

Standardized Testing and Its Current Uses

STS Research Paper
Presented to the Faculty of the
School of Engineering and Applied Science
University of Virginia

By

Matthew McDonnell

04/27/2020

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: _____

Approved: _____ Date _____
Rider Foley, Department of Engineering and Society

Introduction

“A standardized test is any examination that’s administered and scored in a predetermined, standard manner” (Popham 1999 p. 1). According to a 2015 survey administered by the Washington Post, in the United States of America, “a typical student takes 112 mandated standardized tests between pre-kindergarten classes and 12th grade” (Layton 2015 p. 1). Some of these standardized tests, such as the Stanford 10, are used as end of year assessments, much to the dismay of students and some teachers. With standardized tests comes unnecessary stress for students (Groves 2001). In 2000, despite monetary incentives to improve students’ standardized test scores, teachers cautioned parents against letting their children take the Stanford 9 as they did not affect grades (Groves 2001). Despite teachers, students, and parents pushing against standardized testing, the government still mandates them, with many state governments ordering specific tests that only have questions related to the state educational standards.

A topic up for debate is what decisions should be made based on standardized test scores. Every year, schools receive the scores of their students and comparisons to the national average. What the schools do with this data is up to them. Kimberly Moore, the head administrator at the Community Christian Academy, a private K-8 school in Charlottesville, VA, has made staffing decisions based on Stanford 10 scores. When she noticed students were struggling, she brought in a special education expert to work with individual students. She also has switched the subjects that teachers taught based on the scores, because she believed they were not teaching certain topic well.

Meanwhile, in his article, Popham claims that Standardized tests cannot be used to evaluate the quality of an education. Sometimes the material taught in the classroom does not align with the material tested on by the test. There are also inevitably biased questions on the

exams. “Historically, white students have scored higher on standardized tests than minority students” (Thomas 2018 p. 1). Popham also references a question in which students were asked which of the following foods was a fruit. He believes the question is biased against students coming from families of a lower socioeconomic status who may not have access to the same fruits and vegetables as the students from families of a higher socioeconomic status (Popham 1999).

In 2003 a study was done on students to counteract these biases. Researchers helped students fight the inner struggles and self-blame that many minority students have when taking standardized tests, but this required one on one mentoring, a resource that many schools do not have access to. Mrs. Moore wants a program that will help her analyze the Stanford 10 score data of her students so that she can make administrative decisions to better help her students. Making decisions beneficial to students based on their standardized test scores requires more research into what the scores actually demonstrate about the learning outcomes among the student body. After collecting data through standardized testing, how can data benefit the students? Teachers in the United States are dedicated to their students and would not require them to take tests without reason. In my paper I will explore how standardized test scores, specifically those of are being used to improve the education of students in the United States.

Standardized Testing and Soft Technological Determinism

The most applicable framework to standardized tests and their relation to educational decisions is soft technological determinism. Soft technological determinism “claims that the presence of a particular communication technology is an enabling or facilitating factor leading to potential opportunities which may or may not be taken up in particular societies or periods” (Chandler 1995 p. 13). Students take the standardized tests, and their scores are reported to the

schools. No decisions have to be made by the schools or government regarding scores, as the data collected only allows for action without forcing any. As Marx states, “technological change drives social change but at the same time responds discriminatingly to social pressures” (Marx 1994 p. 2). Standardized test scores cannot be used if students do not take them. With this technology comes a choice to make change or deny it. Humans are in control of the use. Looking at standardized testing as a tool through the lens of soft technological determinism allows us to see an imperfect technology with great potential.

In 2001, Martha Groves and Jessica Garrison, reporters for the *Los Angeles Times*, interviewed parents, teachers, and students in California whose schools made them take the Stanford 9, the predecessor of the Stanford 10. “At the tiny Santa Monica Alternative School House, parents of more than 60% of students in 2nd to 8th grades chose not to let their children take the test this month, saying they considered it a waste of time, and in affluent Marin County, north of San Francisco, enough students asked to be exempted from testing earlier this spring to render their two high schools ineligible for state rewards for at least two years” (Groves 2001 p. 1). There are many more stories that are mentioned of students and parents fighting against standardized testing. Eighth graders boycotted tests in New York, parents gathered in Boston to protest, and some high schools in Michigan risked losing state accreditation because many students simply did not come to school to take the tests. The problem in California, according to Michael W. Kirst, a Stanford professor, was that the test did not match the curriculum. (Groves 2001). This problem is noticed by Oliver in his 2011 paper on educational reform and technological determinism. Oliver cites Cuban’s 2002 paper saying “technology has been bought – on a massive scale and over decades – on the assumption that it will cause improvements in learning outcomes and teaching efficiency,” but “in practice, teachers struggle to integrate these

resources into their practice” (Oliver 2002 p. 376), which contributes to the theory that standardized testing and in general all educational technologies only have a soft technological deterministic relationship with educational reform. If technologies were mandated and forced into use, they would have more, but as nothing is forced, teachers, administrators, and students have the final say.

Despite the monetary incentives placed on improving standardized test scores, students, teachers, and parents still disapproved of standardized testing. Kirst claims that test questions fail to test on the topics listed in the educational standards of the state, but in his piece from 2000 for the *New York Times*, Jim Yardley says that the test writers working for the Psychological Corporation, the company that makes the Stanford 9, consult state standards before writing questions. After questions are written, they are then passed through committees who debate whether or not the questions are biased. From there, the questions are tested on students of different backgrounds to make sure students are not favored to do better than others based on race, gender, or any other factors that the CCA web application will track. (Yardley 2000)

Another problem with standardized tests that teachers point to is teaching to the test. In their classes, teachers teach more than the subject material. Teachers are also focused on the social development of their students (Salmon-Cox 1981). Teachers in California also feared this, especially when they claimed the tests did not align with the standards (Groves 2001). In a study from 2000, Bukendahl et al. (2000) interviewed teachers and publishers and exposed a divide in the belief of the validity of standardized tests. On one hand, publishers advocate for the validity of their because they are created in a professional environment that seeks to eliminate bias and create absolutely fair tests. The results when teachers were asked if the tests aligned with the standards showed that teachers did not agree with the publishers. They believed that the tests did

not completely align with the standards (ibid). When learning material is not standard across the nation, we cannot compare students from different states based on their standardized test scores, therefore, we cannot use standardized tests to measure educational quality (Popham 1999).

Case Context

For my Computer Science Capstone class, my team has been directed to produce a web application that will assist the user, our client, Mrs. Moore, in the analysis of Stanford 10 test data. The application will track every student who has attended the school along with their test scores from every year. The user will be able to add student profiles and then upload data for each student. The application will store each student's name, race, gender, English as a second or other language status, socioeconomic status, disability status, parental status, and Stanford 10 test scores from their entire time at CCA. The application will automatically flag students who have scored below their current grade level. The application will also relate each subject tested by the Stanford 10 to the teacher who teaches that subject at CCA. This will enable the user to evaluate teacher performance as well as student performance. With a search functionality, the user will be able to search for and filter students by name, score data, and any of their profile flags.

The primary way the application will let the user analyze the data will be through graphs. The user will be able to select the data they want on both axes and the graph will be displayed. The application will graph student, teacher, and subject performance over the years on both an individual and class basis. For example, the user can view a graph of a student's grade equivalent versus current grade. The user can also graph on the same graph the grade equivalent of the entire class over the same years.

With the web application created, Mrs. Moore plans on analyzing student scores and personal data to best determine how to help every student. After she inputs all the student data,

Mrs. Moore will more easily identify what areas of the school need improvement based on the Stanford 10 scores. She will be able to quickly find the students who are performing below grade level. If a whole class is struggling in a certain subject, she will evaluate the teacher. If a student is far behind their class according to the grade equivalent, she will provide the student with extra help depending on their needs. If the student has a disability, she will send them to a special education expert. By knowing more about the students and having all of their data readily available in one place, Mrs. Moore will be able to make more informed decisions on how to run her school.

My project gives insight into a small slice of how Standardized Test Scores are being used. Mrs. Moore wants an application to perform careful analysis on her students' test scores and make decisions based off of that. While she has said that her analysis has supported her decisions regarding administration, curriculum, and grade advancement, one sample will not be representative of the whole population of schools in the United States. As mentioned in the STS Section, standardized tests fit the theory of soft technological determinism and how Mrs. Moore uses them may work for her small school and not larger ones. By taking a deeper dive into the infrastructure around Standardized Testing in the United States, I hope to find what decisions are safe to make based purely on scores.

Research Question and Methods

The question I sought to answer with my research is: How are standardized tests being used to benefit the students who take them? If we are requiring students to take standardized tests, we should use the data to benefit the students. According to a survey from 1981, "teachers in the schools studied rarely used test information to mold their instruction or curricular content" (Salmon-Cox 1981 p. 4). This was almost 40 years ago. I explored whether or not this has

changed. To answer this question, I surveyed teachers and school administrators from around the United States by posting a survey on several active teacher forums and emailing principals across Fairfax County, Virginia. In addition to the survey, I conducted a policy review of policies in the United States relating to standardized testing, specifically the Every Student Succeeds Act.

I wanted out what standardized tests schools are using, and how they use the data generated from these tests. I also asked what the students and teachers like and do not like about the standardized tests and what they would change about them. With a comprehensive survey and interview of teachers and administrators across the United States and the policy review, I will determine the role of standardized tests in schools around the United States and whether or not students should be taking them.

Results

In 1965, Lyndon B. Johnson signed the Elementary and Secondary Education Act (ESEA), “the federal government’s first general foray into public K-12 education” (Hanna, 2005 p. 1). “ESEA is an extensive statute that funds primary and secondary education, emphasizing high standards and accountability” (Paul, 2016 p. 1). “It created a clear role for the federal government in K-12 policy, offering more than \$1 billion a year in aid under its first statutory section, known as Title I, to districts to help cover the cost of educating disadvantaged students” (Klein, 2015 p. 1). States were required to present plans in order to receive this funding and the Act must be reauthorized by congress every three to five years (Korte, 2015).

In 2002, George W. Bush passed No Child Left Behind (NCLB), the first “update” to ESEA (Klein, 2015, p. 1). “NCLB introduced significant changes in the curriculum of public primary and secondary schools in the United States and dramatically increased federal regulation of state school systems. Under the law, states were required to administer yearly tests of the reading and mathematics skills of public school students and to demonstrate adequate progress

toward raising the scores of all students to a level defined as “proficient” or higher by 2014.

Teachers were also required to meet higher standards for certification. Schools that failed to meet their goals would be subject to gradually increasing sanctions, eventually including replacement of staff or closure.” (Nolen, 2019 p. 1).

In December of 2015, the U.S. Department of Education released the Every Student Succeeds Act (ESSA). ESSA was created to replace NCLB and “pare back the federal role in K-12 education” (Klein, 2015 p. 1). As stated on the United States Department of Education Website, the goal of ESSA is to continue the nation’s “commitment to equal opportunity for all students”. To receive education funding from the federal government, each state must still present a plan that meets certain requirements as stated in ESSA. These requirements include the “implementation of academic standards, assessments, and accountability systems... that apply to all public schools and public school students in the state” (ESSA, 2015). They expand the requirements further saying “each state plan shall demonstrate that the State educational agency, in consultation with local educational agencies, has implemented a set of high-quality student academic assessments” which must be “the same academic assessments used to measure the achievement of all public elementary school and secondary school students in the state... be aligned with the challenging state standards, and provide coherent and timely information about student attainment of such standards” (ESSA, 2015). ESSA gives states more decision power in how and when it tests students while preventing over-testing. “At the State’s discretion [assessments may] be administered through a single summative assessment or be administered through multiple statewide interim assessments during the course of the academic year that result in a single summative score that provides valid, reliable, and transparent information on student achievement or growth” (ESSA, 2015). ESSA also stresses the need for “multiple up-to-date

measures of student academic achievement” other than tests that “may be partially delivered in the form of portfolios, projects, or extended performance tasks” (ESSA, 2015).

ESSA demands state plans that fulfill their requirements, and they have implemented a system of ensuring the requirements are met. Included in ESSA is a section stating how the review process works. “The Secretary shall establish multidisciplinary peer-review teams and appoint members of such teams who are representative of parents, teachers, principals, and other school leaders” to assist in the review of State plans (ESSA, 2015). The peer-review teams may also include “researchers who are familiar with the implementation of academic standards and assessments, how to meet the needs of disadvantaged students, and children with disabilities” (ESSA, 2015). The plan must also be open to review by the public (ESSA, 2015).

Included in a United States Department of Education ESSA web announcement are examples of how states are currently following ESSA. In North Carolina, the Board of Education created a Task Force to reduce testing by reevaluating requirements and assessment strategies. Florida reduced testing by eliminating redundancies in local and statewide standardized testing. In Rhode Island, the Department of Education and local superintendents started developing guidelines to eliminate tests that do not advance teaching and learning. New York has limited the amount of time students can take tests. In Virginia, Standards of Learning (SOL) tests are administered at the end of each grade in subjects that should have been taught during the year and teachers help decide what is tested on. More states have also revised their educational plans to use better assessments and reduce over-testing.

In addition to the policy review I created a survey, but I only received six responses. Three were from an administrator, a library teacher, and a classroom teacher in Albemarle County, Virginia. Two other responses were from two classroom teachers in North Carolina. The

final response was an email saying that I needed approval before sending a survey to teachers in Fairfax County, Virginia. Their deadlines for submission; however, did not coincide with my graduation deadline so I was unable to collect the vast amounts of data that I planned to collect. The three survey responses that were from Virginia all stated that their students took the SOL at the end of the year. One North Carolina teacher had students who took the NC End of Grade Test. The school administrator claimed that “teachers use the information for remedial services and weaknesses in instruction for the next year.” They were divided on their support for standardized tests. The administrator believes that standardized tests should be used jointly with other assessment models, which is the same opinion as the U.S. Department of Education. The library teacher, however, believes that the standardized tests are not a complete picture of a student’s strengths and therefore should not be used in evaluation. The classroom teacher who states they come from an “inner city public housing school” believes that the results are “white biased” as most of their students come from “low income, public housing, one-parent families” and therefore have a low chance of success on the SOLs. They also state that it takes too long for teachers to get specific results to aid the students who took the tests. For these reasons, that teacher does not support standardized testing. The private school teacher from North Carolina had a different testing scenario. They have their class take the Iowa Test of Basic Skills at the beginning of the school year. They then use it to fine-tune instruction for the rest of the year.

Discussion

In my policy review, I discovered that the many concerns, including over-testing, teaching to the test, and bias based on socioeconomic status, are currently being addressed by lawmakers and school districts across the country. There is definitely a commitment to fairness, improvement, and learning for all students in ESSA through their reduced testing policies and

increased test scrutiny. Today, teachers aid in the creation of the tests, and school standards are better communicated as a result of an effort to increase transparency to parents, students, and teachers as mandated by ESSA. Also, ESSA stresses the use of multiple measures for student and teacher performance. They discourage only using one test, which addresses the issue of bad test days and bad tests.

Although there is a greater commitment to equality, I can see from the response I received from the teacher in Virginia who teaches in the inner city that there is still more that can be done to meet the goals of ESSA. One of their complaints was that they could not get the results back in time to help the students who took them, which defies the ESSA requirement of assessment results being returned in a timely manner. The Virginia state plan does have the potential to work as one teacher from Virginia really appreciated the insight the tests gave; however, in actuality, the execution is not flawless.

While NCLB allowed technological determinism and gave the tests themselves too much power by allowing officials to make administrative decisions based purely on standardized test scores, ESSA allows standardized tests to better align with the theory of soft technological determinism. Standardized tests are one of the many student and teacher assessments which means one test is not the only factor in any administrative or student decisions. Schools in the United States have access to the technology that can help bring about educational reform, but are not bound to any decisions based on scores alone. There are many test companies and education experts that can be consulted which may result in good tests, but merely having the technology and blindly following would be unfair. It is up to the administrators and teachers to analyze the data and make adjustments based not only on the questions students answered correctly, but on how the student best learns, their home situation, and other factors that affect daily life that

cannot be learned from a standardized test score report. The people involved drive the change with help from technology.

One limiting factor in my research was the inability to receive enough survey responses to fully understand teacher and administrator opinions and policies on standardized tests. I needed approval to send it to an entire school district, but I would not receive approval in time to graduate. With the policy review I was able to learn what the United States' Government's view; however, teacher and administrator opinion would help me learn more about how exactly scores are used. In my research I was unable to find what other technologies and models are used to analyze these scores to help students and teachers even more. Looking into these methods and technologies will aid to the efforts already being done to create the fair tests and provide another view on what successful standardized testing looks like. Sharing these methods can also help other schools in the nation to better use technologies.

If I were to do more research in the future, I would resend the survey with approval to even more school districts to gain insight into the standardized test score analysis procedures across the country. I believe that this information should be discussed because with greater data analysis could come better strategies for education. The key to successfully implementing these technologies however is to not get carried away in letting technology decide. Ultimately, the teachers and administrators know the students and should be in charge of any decisions on a case by case basis to account for the factors that no test can. With this survey, I would also be able to know teacher opinion and generalize it to the population of the United States.

In my engineering practice of Computer Engineering, I can apply this research when I make my own decisions. Measurements I may read may not always communicate the full depth of a problem, and it must be my responsibility to take into account all possible factors in my

decisions. Secondly, I must take into account peoples' different educational backgrounds. Standards vary across states, so something that may be common knowledge to me may not be common knowledge to another person. I will also keep this research in mind when using data analysis technologies. What I have may not be the exact representation of the population, so I must think critically about the people who can be affected by my decisions.

Conclusion

Many people have expressed concern with standardized testing, but much has changed in both policies and material related to standardized testing since Salmon-Cox's interview of teachers in 1982. Teachers have more say in the material on which standardized tests test. The government has taken conscious efforts to reduce over-testing and increase fairness across socioeconomic and educational backgrounds, which are risks that Popham mentioned. Many of the fears that people have had in the past are being addressed. As long as we do not get carried away and let standardized tests and any other educational technologies determine educational reform, we have nothing to fear. Teachers, administrators, governments, and parents will fight for equality in education. They are the ones who control education and educational reform, not the technology.

I believe that we should continue administering standardized tests to students. Standardized tests, the policies surrounding them, and the decisions based on them will only improve with time. Standardized tests themselves have no say in how we teach. Instead, they are merely a measuring tool that we can use to implement human change for educational reform. This can be expanded further to all testing technologies. No test can paint a complete picture of the person who took it nor their background, therefore making decisions based only on them cannot be fair. The next steps to take are research into analysis of scores and positive decisions

based on the analysis. By knowing methods of standardized test score use, we can improve educational policies and help students even more. This, combined with careful test policies and successful implementation of these policies, can lead to improvement in education.

Works Cited

- Buckendahl, C. W., Plake, B. S., Impara, J. C., & Irwin, P. M. (2000). Alignment of Standardized Achievement Tests To State Content Standards: A Comparison of Publishers' and Teachers' Perspectives. *Journal?*
- Chandler, D. (1995). Technological or media determinism[SOURCE]
- Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, 24(6), 645-662.
- Groves, Martha and Garrison, Jessica. (2001, May 24). State's Standardized Test Spurs Scattered Backlash. *The Los Angeles Times*. <https://www.latimes.com/archives/la-xpm-2001-may-24-mn-2056-story.html>
- Hanna, Julia. (2005, August 18). The Elementary and Secondary Education Act: 40 Years Later. Harvard Graduate School of Education. <https://www.gse.harvard.edu/news/05/08/elementary-and-secondary-education-act-40-years-later>
- Klein, A. (2015, April 10). No Child Left Behind: An Overview. Education Week. <https://www.edweek.org/ew/section/multimedia/no-child-left-behind-overview-definition-summary.html>
- Korte, Gregory. (2015, December 10). The Every Student Succeeds Act vs. No Child Left Behind: What's changed?. USA Today. <https://www.usatoday.com/story/news/politics/2015/12/10/every-student-succeeds-act-vs-no-child-left-behind-whats-changed/77088780/>
- Layton, Lindsey. (2015, October 24). Study says standardized testing is overwhelming nation's public schools. *The Washington Post*. https://www.washingtonpost.com/local/education/study-says-standardized-testing-is-overwhelming-nations-public-schools/2015/10/24/8a22092c-79ae-11e5-a958-d889faf561dc_story.html
- Marx, M. R. S. L. (1994). *Does technology drive history?: The dilemma of technological determinism*. Mit Press, Cambridge MA.

Nolen, Jeannette L. and Duignan, Brian. (2019, September 24). No Child Left Behind. Encyclopedia Britannica. Encyclopedia Britannica Inc. <https://www.britannica.com/topic/No-Child-Left-Behind-Act>

Oliver, Martin. "Technological determinism in educational technology research: some alternative ways of thinking about the relationship between learning and technology." *Journal of Computer Assisted Learning* 27.5 (2011): 373-384.

Paul, C. A. (2016). Elementary and Secondary Education Act of 1965. Social Welfare History Project. Retrieved from <http://socialwelfare.library.vcu.edu/programs/education/elementary-and-secondary-education-act-of-1965/>

Phelps, R. P. (2006). Characteristics of an effective student testing system. *Educational Horizons*, 85(1), 19-29.

Popham, W. J. (1999). Why standardized tests don't measure educational quality. *Educational Leadership*, 56, 8-16.

Rhodes, J. H. (2012). *An education in politics: The origin and evolution of No Child Left Behind*. Ithaca, NY: Cornell University Press

Salmon-Cox, L. (1981). Teachers and standardized achievement tests: What's really happening?. *The Phi Delta Kappan*, 62(9), 631-634.

Strauss, Valerie. (2016). New U.S. rules for standardized testing have been drafted. Here's what they mean for kids. *The Washington Post*. <https://www.washingtonpost.com/news/answer-sheet/wp/2016/04/28/new-u-s-rules-for-standardized-testing-have-been-drafted-heres-what-they-mean-for-kids/>

United States. (1965). Elementary and secondary education act of 1965 : H. R. 2362, 89th Cong., 1st sess., Public law 89-10. Reports, bills, debate and act. [Washington] :[U.S. Govt. Print. Off.]

United States. (2015). Every Student Succeeds Act. S.1177, 114th Cong.

Yardley, Jim. (2000, April 9). A Test is Born. *The New York Times*. p. A4.