

Assessing Student Performance in Geometrical Optics Using Two Different Assessment Tools: Tablet and Paper

Gregory M. Fecho, OD, Jamie Althoff, OD, and Patrick Hardigan, PhD

Abstract

The use of tablets to deliver in-class exams may be appealing to instructors wishing to take advantage of this evolving technology. This study compared two groups: students who took exams using traditional pencil and paper methods, and students who took exams via the SofTest-M application (app) on their iPads. There was no statistically significant difference in exam averages between the two groups. The SofTest-M app provides a valid alternative to pencil and paper for delivering examinations. However, students prefer taking exams using pencil and paper. The results of this study may influence educators' decisions on whether to adopt tablet-based assessments.

Key Words: *tablet, testing, computer-based assessment, paper assessments, student opinions*

Background

Since the release of the first iPad in 2010, tablets have become increasingly prominent in educational settings. According to a 2015 survey by the Pearson Foundation, 51% of college students in the United States own a tablet and use it for academic purposes, while in 2011 only 7% owned a tablet.^{1,2} Another survey, by McGraw-Hill, found that 81% of college students used mobile devices for studying in 2014, a 40% increase from 2013.³ The portability and ease of use of tablets can make many aspects of teaching and learning more interactive and more convenient. It has become common to use tablets in the classroom for conducting formative assessments such as mid-lecture questions to check for student comprehension.^{4,5} However, no publications have reported on the utilization of tablets for administering summative assessments such as midterm and final examinations. There are many reports regarding the use of personal computers for this purpose, some showing no significant difference in student performance,^{6,7,8} and others indicating that computer-based testing methods should not be considered equivalent to pencil and paper testing.^{9,10} Clariana and Wallace⁹ describe the “test mode effect,” whereby otherwise equivalent tests yield different results when completed with different methods. Their study found that undergraduate students performed better on computer-based assessments than on paper-based assessments. Their results conflict with a previous study in which Lee and Weerakoon¹⁰ discuss the significance of computer anxiety when completing computer-based tests. They found that one-third of health professions students in a microbiology course experienced computer anxiety, and the students performed significantly better on paper-based tests than on computer-based. Leeson¹¹ describes important factors, such as the student’s familiarity with computers, text readability due to screen resolution and font characteristics, and the intuitiveness of the user interface, which may affect student performance on computer-based assessments. Tablet use differs significantly from computer use in these aspects, and there are additional differences such as screen size and the use of a touch screen instead of a mouse. Therefore, the available literature describing computer-based testing should not be generalized to testing on iPads.

Since August 2012, Nova Southeastern University College of Optometry (NSU) has mandated that every incoming student acquire an iPad for use in classrooms and labs. During the first week of classes, a general orientation is given to familiarize the students with the basic functions of the iPad, including suggested apps, e-mail setup, and data backup strategies. Students have embraced the technology for taking notes and viewing interactive class materials. In Fall 2013, the College of Optometry acquired the ExamSoft program and began utilizing the SofTest-M app. Like other colleges of optometry, the NSU program traditionally administers in-class examinations using printed copies and Scantrons. Questions are typically multiple-choice, with the occasional short answer format, depending on the course. The SofTest-M app allows for a secure assessment to be taken on an iPad by downloading an encrypted exam file that can only be opened with a password. Before the exam starts, the iPad must be “locked” with the Guided Access setting, making it impossible for the user to view other apps or capture screenshots while the exam is in progress. Questions and answers can be randomized to improve security, and several types of test questions can be created. The ability to include pictures or videos within questions increases the potential to test the student’s clinical observational skills and allows the instructor to test in ways not possible with pencil and paper examinations. The ExamSoft program utilizes a user-friendly and secure cloud-based test bank creation system, and a powerful reporting system allows instructors to provide more effective and constructive feedback to the student. Tagging test items with categories allows for long-term tracking of student performance in areas of concern, providing valuable information for individual student growth or overall course or program development. Data collected from these reports can allow instructors to gather more meaningful outcome measures from their courses than simple test averages. Depending on how each question in the test data bank is categorized, the instructor can gather outcome measures on specific topics, test question types, and even college-level accreditation outcome measures.¹²

All incoming optometry students in the Fall of 2013 were given an introduction to SofTest-M as part of the general iPad orientation. This orientation included the requirement to take a mock examination to become familiar with the mechanics of the program. After this general orientation, some instructors began using the program to administer small assessments such as lecture and lab quizzes. These initial experiences with the SofTest-M app and ExamSoft program showed great potential for efficient and secure creation, distribution, and analysis of assessments. Despite some reluctance to using ExamSoft expressed by a few vocal students, grades on these quizzes did not seem to be affected. There were no significant technical difficulties relating to test administration, grading or student test reviews when using ExamSoft.

The numerous benefits of the system for both instructors and students were clear during initial trials. Nevertheless, instructors may feel that adopting a new way of administering an examination could negatively influence student performance. Technical issues may arise, or psychological barriers when using technology in the exam process could negatively affect student performance on critical examinations. Therefore, we felt further research was warranted to determine whether testing on iPads is an acceptable and valid alternative to traditional pencil and paper testing. The goal of this study was to determine if any statistically significant difference exists in test scores in Geometrical Optics when tests are completed using the SofTest-M app on an iPad instead of pencil and paper. Student opinions on using this software were also obtained to determine whether student attitude correlates with exam score averages. Results of this study may influence educators’ decisions to adopt this technology in their courses.

Based on our initial experiences with the software, we hypothesized that Geometrical Optics test score averages would not be significantly affected when tests were completed using the SofTest-M app on an iPad instead of the traditional pencil and paper method.

Methods

The Institutional Review Board of NSU determined this study to be exempt from review for the following reasons: Test scores were naturally available to one of us as the instructor of record for the Geometrical Optics course, the OAT scores and GPAs that were required in the comparison of the groups involved in this study were provided to us de-identified and as separate data sets, and participation in the survey was voluntary and anonymous with minimal risk to participants.

This study compared two groups: students who took exams using traditional pencil and paper methods (Fall 2013 group), and students who took exams via the SofTest-M app on their iPads (Fall 2014 group).

During the Fall 2013 semester of Geometrical Optics, three exams were given using pencil and paper and Scantrons. The two midterm exams were each weighted as 20% of the course grade, and the cumulative final exam was weighted as 40% of the course grade. An additional 20% of the course grade was based on homework and quiz scores. Each exam consisted of 26-28 multiple-choice questions with a two-hour time limit. Students used their own calculators for the exams. There were 111 new, first-year students enrolled in the course. One hundred and ten students took the first and second exams, and all 111 students took the final exam. These students had completed the iPad orientation, along with a mock quiz on SofTest-M, at the beginning of the semester, and the instructor administered small-stakes lecture and lab quizzes with the SofTest-M app during this semester to become familiar with test creation, administration and grading. Our Technology Advisory Committee felt it necessary to pilot SofTest-M on these smaller low-stakes quizzes for the 2013 academic year before adopting the new technology for higher-stakes, in-class examinations for Geometrical Optics starting in the Fall of 2014.

During the Fall 2014 semester of Geometrical Optics, three in-class exams were given using SofTest-M on iPads. The exams were identical in weighting, content and order of presentation to the exams given in 2013, with the exception of the last question on the final exam. Upon review in 2013, this item was ambiguous. It was removed from the 2014 final exam, and it was not included in the data for either year. The exams had the same two-hour time limit as the previous year. Students used their own iPads and calculators for the exams. There were 106 new, first-year students enrolled in the course. Two additional students were repeating the course, and their data was not included. One hundred and five students took the first and third exams, and 106 students took the second exam. All of the students had completed the iPad orientation with a mock SofTest-M quiz during the first week of classes, and they had used their own iPads to take at least seven small-stakes SofTest-M quizzes during the early weeks of the Geometrical Optics lecture and lab courses. This allowed all students to become familiar with the iPad testing format before the first exam on the iPad was given. In addition, while the students' first experience with a high-stakes examination using SofTest-M was in this Geometrical Optics course, it is worth noting that they were also using SofTest-M in another optometry course during the Fall 2014 semester for take-home quizzes and in-class examinations.

Protocol for test administration using pencil and paper exams was standard and straightforward during the Fall 2013 semester. However, the use of SofTest-M during the Fall 2014 semester required some changes to the testing protocol. Students were required to download the encrypted exam file onto their iPad. These exam files were available to download several hours before each test began, but the exams were password-protected and therefore inaccessible until the actual start time of the test. There were two instructors present during each test administration who were familiar with SofTest-M and available to address any technical issues during the exam. There were two extra iPads available for student use as needed, as well as several backup paper copies of the exam. Upon entering the exam room, students were given the exam password and several sheets of scratch paper and were expected to prepare their iPads to start the exam. This entails setting the iPad to "Airplane Mode" in order to block internet access and turning on "Guided Access" so that the students are not able to leave the SofTest-M app or take screenshots during the exam. The software requires these settings in order for a secure exam to be opened in the app. Once all students were prepared, they were told to begin the exam. Upon starting the exam, the SofTest-M app began the two-hour timer set by the instructor. At the end of the time limit, the

exam automatically saved and closed. As students completed the exam, they would return their scratch paper to an exam proctor and show the preceptor that their exam was permanently closed and therefore unable to be viewed or modified. Upon leaving the room, students would exit Guided Access mode and reconnect to the internet in order to upload their completed exam file. Any students who forgot to upload their exam file were sent an e-mail reminder, and all exam files were always uploaded and ready to grade within several hours of the end of the exam time.

A quasi-experimental design was employed to look for differences between the two groups: the pencil and paper test group (Fall 2013 group) and the iPad test group (Fall 2014 group). The dependent variable was the raw score on the three Geometrical Optics exams, with a maximum possible score of 26 on the first and second exams, and 28 on the final exam. To control for variation in academic ability between classes, comparisons were made between the pencil and paper test group and iPad test group by undergraduate GPA and academic average OAT score. De-identified GPA and OAT averages for both groups were provided by the Dean of Student Affairs. Descriptive statistics were calculated for all study variables. The Welch t-test was employed to look for differences between the two groups by each exam. Welch's t-test is more robust than a Student's t-test and maintains type I error rates close to nominal for unequal variances and for unequal sample sizes. Furthermore, the power of Welch's t-test comes close to that of Student's t-test, even when the population variances are equal and sample sizes are balanced.¹³ Statistical significance was found at $p < 0.05$ and R 3.1.2 was used for all analyses.¹⁴

An online survey was e-mailed to students at the conclusion of the Fall 2014 Geometrical Optics course to ask them about their experience with the software. Two weeks later, an e-mail reminder was sent. The Fall 2013 students were not surveyed because they were only exposed to small-stakes quizzes and did not take course examinations in class using SofTest-M. The students chose one of five answers for each of the survey questions: "Strongly Agree," "Agree," "Neither Agree nor Disagree," "Disagree," or "Strongly Disagree." We assigned a value of 5 for "Strongly Agree," 4 for "Agree," 3 for "Neither Agree nor Disagree," 2 for "Disagree," and 1 for "Strongly Disagree." The mean response and standard deviation were then calculated for each question. The midpoint value for the 5-point scale was 3.0. For statistical purposes we treated the scale as an approximation to an underlying continuous variable whose value characterizes the respondents' opinions or attitudes toward SofTest-M.¹⁵ In describing the trends of student feelings, we used any mean of 3.5 or above as representing "Agree" and a mean of 2.5 or less as representing "Disagree." For simplicity in presenting the survey responses in **Table 2**, we combined the "Strongly Agree" and "Agree" responses and calculated the percentage of students that generally agreed with a given statement. We also combined the "Strongly Disagree" and "Disagree" responses and calculated the percentage of students that generally disagreed with a given statement.

The survey also included free text boxes so students could share strengths, weaknesses and any other general feedback regarding SofTest-M. We used a grounded theory approach in analyzing these responses, relying on our empirical observations and data to qualitatively evaluate the free text responses and inductively obtain the general themes.¹⁶ The steps we followed in applying grounded theory to our data analysis included open axial, and selective coding.¹⁷ The open coding phase involved listing all of the individual statements collected from the survey, analyzing the comments and summarizing the main point of each comment to create an initial label for each statement. For the purposes of this study, a topic was considered significant when three or more comments were given the same label, while comments that were raised only once or twice were not further considered during axial or selective coding. In the axial phase we analyzed the labels made in the open coding phase for relationships and grouped comments into categories and subcategories according to their commonalities. Finally, in the selective coding phase, we were able to determine the broader overriding themes that emerged from these data.

Results

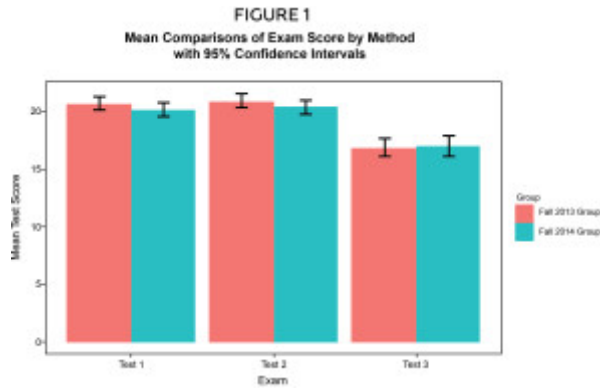


Figure 1: [click to enlarge](#)

One-hundred and eleven students constituted the pencil and paper test group, while 106 were in the iPad test group. To examine if the groups had similar academic abilities, we compared their average undergraduate GPA and OAT scores. Using a Welch t-test we found no statistically significant difference by GPA or OAT score (**Table 1**). Using a Welch t-test we also found no statistically significant difference between the groups for any Geometrical Optics exam (**Figure 1**).

Survey results

TABLE 1
Descriptive Statistics for Study Variables

Group	N	M	SD
all 2013 Group	190	20.70	3.32
all 2014 Group	195	20.18	3.42
all 2013 Group	190	20.63	3.35
all 2014 Group	198	20.42	3.09
all 2013 Group	111	16.88	4.31
all 2014 Group	195	17.03	4.83
all 2013 Group GPA	111	3.33	0.32
all 2014 Group GPA	195	3.33	0.32
all 2013 Group OAT Score	111	314.23	20.12
all 2014 Group OAT Score	195	317.60	21.96

Table 1: [click to enlarge](#)

A total of 60 out of 106 students from the iPad group completed the survey, yielding a response rate of 57%. Survey results are presented in **Table 2** and are divided into two categories: student opinions of the SofTest-M application, and student opinions on comparing in-class examinations on paper vs. in-class examinations on the iPad using SofTest-M. Most students agreed that the basic technical aspects of SofTest-M were simple to use. The students found SofTest-M easy to use (4.22 ±0.80, 90% agreed), found downloading the exams simple (4.52 ±0.60, 98% agreed), and found the software easy to adapt to (3.97 ±0.97, 75% agreed). The average response to the question that there were too many technical issues during test-taking fell slightly above our cutoff for “disagree” (2.55 ±1.17, 58% disagreed, 22% agreed).

TABLE 2
Student Survey Results

Statement	Strongly Agree (%)	Agree (%)	Disagree (%)	Strongly Disagree (%)
all questions regarding the SofTest-M application				
all SofTest-M using features	90%	7%	1%	0%
downloading exams was straightforward to learn	98%	0%	0%	0%
exam content was comparable to exams with other testing software	32%	23%	38%	27%
exam adapted to the size of the test for all device examinations	70%	20%	10%	0%
all questions on comparing in-class examinations on Paper vs. In-class Examinations				
all SofTest-M technical				
all traditional/pencil and paper exams were	90%	4%	0%	0%
that traditional exams are easier to learn on than any other software used in class	90%	0%	0%	0%
that there were additional pencil and paper exams	30%	0%	0%	0%
that students who prefer to complete only in-class SofTest-M	30%	0%	0%	0%
to complete other take the exams on the iPad application after experiencing	90%	0%	0%	0%
that students who prefer to complete exams on the iPad application in class exams	70%	0%	0%	0%
found and presentation of exam content was comparable to pencil exams	30%	0%	0%	0%
exam content was comparable to pencil exams	30%	0%	0%	0%
exam content was comparable to pencil exams	30%	0%	0%	0%

Table 2: [click to enlarge](#)

On average, students prefer to take traditional pencil and paper exams (3.75 ±1.24, 60% agreed).

Student feelings were divided regarding whether SofTest-M negatively affected their exam performance (2.78 ± 1.42 , 32% agreed, 45% disagreed), and whether it took longer to take exams (3.03 ± 1.39 , 43% agreed, 47% disagreed), and they were divided in their feelings that more instructors should use SofTest-M (2.73 ± 1.33 , 32% agreed, 43% disagreed). On average, students disagreed with the statement that they did not have time to finish the exams given on the iPad (2.17 ± 1.26 , 70% disagreed). Student opinions were generally indifferent regarding whether SofTest-M represented an improvement in formatting and presentation of questions (2.75 ± 1.17 , 38% neither agreed nor disagreed, 40% disagreed). Finally, most students agreed that they needed scratch paper for the exams given on the iPad regardless of course content (4.20 ± 0.95 , 80% agreed).

TABLE 3
Student Comments Regarding the Use of SofTest-M
(number of times statement occurred in parentheses)

	Strengths	Weaknesses
ip	Ability to flag questions (14), better image implementation (6), raised question alerts (5)	Graded answer to generate area (1), Potential to send wrong answer (1), Answering from outside frame (2), Issues (1), Click between correct answers on iPad (1)
	Quicker grade comments (18), Detailed score reports (8)	
web	Quick to enter and delete (7), No need to spend time filling out bubbles (3)	
not clear comment		On-screen length to see (3), Easy to understand (1), Added stress (3)
		Minimize cheating (5)

Table 3: [click to enlarge](#)

Thirty-four students provided strengths and weaknesses of SofTest-M, and 12 students provided general feedback regarding their use of the software. In total, these students provided us with 159 individual statements. Each of the statements was assigned a label representing the main idea being conveyed. Because the students were asked to provide strengths and weaknesses of using SofTest-M, the comments naturally fell into these two main categories. Sixty-three cited strength comments, and 68 weakness comments were determined to be significant during open coding. Comments were then divided into six subcategories relating to software design, grading, time management, ergonomics and user comfort, stress and exam security. **Table 3** organizes the comments based on this grounded theory analysis and lists the number of occurrences for each statement.

Discussion

We have found that there are no statistically significant differences in average Geometrical Optics exam scores when comparing the traditional pencil and paper method to the SofTest-M app on iPads. Early studies of computer-based assessments describe a “test mode effect,”^{9,10} but more recently published literature implies that performance on computer-based assessments is equal to performance on traditional pencil and paper tests.^{6,7,8} Our study further shows that the mode of testing does not affect performance when using this more recently developed tablet-based assessment method.

Perhaps the most striking result from our survey is that most students preferred pencil and paper exams in spite of largely agreeing that SofTest-M is easy to use. We feel this preference is partially due to the students’ comfort with pencil and paper assessments and relative unfamiliarity with SofTest-M. Ward et al.¹⁸ found an increased anxiety level in students taking examinations on a computer compared to those who took exams using pencil and paper. They suggested that the increased anxiety level originates from the unfamiliarity with the technology, and we could expect this anxiety to lessen as exposure increases. While computerized test-taking is not an entirely new concept — the OAT and Part 2 of the National Board exam are given on desktop computers — the overall exposure to computerized testing over the entire career of the student is minimal, while paper examinations remain commonplace. The significance of this point might be apparent when comparing our study results to those of Higgins et al.¹⁹ The much younger fourth-grade students in their study preferred computer-based exams over pencil and paper exams, perhaps due to the fact that they had spent fewer years developing a strong preference for pencil

and paper exams. We agree with suggestions by Ward et al. that students' feelings toward a new technology may improve over time. If instructors in subsequent didactic courses utilize SofTest-M, it is possible that students will become accustomed to taking exams on their iPads, and this increased exposure may reduce the stress of using a new method of testing. Hanson et al.²⁰ did in fact find that students at the Indiana University School of Medicine reported fewer concerns with computerized testing after taking two tests via ExamSoft on desktop computers than before taking those tests. Stress from using a new test-taking modality was not frequently mentioned in our survey, but more students might have expressed this concern if we had included this topic in the survey questions.

It is also possible that student anxiety manifested as, or resulted from, the perception of inadequate exam time while using SofTest-M. Although most students indicated that SofTest-M did not prevent them from finishing their exam in time, there was a notable minority that felt this was an issue. Our study did not attempt to analyze any differences in speed between groups, and it is unclear whether these students' perceptions were accurate. However, research currently indicates that computerized testing offers an advantage for timed tests because answers do not have to be recorded on a separate answer sheet.^{21,22}

Responses to the survey questions also emphasized the students' strong desire for scratch paper during exams. This was also one of the most commonly cited weaknesses of SofTest-M in the free-response section of our survey. This was likely due to the inability of SofTest-M to allow annotation and highlighting of questions and answer choices, a practice that students are accustomed to on pencil and paper exams. Blazer²³ discussed the importance of this feature in a 2010 report for Miami-Dade Public Schools and referred to the use of an electronic marker to highlight passages on computer-based exams. The ability to highlight words and phrases in the question stem was added to the software after our study was completed. Perhaps as the software matures more annotation features such as drawing and writing may be added, and scratch paper will become less of a demand.



Image 1. For long questions or questions with many answer choices, students may have to scroll or adjust the window size to view all necessary information. On the right side of this screen, the Questions List is opened and showing all Flagged Questions from this exam.

[Image 1: click to enlarge](#)

The free-response section of our survey implicated the visually crowded screen as another significant weakness of SofTest-M. Students complained that information can get crowded on the screen, and that it is inconvenient to scroll through question text and answer choices for some of the longer questions (**Image 1**). When using SofTest-M, images or attachments within individual questions may also compete for screen space, although these attachments can be closed and reopened as needed (**Image 2**). The impact of scrolling on performance in computer-based assessments has been studied in various settings, and some studies suggest that the ability to view all question content without scrolling results in

higher test scores.^{19,21,24} However, we hesitate to assume these findings apply to SofTest-M exams because the process of scrolling on a computer is different than on a tablet. Scrolling through information on a computer screen involves a scroll bar, mouse wheel or page up/down keyboard keys. On a tablet, the students place their finger directly on the screen to move the content up or down to the desired location.



Image 2. A question attachment such as an image or diagram opens automatically when the question is opened. The image may obscure parts of the question text or answer choices, but it can be moved or closed and reopened as needed.

[Image 2: click to enlarge](#)

The greatest strength of SofTest-M that was mentioned in the free-response section of the survey was quickly receiving exam scores with detailed category reports (**Image 3**). Faster scoring and more specific feedback are commonly mentioned benefits of computer-based assessments,^{25,26} and Bennet²⁷ has cited these as important reasons that state education agencies are employing more online tests. At the NSU College of Optometry, Scantrons are scored using an external grading center, and any errors such as a mis-key can cause significant grading delays. ExamSoft, on the other hand, allows for direct control of the grading process with the ability to easily adjust and re-score items and assessments as needed. This has decreased the time it takes for faculty to distribute grades to the students. The ability to tag and categorize questions also enables the instructor to provide customized, detailed feedback regarding a student's performance (**Image 3**) and can help guide the student in determining particular topics that need to be reviewed.

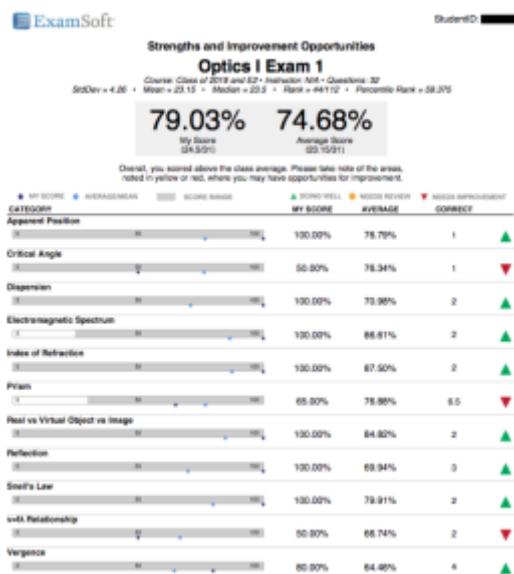


Image 3. Student score reports can be customized to

include more or less information according to instructor preferences.

[Image 3: click to enlarge](#)

The ability to flag questions was another commonly mentioned strength of the SofTest-M design. The software has the ability to flag questions so that the student can return directly to the questions by accessing a question list menu (**Image 1**). This menu can show a quick overview of the unanswered and flagged questions, and a simple tap returns the student directly to these questions. Russell et al.²⁸ and Johnson and Green²⁹ confirmed the importance of features such as reviewing questions, skipping questions and changing items that have already been answered. These features allow students to utilize test-taking strategies similar to those they have developed on pencil and paper exams.

The use of any technology introduces the potential of technical problems and software bugs, which could cause serious problems for the administration of an examination. Rabinowitz and Brandt³⁰ articulated this point, stating that computer crashes are more difficult to resolve than broken pencils. However, during our study, technical issues were rare and never prevented a student from completing and submitting an exam. Our policy is to bring two backup iPads and five paper copies of the exam for student use if needed. The backup iPads were used on occasion for students whose iPads were either stolen or broken, but the paper backup copies were never needed. During examinations for another course, we encountered software bugs in the form of app crashes that exit a student out of the exam. Because the exam is frequently and automatically saved locally to the iPad while the student is taking the exam, no exam progress is lost in the event of an app crash. Resuming an examination is trivial and simply requires a “resume code” entered by the proctor. We did not experience any of these crashes during the examinations in this study. In order to minimize technical issues with the software, we instructed students to keep both the SofTest-M app and the iPad operating system (iOS) updated to their most current versions. However, because the major annual iOS updates (e.g., from iOS 9 to iOS 10) are more likely to create compatibility issues, students are told to wait for approval from our Technology Advisory Committee before completing these particular updates.

Because SofTest-M can only be used on the iPad, and no other available apps allow for secure tablet-based exams, we did not attempt to separate student feelings regarding taking exams using SofTest-M and taking exams using iPads. While the survey focused on gathering student opinions on SofTest-M, distinguishing between feelings toward iPad testing and feelings toward SofTest-M as well as asking questions regarding the general use of the iPad during testing would have been informative. For example, feedback regarding glare from the screen, asthenopia from prolonged viewing of the display, and legibility of the print would have been valuable. It is also worth noting that the survey was only available online, and although it was simple to complete, this may have discouraged participation by students who are uncomfortable with technology and created bias in survey results. A final limitation of the study was the unavoidable use of a quasi-experimental design and the relatively small sample size. Therefore, we suggest others replicate the study.

Further research to compare the anxiety level of students taking examinations using SofTest-M and students using pencil and paper would be useful to increase our understanding of students’ mindsets during exams. Studies evaluating student performance with different question types, especially those available only with computerized testing, such as video questions, can help better define the usefulness of such unique questions in the examination. Students were divided on whether it took longer to complete their examinations with SofTest-M, and this is another possible topic for future study. Lastly, faculty experiences including test creation, category management and exam implementation experience could be explored in future research. Since implementing this study, many more faculty at the NSU College of Optometry have begun using ExamSoft, and initial anecdotal experiences have been largely favorable.

In light of our findings, instructors can confidently implement ExamSoft into their curricula in lieu of paper

and pencil exams for multiple choice-type examinations without fear of a negative impact on student performance. Findings from this study can be used to assure students that their performance will not suffer if an instructor decides to use this modality of testing. The instructor can also inform students that SofTest-M is an easy and approachable system to use. We strongly urge the use of scratch paper for calculation-based examinations because of the limited annotation capabilities of the software. The importance of scrolling completely through both the question stem and answer choices should be emphasized to students. Instructors might also consider making adjustments to test items, such as limiting answer choices or reducing font size, to fit more information on the screen at one time. Students' experiences will be optimal if they are made aware of all software features, such as question flagging, and if instructors take the time to categorize all questions and provide individual student feedback.

Conclusion

We have concluded that there were no differences in test averages between students who took Geometrical Optics examinations using pencil and paper compared with those who took exams via the SofTest-M app on their iPads. Students were divided on how they viewed their exam performance when using SofTest-M and overall preferred to take traditional pencil and paper examinations. Adoption of new technology in education should be implemented not because it may be trendy, but because it has the potential to produce effective outcome measures that can help strengthen students' understanding of the material in addition to strengthening the curriculum. ExamSoft has the potential of accomplishing these goals, and so far we have had a positive experience using the system. Students may not readily appreciate some of the most significant benefits of ExamSoft, such as the ability to categorize each question, generate individual reports highlighting a student's strengths and weaknesses, and eventually track a student's performance in each of these course topics over the entire curriculum. The advantages of such tailored feedback may become more apparent to students as more instructors take advantage of these features of ExamSoft. We hope the findings from our research help educators make informed decisions regarding the implementation and adoption of tablet-based assessments into their courses and curricula.

Acknowledgement

This project was partially funded by a 2014 Educational Starter Grant from the Association of Schools and Colleges of Optometry.

References

1. Pearson Foundation. Pearson Student Mobile Device Survey [Internet]. 2015 June [cited 2016 Apr 6]. Available from: <https://www.pearsoned.com/wp-content/uploads/2015-Pearson-Student-Mobile-Device-Survey-College.pdf>.
2. Pearson Foundation. Survey on Students and Tablets [Internet]. 2012 March 15 [cited 2016 Jan 6]. Available from: https://www.colby.edu/administration_cs/its/instruction/cstrain/upload/PF_Tablet_Survey_Summary_2012.pdf.
3. Belardi B. Report: New McGraw-Hill Education Research Finds More than 80 Percent of Students Use Mobile Technology to Study [Internet]. McGraw-Hill Education; c2015 [cited 2016 Jan 6]. Available from: <https://www.mheducation.com/news-media/press-releases/report-new-mcgraw-hill-education-research-finds-more-80-percent-students-use-mobile.html>.
4. Calma A, et al. Improving the quality of student experience in large lectures using quick polls. *Aust J Adult Learn*. 2014;54(1):114-136.
5. Glassman NR. Texting during class: audience response systems. *J Electron Resour Med Libr*.

- 2015;12(1):59-71.
6. Randall J, Sireci S, Li X, Kaira L. Evaluating the comparability of paper- and computer-based science tests across sex and SES subgroups. *Educ Meas*. 2012;31(4):2-12.
 7. Escudier MP, Newton TJ, Cox MJ, Reynolds PA, Odell EW. University students' attainment and perceptions of computer delivered assessment; a comparison between computer-based and traditional tests in a 'high-stakes' examination. *J Comput Assist Lear*. 2011;27:440-447.
 8. Frein S. 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.
 9. Clariana R, Wallace P. Paper-based versus computer-based assessment: key factors associated with the test mode effect. *Brit J Educ Technol*. 2002;33:593-602.
 10. Lee G, Weerakon P. The role of computer-aided assessment in health professional education: a comparison of student performance in computer-based and paper-and-pen multiple choice tests. *Med Teach*. 2001;23:152-157.
 11. Leeson H. The mode effect: a literature review of human and technological issues in computerized testing. *Int J Test*. 2006;6(1):1-24.
 12. Resource Center | ExamSoft [Internet]. ExamSoft Worldwide, Inc; c2016. [cited 2106 Jan 6]. Available from: <https://learn.examssoft.com/resources>.
 13. Ruxton GD. The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. *Behav Ecol*. 2006;17:688-690.
 14. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing [Internet]. Vienna, Austria. 2015. [cited 2015 Sept 16]. Available from: <https://www.R-project.org>.
 15. Clasen DL, Dormody TJ. Analyzing data measured by individual Likert-type items. *J Agr Educ*. 1994;35(4):31-35.
 16. Glaser BG, Strauss AL. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago, IL: Aldine Publishing Company; 1967.
 17. Smith K, Biley F. Understanding grounded theory: principles and evaluation. *Nurs Research*. 1997;4(3):17-30.
 18. Ward TJ, Hooper SR, Hannafin KM. The effect of computerized tests on the performance and attitudes of college students. *J Educ Comput Res*. 1989;5(3):327-333.
 19. Higgins J, Russell M, Hoffmann T. Examining the effect of computer-based passage presentation on reading test performance. *J Technol Learn Assess* [Internet]. 2005 Jan [cited 2016 Sep 27];3(4). Available from: <https://ejournals.bc.edu/ojs/index.php/jtla/article/view/1657/1499>.
 20. Hanson D, Braun M, Bauman M, O'Loughlin V. Attitudes toward the implementation of computerized testing at IU School of Medicine [abstract]. *FASEB J*. 2014;28(1) Suppl 533.6.
 21. Pommerich M. Developing computerized versions of paper-and-pencil tests: mode effects for passage-based tests. *J Technol Learn Assess* [Internet]. 2004 Feb [cited 2016 Sep 27];2(6). Available from: <https://ejournals.bc.edu/ojs/index.php/jtla/article/view/1666/1508>.
 22. Pomplun M, Frey S, Becker DF. The score equivalence of paper and computerized versions of a speeded test of reading comprehension. *Educ Psych Meas*. 2002;62(2):337-354.
 23. Blazer C. Computer-based assessments. Information capsule [Internet]. Miami (FL): Research Services, Miami-Dade County Public Schools (US); 2010 Jun [cited 2016 Sep 27]. 18 p. Vol No.: 0918. Available from: <https://drs.dadeschools.net/InformationCapsules/IC0918.pdf>.
 24. Bridgeman B, Lennon ML, Jackenthal A. Effects of screen size, screen resolution, and display rate on computer-based testing performance [Internet]. Princeton (NJ): Educational Testing Service (US); 2001 Oct [cited 2016 Sep 27]. 23 p. Report No.: RR-01-23. Available from: <https://www.ets.org/Media/Research/pdf/RR-01-23-Bridgeman.pdf>.
 25. Paek P. Recent trends in comparability studies [Internet]. San Antonio (TX): Pearson Educational Measurement (US); 2005 Aug [cited 2016 Sep 27]. 30 p. Report No.: 05-05. Available from: <https://assets.pearsonglobalschools.com/>.
 26. Puhan G, Boughton K, Kim S. Examining differences in examinee performance in paper and pencil

- and computerized testing [Internet]. J Tech Learn Assess. 2007 Nov [cited 2016 Sep 27];6(3). Available from: <https://ejournals.bc.edu/ojs/index.php/jtla/article/view/1633>.
27. Bennett RE. Online assessment and the comparability of score meaning [Internet]. Princeton (NJ): Educational Testing Service (US); 2003 Nov [cited 2016 Sep 27]. 19 p. Report No.: RM-03-05. Available from: <https://www.ets.org/Media/Research/pdf/RM-03-05-Bennett.pdf>.
28. Russell M, Goldberg A, O'Connor K. Computer-based testing and validity: a look back into the future. Assess Educ. 2003;10(3):279-294.
29. Johnson M, Green S. On-line mathematics assessment: The impact of mode on performance and question answering strategies [Internet]. J Tech Learn Assess. 2006 Mar [cited 2016 Sep 27];4(5). Available from: <https://ejournals.bc.edu/ojs/index.php/jtla/article/view/1652/1494>.
30. Rabinowitz S, Brandt T. Computer-based assessment: Can it deliver on its promise? [Internet]. San Francisco (CA): WestEd (US); 2001 [cited 2016 Sep 27]. 8 p. Report No.: KN-01-05. Available from: https://www.wested.org/online_pubs/kn-01-05.pdf.

Dr. Fecho [fecho@nova.edu] is an Assistant Professor at Nova Southeastern University College of Optometry. He teaches Optometric Theory and Methods and a course on Anomalies of Binocular Vision. He also provides patient care in the Eye Care Institute's Vision Therapy and Binocular Vision Service.

Dr. Althoff is an Assistant Professor at Nova Southeastern University College of Optometry. She teaches Geometrical and Physical Optics and Clinical Gerontology and she provides patient care in the Eye Care Institute's Primary Care and Geriatric clinics.

Dr. Hardigan is a Professor in the College of Medicine at Nova Southeastern University. Dr. Hardigan earned both his masters and doctoral degrees at the University of Wyoming, where he studied psychometrics. He has published more than 120 peer-reviewed abstracts and publications.