California State University, Monterey Bay

Digital Commons @ CSUMB

The Pros and Cons of Computer-Based Standardized Testing for Elementary Students

Sierra Gonzalez
California State University, Monterey Bay

The Pros and Cons of Computer-Based Standardized Testing for Elementary Students
Sierra Gonzalez
Liberal Studies Department
California State University Monterey Bay

Abstract

The purpose of this study was to find the pros and cons of computer-based standardized testing for elementary school students. This research can also be used by educators or administrators to understand the effectiveness of computer-based testing (CBT) compared to the traditional paper and pencil testing (PPT) method. There was little to no research on standardized tests specifically, so general research was found, and conclusions were drawn with the available information. The research in this paper referenced multiple studies which compared CBT and PPT scores according to age, race, gender, and ability to speak English. In addition, this research evaluated the effectiveness of CBT for students with disabilities, and it surveyed students to find their overall opinion on CBT compared to PPT. With this information, it was found that CBT scores rated roughly the same for general education and special education elementary students; the only difference that was found was that students who speak English performed better on CBT than students who did not speak English. With this information, it was concluded that CBT was an effective method for the standardized testing of elementary students. There were some concerns surrounding technology malfunctions, but with proper preparation from administrators and teachers, CBT should produce accurate results of student success.

The Pros and Cons of Computer-Based Standardized Testing for Elementary Students

I remember taking my first computerized test in high school and being nervous because paper and pencil testing (PPT) was the norm growing up. Luckily, I was familiar with computers so I could adapt easily, but that is and was not the case for many students. The school also didn't prepare anyone for computer-based tests (CBT), so many students and teachers were confused. My observations and frustration in high school led me to conduct research on the effectiveness of computer-based standardized testing; it was important to understand this testing method because of its increasing popularity and effect on larger numbers of people every year. Some groups affected by CBT were students who took the tests, teachers who administered and prepared their class, and administrators who trained teachers and were responsible for policies and procedures. If these groups were educated on the pros and cons of CBT, especially ones higher up in the education system, they could then take the correct steps to ensure that everyone was as successful as possible. The first step in understanding this method of testing was to understand why and when it began.

The need for technology in the classroom and eventually CBT began many years ago when standardized testing became compulsory with the Elementary and Secondary Education Act (ESEA) of 1965. This was a law that was signed by President Lyndon B. Johnson in hopes of creating a more equal education for all students in the United States. Standardized testing was implemented so the government could ensure students were being given an equal education; if the scores reported anything less than that, they were able to give schools the tools they needed to improve (Brenchley, 2015). Approximately 45 years later, this act was replaced by the No Child Left Behind Act (NCLB) of 2001 (Alcocer & NEA, n.d.), which was signed by President

George W. Bush to help further close education gaps between students in classrooms all over the

United States. This act still required that schools administered standardized tests and forced them

to submit their scores to see whether the gaps in the education system were being closed. As if

there was not already enough pressure on schools to have their students perform well on these

tests, NCLB expected scores to be turned in quickly. So, to get a more accurate reading of

student success and produce quicker results, computerized testing started becoming more

popular. However, there were some issues with this program, and just like with computerized

assessments, NCLB had its pros and its cons.

To replace this controversial act, President Barack Obama passed the Every Student

Succeeds Act (ESSA) in 2015. This act was basically a revision of ESEA and created an even

more equal education system. A letter to Chief State School Officers described that this act

would hold all students to a high standard to help prepare them for their futures. To do this, they

still required standardized assessments to ensure students' progress and improvement of policies

(Duncan & Kind, 2015, par. 1). While standardized tests were still mandatory in this act and

schools still had to produce quick test results, the information received by the government about

student progress was shared with educators so that they could create an effective learning plan.

States were then required to create a plan for their education system to be reviewed by the

federal government and later implement into schools. While these acts were working toward

creating a more equal education system, they were forcing schools to come up with faster

methods of test-taking and score reporting without fully analyzing the benefits or challenges that

surrounded CBT. For this reason, it was important for researchers to examine how well the

current system of computer-based standardized testing was working for elementary students

across the United States.

Unfortunately, the effectiveness of computer-based *standardized* testing in elementary schools is a subject that is relatively new to the education system and has yet to be supported by long term research. The primary question that I planned to research was: what are the strengths and weaknesses of Computer-Based standardized testing for elementary school students? Other secondary questions that were necessary to answer the primary question were: Why are schools shifting to CBT? Do factors such as age or gender affect the success of CBT? How are students with disabilities performing on CBTs? Does gender, race, or age affect a student's scored on CBT? Does language affect student performance on CBT compared to PPT? Since technology is new in comparison to the "education timeline", extensive research was extremely limited and hard to find. However, it became evident that with proper preparation and practice tests for students, CBT was an efficient and effective way to track student progress and success.

Literature Review

The focus of this paper revolved around the effectiveness of CBT specifically regarding standardized testing. Due to how quickly technology advances, it was a priority to find newer studies that were not out of date. However, such specific and new research was difficult to find, so I had to modify the research. Instead of finding articles about standardized testing specifically, I had to analyze the effectiveness of any computerized testing *in general*. In addition, I often had to look at studies conducted outside of the United States, but still on elementary-aged students. With this general information, I was able to make inferences about the pros and cons of CBT for standardized tests and could offer some answers to my primary and secondary questions. It may be years of studies and research to show how well CBT is performing and create a system that will best benefit students, teachers, and schools.

Many studies produced inconclusive results on the overall effectiveness of CBT compared to (PPT) and even after much analysis, no correlations were found between which students perform better on either type of test. One of the most helpful studies conducted by the American Association for the Advancement of Scientists (AAAS; 2017), assessed students between 4th and 12th grade on their understanding of how energy in science worked (respective to their grade level). While this was technically not a standardized test like the focus of this research, AAAS's purpose for this assessment was to compare the effectiveness of several different test-taking methods. So, using two computer-based tests and a paper-and-pencil test (PPT), a total of 33,422 students participated in the study. It was found that elementary students performed better on one of the CBT tests administered, with PPT close behind, and a different CBT coming in last (Hardcastle, Herrmann-Abell & DeBoer, 2017). This essentially meant that one testing method did not do exceedingly well compared to the other since the CBT produced the best scores for some but the worst for others.

The results of the tests conducted by the AAAS were also analyzed by gender, grade level, and ability to speak English; it was found that there was no correlation between gender and test modality. The study reported that male and female students performed lower on their CBT tests compared to PPT. In addition, they reported that students whose primary language was English performed about equal on the PPT and CBT. However, non-English speakers performed well only on the PPT and poorly on both CBT's. (Hardcastle, Herrmann-Abell & DeBoer, 2017, p. 10). Similar results were found in a study on elementary-aged students who took a German developed math test in Sweden; the authors research showed that "Neither age, nor ethnicity, or gender seems to have a significant impact on the performance on paper-, versus computer-based tests" (Hassler Hallstedt & Ghaderi, 2018, p. 197). These first two studies confirmed that there

was no preferred method between CBT or PPT regarding age, gender, or ethnicity. The only

concrete results that were produced were that English-speaking students performed equally as

well on CBT and PPT while non-English speaking students performed better on PPT.

The methodology of comparing the performance of PPT were a common factor in almost all the

articles that were researched. Students were essentially administered identical tests in the form of

a PPT and a CBT and the researchers compared the results of both. However, one study by

authors Husein Taherbhai[a], Daeryong Seo[a], and Trinell Bowman[b] of Pearson[a] and Maryland

State Department of Education[b], looked at students with disabilities while others looked at

children in general education. In the first study, the authors concluded that "for students with

learning disabilities, the computer is not a hindrance and initiates a similar pattern of behavior

from these students on the [online] version as with respect to the [PPT] mode of administration"

(Taherbhai, Seo & Bowman, 2012, p. 72). Similarly, a study by John Poggio, Douglas Glasnapp,

Xiangdong Yang, and Andrew Poggio, authors from the Journal of Technology, Learning, and

Assessment (Poggio et. all, 2015) was conducted to look at the scores between PPT and CBT.

However, they conducted their study on children in general education and found that there was

not any difference in the scores produced by CBT compared to traditional paper methods. It is

clear from the four articles above that there are no numbers to support one testing method or the

other. Regardless, it is important to look at some of the difficulties that accompanied CBT so as

to be prepared for if and when they appear.

It was established that there was not statistically one better testing method than the other.

However, it was still necessary to acknowledge the potential issues that may have risen with

CBT. The next two articles had similar themes of addressing problems that one may have faced

when working with computerized tests. The first article explained that there were three factors to

consider when analyzing the potential setbacks of CBT. The first factor was the student and their potential special needs, wandering eyes, and adjustment to a new environment. Next was the site of the testing and the possibility of full labs, or a school having no computer lab and therefore bussing their students to other lab sites. Lastly, they advised keeping in mind system setbacks that could occur such as connection problems, loss of data, and data security (McHenry, Griffith & McHenry, 2004). Another author addressed possible concerns of CBT by explaining some issues to be as follows: an adequate number of computers, adequate bandwidth, the number of students testing with consideration to the days available for testing and computers available, Network security and stability, staffing, budget, and overall readiness (Schaffhauser, 2011).

Lastly, it was essential to understand the opinions of the students who were taking the tests. One researcher from Turkey administered an assessment to elementary-aged students asking them their opinions on CBT. He found that "students generally present positive attitudes ... towards [CBT]s. ... On the other hand, considering students' gender and schools they enroll at, no significant difference was found between the attitudes of boys … and girls. (Yurdabakan & Uzunkavak, 2012, p.7). A different study took a slightly different approach to find student opinions. They administered CBT and PPT and first studied which method was better and found that there was no difference in scores. Then, they surveyed the participants to find their opinions on which testing method they liked better. The results showed that students had positive feelings toward CBT due to the ease of use and comprehensive and instant feedback. Overall, they found that the most favored test was the CBT test and the least favored test was PPT (Karadeniz, 2009). Students seemed to like the computerized testing methods over the traditional paper testing method. Although there were some problems that can arise with technology, it is still becoming the norm and is what many students are learning to be comfortable with.

Methods and Procedures

Most of the research for this paper was qualitative and was found electronically on the Education Resources Information Center (ERIC), which was located within the EBSCO Information Services database. I had access to this database through CSUMB's library website, but it could also be found on the EBSCO website by searching "ERIC". From there, one could request pricing for access to the database; ERIC asked about which institution or occupation you are a part of and left a comment section for why one would like to request access to the database (in order to create a price depending on the individual).

As stated before, the research for my primary questions was very limited, so I had to research more broad questions and form predictions with the information. I began by researching keywords such as "computer-based testing" and limiting the results to "peer-reviewed" and the subject to "elementary education". I then kept "computer-based testing" for the first subject, added "paper and pencil" into the second subject field, clicked "peer-reviewed", and set the subject to "comparative analysis." The results produced did not answer the exact question regarding standardized testing for specifically elementary students, but they did offer insight on the benefits of computerized testing in general. In addition, they were not all conducted in the United States, but they did still focus on elementary school students. One quantitative research article also touched on CBT depending on gender, race, language, and age. With the general knowledge of the pros and cons of CBT and reading studies based on the opinions of students themselves, I was able to come up with results on CBT on standardized tests for elementary students.

Results

Although it was difficult to find research articles in the beginning, I was able to pull from various sources to produce results for my primary and secondary questions. The questions I initially proposed are answered below.

1. *What are the strengths and weaknesses of Computer-Based standardized testing for elementary school students?*

There were many pros and cons to CBT. Some of the negative aspects that were found were unpredictability of technology; system setbacks that could occur such as connection problems, loss of data, and data security, or lockouts form accounts. It was also important to factor school funding in terms of an adequate number of computers, compared to the number of students testing and consideration to the days available for testing. There were also some concerns about proper staffing, budget, and overall readiness of staff and students.

Regardless of these possible setbacks, one of the pros to CBT was that students had continuously reported that they preferred CBT over PPT. In addition, there were no studies showing that PPT was any more effective than CBT except with non-English speaking students. It also produced quicker test results in order to meet the standards of the state and federal governments. If students preferred CBT as their test-taking method and their schools had the means to provide it to them, it should be a fine method to use.

2. *Why are schools shifting to CBT?*

Schools were adopting CBT because it was becoming more and more necessary to keep up with the standards set by the ESSA of 2015. This education act made state

testing mandatory to track student, teacher, and school achievement. It also required

quick test results, which was why schools were shifting to computerized methods.

Computers could produce and analyze results very quickly; the information it produced

was then sent back to school so adjustments to procedures or curriculum can be made.

3. *Does gender, race, or age affect a student's scored on CBT? Does language affect*

    student performance on CBT compared to PPT?

      There were no factors that showed any difference in scores between CBT and

PPT. All the studies analyzed showed no significant gender, age, or race to affect testing

scores on CBT. The only difference that was found was that students who were non-

English speakers performed slightly better on PPT than on CBT.

4. *How are students with disabilities performing on CBTs?*

      Students with disabilities also showed no difference in test scores between CBT

and PPT. They may have required more testing time or modifications, but when those

changes are applied to both testing methods, there were no difference in scores.

Discussion

Before conducting any research, and largely due to my own experience with CBT, I

thought that traditional methods of testing with PPT would produce better test results. I felt this

way because students could be more hands-on while taking the test and could see the whole

assessment in front of them. However, after much research I found my opinion to be incorrect.

Rather, it was true the CBT was a safe and effective method of testing for elementary school

students. Regardless, it was important to understand why this was true and what were some

possible problems that may have arisen.

As a future educator, it was necessary to understand that students were beginning to

prefer CBT over PPT; it is a reality that technology is becoming more and more prominent in

children's lives and they may even be more familiar with it. It could also have been beneficial for

administrators and districts because it produced expeditious results that allow for quick

adjustments in instruction. It was also important for all school staff to understand some of the

obstacles that they may have faced so as to be best prepared for any scenario that occurs.

There were some setbacks at the beginning of the research because there was very limited

data on Computer-Based Standardized testing. However, I was able to pull from sources that

study the broader subject of CBT, student opinions, and possible "pitfalls" to form an opinion.

The information convinced me that CBT was, in fact, a safe testing method for students that

would produce accurate scores.

When forming an opinion, it was important to remember that the studies analyzed in this

research were conducted in controlled settings. In a regular classroom, there may have been more

distractions, technology issues, or other unpredicted setbacks. It was also important to note that

there were no studies that answered the direct question of this paper, but rather ones that touched

on parts of the question. With this information, I was able to form results and a conclusion, but

there was not a study dedicated to my primary question. In addition, there was still a question of

how non-traditional or low-income schools would prepare their students to take CBT in the case

that they did not have computers to practice on beforehand. I could not find any research on

performance for these schools, probably because these schools were still administering PPT for

their students.

Conclusion

Schools began shifting to CBT because the world is becoming increasingly reliant on

technology. Education acts in the United States required that all students complete standardized

tests to ensure that they were staying on track with their grade level. In addition, they required

quick test results so that feedback could be sent back to schools about the results and ways to

improve their numbers. Due to the quick results and easy analysis of scores, schools were relying

on CBT over PPT. It has been a concern to some if CBT was an effective testing method

compared to the original format of PPT. However, research consistently showed that elementary

students perform comparably well on paper-and-pencil standardized tests as they do on CBTs.

Regardless of gender, age, race, or ethnicity, students were performing equally well on

both testing methods; the only factor that made a detectible difference in score was that non-

English speaking students tested better with PPT than CBT. Overall, there were rarely ever any

measurable differences in scores on PPT and CBT in any published studies. However, it was also

found that there were a variety of issues that could arise surround CBT. Some of those issues

included internet crashes, glitches in programs, internet connection issues, data security, and lack

of funding. Regardless, it was true that students preferred testing on computers rather than with a

pencil and paper. This could have been due to quick test results and familiarity with technology.

So, in terms of numbers, there was no one method that was better than the other - there were

simply pros and cons to each. With today's rapidly growing use for technology, it seems that

CBT is an effective method of testing for students in elementary schools.

References

Alcocer, P., & NEA. (n.d.). History of standardized testing in the United States. Retrieved

from http://www.nea.org/home/66139.htm

Brenchley, C. (2015, April 08). What is ESEA? Retrieved from

https://blog.ed.gov/2015/04/what-is-esea/

Duncan, A., & Kind, J. B. (2015, December 10). Retrieved from

https://www2.ed.gov/policy/elsec/guid/secletter/151210.html

Hardcastle, J., Herrmann-Abell, C. F., & DeBoer, G. E. (2017). *Comparing student

performance on paper-and-pencil and computer-based-tests* (Rep.). Retrieved December

5, 2018, from https://eric.ed.gov/?id=ED574099 (ERIC Document Reproduction Service

No. ED574099)

Hassler Hallstedt, M., & Ghaderi, A. (2018). Tablets instead of paper-based tests for young

children? Comparability between paper and tablet versions of the mathematical

Heidelberger Rechen test 1-4. *Educational Assessment*, *23*, 195–210.

Karadeniz, S. (2009). The impacts of paper, web and mobile based assessment on students'

achievement and perceptions. Scientific Research and Essay, 4(10), 984 – 991. Retrieved

May 15, 2011 from www.academicjournals.org

McHenry, B., Griffith, L., & McHenry, J. (2004). The potential, pitfalls and promise of

computerized testing. *T.H.E. Journal*, *31*(9).

Poggio, J., Glasnapp, D. R., Yang, X., & Poggio, A. J. (2005). A comparative

evaluation of score results from computerized and paper and pencil mathematics testing

in a large-scale state assessment program. Journal of Technology, Learning, and

Assessment, 3(6). Available from http://www.jtla.org

Schaffhauser, D. (2011). High-Stakes Online Testing. T.H.E. Journal, 38(6), 28–32.

Taherbhai, H., Seo, D., & Bowman, T. (2012). Comparison of Paper-Pencil and Online

Performances of Students with Learning Disabilities. *British Educational Research*

Journal, *38*(1), 61–74.

Yurdabakan, I., & Uzunkavak, C. (2012). Primary School Students' Attitudes towards Computer

Based Testing and Assessment in Turkey. Turkish Online Journal of Distance Education,

13(3), 177–188.