

Loyola University School of Law

A Failed Means of Assessment: Alternatives to Standardized Testing

Sean Lenckus

Education Law and Policy

Dean Michael Kaufman

May 17, 2015

Standardized testing can now be considered an essential aspect of the public school system in the United States. There is a lot of money and time dedicated to ensuring that the standardized testing process works as efficiently as possible. The majority opinion on standardized testing appears to be that it is an effective means of measuring student performance. So of course, the supporters of it justify the time and money spent based on this reason. These individuals argue that standardized testing is efficient and perhaps the most efficient means of measuring academic success. However, there are others who assert that standardized testing is not the most efficient means of measuring academic performance and success. They claim that the testing process does not benefit all individual students effectively. These individuals assert what is a minority viewpoint. But perhaps they are correct. I argue that academic success can be achieved and measured more effectively in forms other than standardized testing. I will support my argument using scholarly articles and research. I will begin with explaining how standardized testing is currently flawed. I will then introduce other means of measuring student performance in schools. I will next go into more specific details regarding this topic and provide research examples to support my argument. Then, I will discuss a case in which alternative measures were attempted and describe the reaction by a testing company and a prestigious university. I will conclude with a prediction of the degree of increase in academic success if standardized testing is abandoned.

Standardized testing is not as ideal as the school system may make it seem. Testing can be said to reflect the traditional instruction of public education. The individuals in American society who held power in the past created a strategy which spread rampantly. A focal element of this strategy is standardized testing. Another entity in society eventually gained power over testing: corporations. Corporations that manufacture and score exams including Pearson and thus reap great profit (estimated at nearly \$250 million in 1999 and likely has increased). These same companies then sell teaching

materials designed to raise scores on their own tests. Politicians gain power utilizing testing by claiming they are concerned about school achievement. They use test scores as a means of accomplishing this goal because they see it as a way to chart progress. So, testing is still used as a strategy by those who hold substantial degrees of power in society.

The supporters of the traditionalist testing system stress this idea of “objectivity”. They assert that standardized testing is the best way to ensure an objective method of assessing a student’s academic progress. But what must be kept in mind is that individuals are being tested. Individuals have different subjective beliefs about many aspects of life. Test results may be measured in terms of numeric value but this value is based on the effort of the individuals who create the tests and the students who take them. The individuals who create the tests may be biased or not quite intelligent themselves. In regards to the students taking the tests, testing anxiety is something to be considered. Anxiety produces reactions that could greatly alter the validity of test scores. Furthermore, not all students take tests seriously. Some even fill in random answer signifiers.

Another aspect of reality that undermines the objective credibility of testing relates to the disparity in test results from rich compared to poor communities. Research has repeatedly found that the amount of poverty in the communities and other variables that have nothing to do with what occurs in the classroom produces the great differences in test results. Thus, tests are not a reliable method of measuring school effectiveness.

Also, the subjects tested are not tested in a form that relate to what is taught by the individual schools. A test of reading comprehension is just a series of separate questions about short passages on random topics. The questions rarely examine how students interrelate parts of the text and do not require justifications to support interpretations. In regards to mathematics, an analysis found that only 3% of the questions required “high level conceptual knowledge” and only 5% tested skills like problem

solving and reasoning. Typically, memorization is the focus of these tests. Science or social studies tests typically just focus on obscure facts and definitions. Tests may also be exceptionally difficult. So, one can assume that the means of testing regarding individual subjects hardly relate to classroom experience.

There is also a psychological component related to standardized testing. This relates to extrinsic and intrinsic motivation. Intrinsic motivation means that an individual sees a task as valuable in its own right. Extrinsic motivation means that an individual sees a task as merely a means to an end. Extrinsic motivation is unlikely to compensate for an absence of genuine enthusiasm. Punishments and rewards are examples of extrinsic motivation. Students are unlikely to respond positively in the testing environment in regards to both rewards and punishments. In fact, in a study conducted in New York, teachers were told that the school must meet a certain “standard” for testing. The teachers essentially turned into drill sergeants. So, when the teachers as well as the students were faced with this extrinsic motivating factor of punishment, the teachers responded by becoming controlling and the students responded by becoming controlled. This relates to my next point which is testing under “high-stakes”.

“High-stakes” testing refers to testing under conditions in which the school was expected to raise scores. The first issue found with this condition was that it lowered standards. The second issue is it may set teachers against each other by becoming competitive and defensive. The third issue is high-stakes testing has led to cheating. Students in these environments have repeatedly been caught cheating. The fourth issue with this condition is teachers begin to view students as a liability. The fifth issue with this condition is it leads to overspecialization in certain areas (i.e. teachers will focus more time on math than science). The sixth issue with this condition is that it narrows the conversation about *education*. The more test scores are emphasized, there is less discussion about proper goals of

schooling in the environment. Finally, and perhaps most importantly, high-stakes testing has created the idea that “teaching the test” gains importance over all other educational goals.¹

So, due to the above-mentioned flaws, it would be reasonable to assume that the majority of the public would be outraged right? However, as previously mentioned, this problem has not gained a substantial amount of exposure. Teachers, parents, and students in public schools are essentially forced to comply with this arguably outdated method. If one turns to the court system for relief, the results seem eerily similar. The courts are reluctant to invalidate standardized tests even on the ground that tests discriminate on a racial basis.² In *Debra P. v. Turlington*, the Fifth Circuit court held that Florida’s literacy test would violate the equal protection clause unless the state could establish (1) “the instructional validity of the test” and (2) that the test does not perpetuate the effects of racial discrimination in education.³ So, the court did not invalidate the test, it just placed certain standards on it.

So what can be some alternatives to this empirically flawed method of testing? One alternative form of assessment is known as “portfolio-based assessment.” In this means of assessment, teachers keep their own records of students’ work. These “portfolios” show student progress in specific areas of instruction. For example, one folder will be kept for math and one for science. Students are typically encouraged to reflect on the work that has been selected. According to some approaches, the teacher examines the portfolio at the end of a “marking period”. The teacher will record a score known as a “learning record” which is used by both the teacher and the parent in determining how well a student is progressing. Teachers can use random sampling where a number of student learning records and

¹ Kohn, Alfie. “The case against standardized testing: raising the scores, ruining the schools.” Partnership for Smarter Schools. 2000. 23. Web.

² Kaufman, Michael and Kaufman, Sherelyn. “Education Law, Policy, and Practice”. Third Edition. 2013. 1024. Print.

³ *Debra P. v. Turlington*, 644 F.2d 397 (5th Cir. 1981).

portfolios are selected randomly from each classroom. An independent group of teachers from other schools and/or members of the community reviews the records and portfolios. If there is a large disparity between the conclusions reached by the teachers and the independent readers, a third group might be called in or a larger sample from the classroom might be taken. The purpose is to determine how consistent the individual teacher applies the assessment guidelines.

This approach has several advantages. The evaluation is based on a body of student work that is completed over a long period of time as opposed to an assessment by a single test that is taken in hours or even in an hour. The approach also encourages schools and districts to invest in the professional development of teachers and outside evaluators. It pushes teachers to reflect more consistently on the quality of student work in their classroom.

The approach is also criticized in several aspects. One of the criticisms is that it works best when there are quality teachers. While this may be true, even teachers who are not considered to be of the “best” quality have the opportunity to improve when they collaborate with students in the classroom over time. More skilled teachers are also able to assist less experienced ones and this can help the less skilled teachers better evaluate their students’ work. The two types of teachers can engage in meaningful dialogue about strategic approaches in particular classrooms that can help students do better work.

Another criticism is that the portfolio requirements can place too much emphasis on teaching particular topics. For example, if a district mandates a requirement to be taught by English teachers like requiring teachers to teach composing a memo, the teachers may place too much focus on this. The teachers may spend less time focusing on other beneficial areas in English like poetry and literature analysis. There are solutions for this though. One obviously includes requiring teachers to have students focus on a wide range of topics in the general subject. For example, in the aforementioned

example, the teacher could focus on a variety of writing styles as opposed to just one. Another criticism is that it relies too much on the individual judgment of teachers and this could lead to subjective evaluation. Outside evaluators can help solve this problem. If the evaluators are sensitive to this problem, the assessment can be used to identify teachers who fail to have students work on engaging projects.

Despite its criticisms, portfolio-based assessments appear to be central to having a high quality education. It focuses on collaboration among teachers, quality work of students, and providing information to the community regarding the quality of the school. This allows for an environment in which students can arguably gain more knowledge and experience than through standardized evaluation.

Another alternative method of assessment are performance exams. These are exams but they are different from the standardized version of exams. They are tests given to all students which require students to perform a certain task like writing an essay or conducting a science experiment. Teachers assess the task and the final score is based on the sum of the two readers. A third reader might score the task if there is a difference of more than one point between the two evaluations. The benefits of such methods are reflected in the students' learning. Students actually demonstrate greater knowledge in say science after performing hands-on experiments as opposed to being given a test on the topic at hand. Teachers learn more about how to develop interesting and valuable projects for students. Performance exams can also influence teachers to utilize a wider variety of projects in the classroom. This can provide a greater and more meaningful assessment of information that other methods of instruction.

One of the problems of performance exams is that they take considerable time for the teacher and the district. Furthermore, some teachers find that the exams disrupt the flow of classroom work.

However, this is not likely to be the case if the assessments are aligned with good instruction practices. Another challenge is teachers may resort to “teaching the test” and focus on memorization skills as opposed to other skills that are meaningful to learning.

The educators should value acquiring knowledge that goes beyond what a single test can demonstrate. Sampling can be used with performance exams. The National Assessment of Educational Progress uses a sampling approach. When the NAEP reports on a group of students, the data is based on a sample of students. Maryland also uses this approach. The Maryland State performance Assessment Program covers subjects like writing, reading, math, science, and social studies. It also includes exams but each student is given an exam in only one subject area. The school is given a score that covers all subject areas and provides comprehensive data. Ultimately, this shows that these performance exams can be useful to the school and educators if integrated into the curriculum and not isolated from topics covered in the classroom on a daily basis.

Proficiency exist standards combine the approaches of portfolio-based assessment and performance exams. Students have to meet certain standards in order to be promoted to the next grade. The focus is on four broad areas: math, science, communication, and a research project. Students are given several options to demonstrate proficiency in these areas.

Exhibitions and parent teacher conferences can also be important alternative means of assessment. The science fair is one example of an exhibition. The strength of this approach is that it provides numerous ways in hopes of allowing all students to succeed. Parent teacher conferences allow for the opportunity to let parents know how their children are progressing. Teachers are allowed to engage in a dialogue with parents about the student’s experiences in class. Teachers can relay information about any concerns they might have about the students directly and the parents then have the opportunity to address these concerns themselves. The approach should be careful, however.

Teachers should explain a clear idea of the curriculum to the parents as well as the school's view about general child development. The latter can be very important because in the elementary school years, children can attach a certain set of expectations as to their achievement. And in regards to the adolescent ages, the children go through developmental changes which can affect school performance. The students/children can also be involved in these conferences. This approach should allow these individuals to reflect on what they learned and critique their progress. Parents should ultimately feel comfortable enough to raise any concern that they see as relevant at any point in the school year and not ONLY at conference time.

In regards to the school as a whole, report cards can be used to assess the effectiveness of the school's curriculum. In 1999, over two-thirds of states required such report cards. Perhaps the amount has increased since then. The information in such a report can include attendance, average grade point, the number of Advanced Placement courses, discipline issues, and school governance structure.

One cautionary measure here is that the data can be manipulated or omitted. But overall, these report cards should reflect a richer view of learning. The report cards should strive to demonstrate the type of learning inherent in individualized student portfolios and exhibitions.

Another way to measure school-wide issues is through a "school quality review team". Teams of trained educators and community members visit the school here. The teams observe classrooms, follow students, examine the curriculum, and interview parents and teachers. They then write a report which may include recommendations for improvement. The teachers and parents should be involved for this method to be successful. However, this may require additional time and resources.

Furthermore, the team reviews a school based on its own mission. If the mission is inadequate, this may not be noted in the final report. ⁴

Dr. Walter Stoup is a professor and coordinator at the University of Texas at Austin. Stoup found a way to expose children to basic math concepts using calculus and cloud computing. During the summer of 2012, he testified at the Texas Capitol regarding standardized testing. During his testimony, he used science to discuss the idea that the tests measure not what the students have learned but how well students take tests. He implemented a program at a school in Dallas in 2006 and the results were impressive. The students were learning math. However, when it came to testing, the students' scores only rose by 10%. Stoup investigated this. Stoup noticed that most students' test scores remained the same no matter what grade the students were in or what the subject tested was. He concluded that Texas' version of standardized tests were 72 percent "insensitive to instruction" which means that the tests do not measure what is specifically taught in the classroom. However, he also knew that students could improve their mastery of a subject by more than 15% in a school year but the tests couldn't measure this change. His testimony drew conflict with Pearson testing company. Pearson sells the most tests in America. In 2009, Pearson created a \$1 million endowment at the College of Education. This allowed the creation of the Pearson Center for Applied Psychometric Research, an endowed professorship, and an endowed faculty fellowship. In January 2013, Dr. Stoup was given an unsatisfactory rating at the university. There is a strong implication that because Stoup went against the interests of the testing company which the university supports, he was given such a rating. What should be noted here is that challenges to such testing companies may not always result in success. Dr. Stoup was not asked to testify at the capitol after this incident.⁵

⁴ Neill, Monty and Petterson, Bob. "Alternatives to Standardized Tests". Rethinking Schools Online. 1999. 09. Web.

⁵ Stanford, Jason. "Mute the Messenger." Texas Observer. 2014. 05. Web.

Despite this actual example, as Monty Neill discussed, success using alternative measures is possible. However, one should not ignore the idea that alternatives take time to develop. Standardized testing has had a longstanding role in our society. Any alternative measure can be implemented but one should not expect there to be immediately successful results. Furthermore, standardized testing appears to be a problem but it is only one of the many issues that our education system faces. But what appears to be extremely important in the realm of alternatives to standardized testing and education in general is parent and teacher involvement. Parents should want to have an active role in their children's' education and teachers should want to look at the student as an individual and implement these alternatives in accords with a comprehensive curriculum.

Another issue that could arise is resources. Obviously, some school districts have more resources than others and implementing alternatives will require the use of additional resources. If more attention is brought to the benefits of these alternatives, it is likely that more resources may be provided.⁶

My prediction is that many individuals and communities will eventually become fed up with standardized testing. These individuals will then truly want to implement alternative measures of testing. After the testing measures are implemented, it is unlikely that any beneficial results will immediately be apparent. However, over time and when the teachers combine the alternative means of assessment with the daily classroom tasks, I predict that there will be a distinct and substantial increase in academic success. I also predict that companies like Pearson will be forced to adapt and will work with schools with alternative assessment measures. Once there is a noticeably successful

⁶ Neill, Monty and Petterson, Bob. "Alternatives to Standardized Tests". Rethinking Schools Online. 1999. 09. Web.

implementation of the above, other communities will follow and soon, there will be a positive change in this aspect of the educational system.

To conclude, I provided evidence as to why standardized testing is flawed despite being considered by many to be an effective means of measuring knowledge of a topic. I provided background as to how the legal system reacted to standardized testing. I discussed alternatives to standardized testing and the challenges involved in implementing these alternatives. Particularly, I gave an account of a professor who criticized standardized testing and faced backlash from a large company with a demonstrably large influence.

In my opinion, education is one of the most important means of attaining self-awareness and awareness as an individual in a community. Standardized testing has become an obstacle in achieving both goals. Based on personal experience, any standardized test was not an accurate measure of my intelligence and the school I attended even made students feel guilty about possibly doing poorly on our ACTs. Now, I am glad that I realize that alternatives are possible. However, it will take time, resources, and attention from many members of these communities. But I believe that these challenges can be overcome. But awareness of the alternatives and challenges is key. Someday, standardized testing can hopefully be looked at as a laughable aspect of an undeveloped educational system. Until then, people should expose the flaws as much as possible and support the positives of other means of assessment.