test time. Across all 3 years of the didactic pharmacy curriculum, no adverse effect on scores or testing time was found.

REFERENCES

1. Al-Saleem SM, Ullah H. Security considerations and recommendations in computer-based testing. *Sci World J*. 2014;2014:562787. doi:10.1155/2014/562787.
2. Thurlow M, Lazarus SS, Albus D, Hodgson J. Computer-based testing: practices and considerations. *National Center on Educational Outcomes, University of Minnesota*. 2010;Synthesis Report 78.
3. Luecht RM, Sireci SG. A review of models for computer-based testing. *College Board*. 2011;Research Report 2011-12. https://files.eric.ed.gov/fulltext/ED562580.pdf. Accessed July 20, 2020.
4. International Test Commission (ITC). The international test commission guidelines on the security of tests, examinations, and other assessments: international test commission (ITC). *International Journal of Testing*. 2016;16(3):181-204.
5. Ray ME, Daugherty KK, Lebovitz L, Rudolph MJ, Shuford VP, DiVall MV. Best practices on examination construction, administration, and feedback. *Am J Pharm Educ*. 2018;82(10):Article 7066.
6. Hardcastle J, Herrmann-Abell CF, DeBoer GE. Comparing student performance on paper-and-pencil and computer-based-tests. 2017. https://www.aaas.org/sites/default/files/s3fs-public/Comparing%20Student%20Performance%20on%20Paper-and-Pencil%20and%20Computer-Based-Tests-AAAS%20Project%202061%20AERA%202017.pdf. Accessed July 20, 2020.
7. Eaves RC, Smith E. The effect of media and amount of microcomputer experience on examination scores. *J Exp Educ.* 1986;55(1):23-26.
8. Harvey AL. Differences in response behavior for high and low scorers as a function of control of item presentation on a computer-assisted test. [Ph.D.]. University of Nebraska Lincoln; 1987.
9. Pommerich M, Burden T. From simulation to application: examinees react to computerized testing. Paper presented at: Annual Meeting of the National Council on Measurement in Education; April 25-27, 2000; New Orleans, LAWise SL. Examinee issues in CAT. 1997. https://files.eric.ed.gov/fulltext/ED408329.pdf. Accessed July 20, 2020.
10. Caetano ML, Pawasausakas P. A retrospective analysis of the impact of disabling item review on item performance on computerized fixed-item tests in a doctor of pharmacy program. *Curr Pharm Teach Learn.* 2020;12(5):539-543. doi:10.1016/j.cptl.2020.01.018.
11. Naglieri JA, Rojahn J. Gender differences in planning, attention, simultaneous, and successive (PASS) cognitive processes and achievement. *J. Educ. Psychol.* 2001;93:430. doi:10.1037/0022-0663.93.3.430.
12. Imlach AR, Ward DD, Stuart KE, et al. Age is no barrier: predictors of academic success in older learners. *NPJ Sci Learn*. 2017;2:13. Published 2017 Nov 15. doi:10.1038/s41539-017-0014-5. Accessed July 20, 2020.
13. Dodeen H. Teaching Test-Take Strategies: Importance and Techniques. *Psychology Research.* 2015;2(5):108-113 doi:10.17265/2159-5542/2015.02.003.

Table 1. Student Demographics

	2018	2019	*p*-value
1st Year Students			
Age (Years)	22.8	22.9	0.78
Female (%)	75	66.7	0.33
2nd Year Students			
Age (Years)	23.9	23.3	0.41
Female (%)	67.5	73.3	0.17
P1 GPA	3.37	3.38	0.9
3rd Year Students			
Age (Years)	22.3	22.9	0.26
Female (%)	70.7	66	0.28
Female (%)	3.45	3.40	0.53

Student age and GPA were compared using the Student's t-Test; Gender was compared using the Chi-squared test.
P1=first professional year; GPA=grade point average

Table 2. Comparison of Identical Questions

Exam	Program Year	# Identical Questions	2018 Exams Mean Correct (%)	2019 Exams Mean Correct (%)	% Difference between 2018 & 2019 for Identical Questions	*p*-value[a]
Pharmacy & Healthcare	P1	18	13.8 (76.6)	14.4 (80.0)	+ 3.3%	0.18
Pharmaceutical Care	P1	19	17.6 (92.6)	17.1 (90.0)	- 2.6%	0.18
PT1 – Nutrition Exam 1	P2	25	22.2 (88.8)	21.8 (87.2)	- 1.6%	0.34
PT1 – Nutrition Exam 2	P2	19	15.6 (82.1)	15.9 (83.7)	+ 1.5%	0.42
PT3 – Endocrine	P3	34	29.2 (85.9)	28.6 (84.1)	- 1.8%	0.25
PT3 – Infectious Disease	P3	35	28.5 (81.4)	28.3 (80.9)	- 0.5%	0.71

[a] Student's t-Test

P1=first professional year; P2=second professional year; P3=third professional year; PT=Pharmacotherapy

Table 3. Comparison of Full Exams

Exam[a]	Program Year	2018 Exams Total Exam Takers	2018 Exams Total Exam Questions	2018 Exams % Correct (SD)	2019 Exams Total Exam Takers	2019 Exams Total Exam Questions	2019 Exams % Correct (SD)	*p*-value[b]
Pharmacy & Healthcare	P1	61	42	87.0 (6.9)	54	48	88.0 (6.3)	0.42
Pharmaceutical Care	P1	60	50	87.0 (6.3)	54	49	87.0 (7.9)	1.00
PT1 – Nutrition Exam 1	P2	57	72	89.0 (6.2)	60	63	89.0 (6.1)	1.00
PT1 – Nutrition Exam 2	P2	57	57	78.0 (9.9)	60	64	82.0 (8.7)	0.02
PT3 – Endocrine	P3	58	82	82.0 (7.7)	53	78	83.0 (6.9)	0.47

[a] The PT3 Infectious disease exam excluded because 2019 questions were drawn from 2 exams given during the 2018 semester.
[b] Student's t-Test
SD=standard deviation; P1=first professional year; P2=second professional year; P3=third professional year; PT=Pharmacotherapy

Table 4. Comparison of Total Examination Completion Time

		2018 Exams			2019 Exams			
Exam[a]	Program Year	Total # Questions	Median / 90% / Max (Minutes)	Mean Time / Question (Seconds)	Total # Questions	Median / 90% / Max (Minutes)	Mean Time / Question (Seconds)	Average Time / Question Comparison *p*-value[b]
Pharmacy & Healthcare	P1	42	39/54/76	58.20	48	43/56/67	54	0.16
Pharmaceutical Care	P1	50	34/47/58	48.60	49	31/40/74	39.6	<0.001
PT1 – Nutrition Exam 1	P2	72	59/93/126	55.20	63	57/77/100	53.4	0.59
PT1 – Nutrition Exam 2	P2	57	127/198/204	142.20	64	112/152/169	106.8	<0.001
PT3 – Endocrine	P3	82	100/116/127	72.00	78	97/121/157	71.4	0.78

[a] The PT3 Infectious disease exam was excluded because 2019 questions were drawn from 2 exams given during the 2018 semester.

[b] Student's T-test

Max=maximum; P1=first professional year; P2=second professional year; P3=third professional year; PT=Pharmacotherapy

Table 5. Adjusted Change in Identical Question Scores Between 2018 and 2019 Exams[a]

Exam	Program Year	Unadjusted[b] Change in Mean Score	Unadjusted[b] *p*-value	Adjusted[c] Change in Mean Score	Adjusted[c] *p*-value
Pharmacy & Healthcare	P1	0.58	0.18	0.58	0.18
Pharmaceutical Care	P1	-0.43	0.18	-0.40	0.22
PTI – Nutrition Exam 1	P2	-0.39	0.34	-0.39	0.23
PT1 – Nutrition Exam 2	P2	0.28	0.42	0.17	0.60
PT3 – Endocrine	P3	-0.57	0.25	-0.39	0.36
PT3 – Infectious Disease	P3	-0.27	0.71	0.08	0.89

[a] Linear regression results. Exam Year regression coefficient represents the change in score between 2018 and 2019
[b] Unadjusted models only include exam year
[c] Adjusted P1 models include exam year, age and sex. All other adjusted models include exam year, age, sex, and P1 GPA.
P1=first professional year; P2=second professional year; P3=third professional year; PT=Pharmacotherapy; GPA=grade point average.