# Measuring Teacher Quality:

# Are Student Standardized Test Scores an Accurate Measure?

# Lisa Rankins

Applicant to Master of Public Administration Program

# The Evergreen State College

# Abstract

In 2009, legislation promoted by the Obama Administration created a new generation of education reform titled “Race to the Top” (RTP). It builds on reforms enacted during the George W. Bush Administration under “No Child Left Behind” (NCLB). RTP ties award of competitive federal grants, in part, to use of students’ standardized test scores require in teacher assessment. This paper examines arguments for and against using student standardized testing results to measure teacher performance and recommends approaches that should be championed to replace these flawed reforms.

# Measuring Teacher Quality:

# Are Student Standardized Test Scores an Accurate Measure?

“No Child Left Behind” (NCLB) and “Race to the Top” (RTT), the latest Federal policies in education reform have both identified standardized testing as the gold standard for measuring student, and now teacher, success. At a time when teachers face continued challenges of mainstreamed classes with ever-increasing numbers of English Language Learners, increasing class sizes, and cuts to the Arts and other supplemental courses that mean less time during the day to prepare lessons and mark student work, they now also face somewhat arbitrary performance ratings based on student results on standardized tests.

RTT ties federal grants of education funding to State and Local agreement to make student test scores a component of teacher evaluations (Mikulec & Miller, 2012). Since few members of Congress, the Cabinet, or the White House, have ever spent years in a classroom, it is easy to understand their desire to overlay business paradigms on education. The product is an educated workforce. Teachers are the factory workers assembling the product. If anything goes wrong, look at the workers and the factory management to find the reason. Quality is quantifiable and easy to measure with standardized tools. Stated this way, the idea may seem absurd, but this clearly illustrates the philosophy behind many of the reforms of the past two decades.

# **Standardized Tests as a Measure of Student Success**

According to Mella Baxter (2013), a reading coach at Matanzas High School in Palm Coast, Florida, “Standardized testing provides a standards-based way of assessing students and teachers. In our transient society, consistency across the state and country is growing more important. Common standards help ensure consistency.” However, standardized testing, at best, provides a moment-in-time snapshot of learning that may be used to assess learning progress when used in conjunction with other measures. It cannot capture the full-range of learning. As noted by Ravitch (2011, p. 151), a student testing on the same test on different days may score similar but different enough to cross the line between “proficient” and “not proficient.” She further notes:

Testing experts frequently remind school officials that standardized test scores should be used not in isolation to make consequential decisions about students, but only in conjunction with other measures of student performance, such as grades, class participation, homework, and teachers’ recommendations. Testing experts also warn that test scores should be used only for the purpose for which the test was designed: For example, a fifth-grade reading test measures fifth-grade reading skills and cannot reliably serve as a measure of the teacher’s skill. (2011, p. 152)

Given that testing experts have acknowledged the need to use test results in conjunction with other measure, why would schools agree to do so? They have agreed because they cannot afford to not do so. Schools need funds to function, and those holding the purse want to see effective use of the money. Standardized test results make it easy to compare success levels; therefore, those measures are preferred by many. They are simple to read, simple to compare, and simple to critique. To a point they do measure a student’s ability to compute, read, and write.

Standardized tests of grade level learning are based on the assumption that learning is linear; however, educational professionals recognize the opposite is often true and varies from student to student. In examining the function of standardized testing in Australia, Dreher (2012) notes that all of the teachers and principals interviewed in her research felt discomfort with the concept of standardized linear learning benchmarks with one principal stating “Kids have their own trajectory and progress at different rates.”

**Using Student Results to Evaluate Teaching**

In my own teaching and learning experiences, I have experienced times where, externally, it appeared that a student was just not learning. I tutored Jonathan in reading and phonics. For weeks we seemed to get nowhere, but we continued to practice sound patterns and simple words with flash cards and repetition. Day after day, he seemed to make no progress. Had he been tested on day 30, I would have been considered an ineffective teacher. On day 31, he took my stack of phonics flashcards from my hand after expressing boredom with the lesson. He promptly whizzed through all of them and correctly identified all of them. Within another week he was reading simple books on his own. Children can learn without showing outward evidence of that learning during the process.

Most everyone has had a day where they had an “Ah ha!” experience – where multiple pieces of information and partial understanding all combine to create a moment of significant growth in learning. Who is responsible for the effective learning? Only the teacher who provided the final piece of information? The one who coached you through the first stages? What about those in the middle? The first major problem with using standardized tests to assess teacher effectiveness is the basic variance in how children learn and the multitude of contributing factors. Home environment, cognitive development stages, pre-school learning, language, nutrition, behavior expectation and training, role models, parental expectations, previous teachers, care-givers, materials, access to technology – the list is endless but each of these contributes to how and when a child learns a new concept. Holding only current teachers responsible for a student’s learning is like saying that the doctor who birthed a baby is responsible for it being healthy, whole, and male or female!

Assessing credit or blame to individual teachers has additional problems. Since in the United States, the national requirements for testing cover math and language arts, States differ in how to evaluate teachers of other subjects. According to Mikulec and Miller (2012), in Tennessee, the policy that won them a “Race to the Top” grant, requires “all teachers of all content areas to prepare students for standardized testing in math and language arts.” All teachers are evaluated even if their subject at the high school level isn’t one for which there is a standardized test. In Florida, Baxter (2013) notes that nearly one-half of a teacher’s performance is based on student scores on state tests even if the teacher’s subject is not one of those for which there is a test.

Using tests to assess teachers in such a high stakes manner leads to additional problems. For teachers, principals, and school districts, the pressure to show progress can influence otherwise ethical persons to engage in cheating to save their jobs as evidenced by the scandals in the District of Columbia and Atlanta in 2010 (Ravich, p. 265-266). The focus can also radically narrow the curriculum, whether that is to use time to coach students in test-taking skills or simply to emphasize instruction in the subjects that will be tested. In some states, “entire subjects fall away as testing subjects are kept front and center: ‘Social studies is completely gone from the curriculum,’ says Helen Gym, a parent in Philadelphia. ‘There’s very little science, and laboratory work is almost non-existent” (“Learning is More,” 2013). In Australia, Dreher (2012) notes that teachers “offered comments about the ‘push-pull’ relationship between teaching to the test and responding to their own ethical beliefs of good professional practice.” Valerie Strauss, an education writer for the *Washington Post*, writes:

Teaching to the test also narrows the curriculum, forcing teachers and students to concentrate on memorization of isolated facts, instead of developing fundamental and higher order abilities. For example, multiple-choice writing tests are really copy-editing tests, which do not measure the ability to organize or communicate ideas. Practicing on tests or test-like exercises is not how to learn even the mechanics of English, much less how to write like a writer. (2010)

**Conclusions and Actions**

Despite strong evidence against using student standardized test scores as a major component in teacher evaluations, States have responded to Federal pressures to do so. Unless citizens – parents, teachers, administrators, and other concerned individuals – engage in the process of negotiating for common sense legislation to counter this effort, effective teachers will leave our schools because of poor evaluations, discouragement, and stress. Instead of simply criticizing the ineffectiveness of the current trend, citizens need to unite to propose more effective means of improving schools and assessing teachers. By creating a more consistent curriculum base across the nation, comparisons between effective and ineffective schools and teachers can be made more logically. Limiting test preparation to a tightly-structured annual workshop instead of infringing on regular instructional time will help ensure schools maintain a rich curriculum. Developing strong mentoring relationships between teachers from different schools would increase sharing of workable strategies in teaching. Completely eliminating standardized testing would be throwing out one of many tools in the educator’s toolbox, but it is time to let each tool be used as designed, not as the ultimate multi-tool.

References

Common core assessments: More tests, but not much better. *FairTest (2012.)*Retrieved from <http://fairtest.org/common-core-assessments-more-tests-not-much-better>.

Learning is more than test score: AFT campaign favors in-depth learning over excessive testing. *Psrp reporter, Jan.-Feb. 2013* (8). *Academic OneFile.* Retrieved from http://go.galegroup.com/ps/i.do?id=GALE%7CA315507022&v=2.1&u=cfsc\_remote&it=r&p=AONE&sw=w

Baxter, M. (2013). Standardized tests: The good, the bad, and the very ugly. *TakePart, LLC.* Retrieved from http://news.yahoo.com/standardized-tests-good-bad-very-ugly-194700289.html on 2/27/2013.

Drehen, K. (2013). Tests, testing times and literacy teaching. *Australian Journal of Language and Literacy, 35.3* (2012) *Academic OneFile.* Retrieved from http://go.galegroup.com/ps/i.do?id=GALE%7CA306359315&v=2.1&u=cfsc\_remote&it=r&p=AONE&sw=w

Mikulec, E., & Miller, P. C. (2012). The odd couple: Freire and the InTASC teacher education standards. *Journal of Thought*, *47.3 (Fall 2012),* 34. Retrieved from http://go.galegroup.com/ps/i.do?id=GALE%7CA318493853&v=2.1&u=cfsc\_remote&it=r&p=AONE&sw=w

Ravitch, D. (2011). The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education. Basic Books. Kindle Edition. Retrieved from http://army.lib.overdrive.com/B6E471AB-49F4-4216-86A2-EA095B69D955/10/50/en/ContentDetails.htm?id=fd2e31c8-fa5e-46fd-b812-4830f602e147

Strauss, V. (2010). What’s wrong with standardized tests? *The Washington Post.* Retrieved from http://voices.washingtonpost.com/answer-sheet/standardized-tests/whats-wrong-with-standardized.html