Final Literature Review

Individual’s Standardized Test Scores and Teacher Effectiveness Ratings

Cortney Kittridge

EDF671-02

Educational Policy and Practice

Dr. Olivia Williams

Grand Valley State University

Fall, 2015

**INTRODUCTION**

As the old saying goes, the only thing constant in life is change, and the same is true in the world of education. Many would describe the world of education in the 21st century as collaborative, digital, individualized, and high stakes test driven. While many of the changes are for the better of our students, some would argue the focus on high stakes testing is forcing our schools to fail. While the federal government currently uses these scores for several accountability measures perhaps the largest area of question by researchers currently is the validity in tying student achievement scores to teacher evaluations.

Many attribute the major shift in education to the No Child Left Behind act which was passed in 2002. This bill is based on the premise that setting high standards and establishing measurable goals will improve student outcomes in education. The NCLB act requires states to create standard assessments on basic skills appropriate to the grade level. Federal funding is then tied to these assessments and is not received by a school unless all students participate in this standardized testing at the appropriate grade levels (MDE, 2015).

While each state develops its own standards, NCLB dramatically increased the federal government’s role in education. The greater emphasis on annual testing directly correlates to a greater emphasis on individual academic progress, district report cards, teacher qualifications, and most significantly, teacher evaluations and funding. In passing, one of the strong points of the bill was the increased accountability for each school, teacher, and student. However, after thirteen years of living this bill out many in education feel the negative impacts of NCLB are far greater than the positives. Rather than improving our school conditions, most recently many outside of the educational world lack respect and support for the teaching profession and feel our schools need ‘better teachers’ to improve test scores.

In 2011 state lawmakers passed a law that was intended to toughen up public educator’s evaluation system. While this law has been amended since then overall, there are four primary components to senate bill 103. This bill requires all school boards and administrators to adopt and implement for all educators a rigorous, transparent, and a fair performance evaluation system. This evaluation system must evaluate teachers annually at a minimum, include relevant student growth data, assess teachers on their job performance, and be used to make decisions around coaching and tenure (MDE, 2015

In 2015, the teacher evaluation law was amended to reflect effectiveness ratings based more highly on student achievement data. For the 2015-2016, 2016-2017, and 2017-2018 school years 25% of teacher effectiveness ratings should be based off of student growth data. Beginning in 2018-2019 school year, 40% of effectiveness ratings will be based off of student growth data. Additionally for the 2018-2019 school year, 50% of the student growth data used for the evaluation must be based off state assessments for grade levels and content areas in which they are administered. The current bill does allow for exemption of student growth data for a particular student for a school year upon the recommendation of the school administrator or superintendent (MDE, 2015).

In the event a teacher receives an ineffective rating for three consecutive years, the current bill states that the educator must be terminated from their teaching position. In contrast, if an educator is evaluated highly-effective three years in a row they may be moved to a biennial evaluations. Regardless of their effectiveness ratings, during the 2016-2017 school year all teachers shall be provided training around their evaluation tool or tools used by the district either in house or by an area intermediate district (MDE, 2015).

Currently, the state of Michigan uses the MSTEP as the statewide summative growth assessment. While specific content areas vary depending on the grade level, students from third to eleventh grade participate in the MSTEP testing. English language arts and mathematics will be assessed in grades 3–8, science in grades 4 and 7, and social studies in grades 5 and 8. It also includes the Michigan Merit Examination in 11th grade, which consists of a college entrance exam, work skills assessment, and M-STEP summative assessments in English language arts, mathematics, science, and social studies. These assessments may be paper or web based and take place on grade level content between April and June of the given school year.

With a growing emphasis on student achievement data as a key factor in teacher effectiveness ratings one must ask the question, what is the relationship between student assessment scores and teacher effectiveness? Does the research find a positive correlation? Or is there a change that needs to be made to our evaluation bill in order to truly measure teacher effectiveness?

**COMPETING PERSPECTIVES**

When asking whether a teacher will be evaluated effective there are several items to take into consideration. The first is knowing what education looks like for today’s teacher. Another is how teachers are evaluated and the final is the use of standardized testing in schools. As stated earlier, the role of standardized test data in teacher evaluations continues to grow as our nation searches for mechanisms to hold teachers, schools, and students accountable.

While the federal mandate is that all teachers be evaluated, the method for evaluation is chosen by each school district as long as it adheres to the required guidelines. Currently schools have selected a variety of systems focusing on a variety of components such as student engagement, focused instruction, professionalism, classroom culture, use of technology, and data. However, as stated earlier, the role of data in evaluations is set to continue to increase according to federal regulations.

Darling-Hammond did research on evaluating teacher evaluations in 2012 and what she found was quite surprising. Their research looked at the factors most effective in determining success in a child’s educational career. These factors included teacher effectiveness, class size, curriculum, instruction time, home life, community supports, peer culture, previous effective teachers, teacher preparation, summer learning loss, and the type of assessment used. While the current evaluation system focuses on student achievement being directly related to the effectiveness of the current teacher, Darling-Hammond’s research found this to be a large contributor of academic success they sited all of the other factors listed as more important. She goes on to talk about a Value Added Model which gives a percentage value to each of these factors all of which help play into an overall effectiveness rating (Darling-Hammond, 2012).

While student’s test scores appear promising as a contributor to teacher effectiveness, they carry many risk factors as well. Many argue that standardized tests scores are inappropriate to use for evaluation purposes however, very few alternatives have been suggested. Historically, when these tests are used, curricular shifts and poor teaching practices are often encouraged unintentionally (Haertel, 2009).

Their research suggests some of the problems encountered when using student test scores to evaluate teachers and to propose a model for teacher evaluation that is effective, affordable, fair, and politically acceptable. The proposed system looks at pretesting and statistical adjustments for diverse populations as well as student attendance data and portfolios of student work. While one evaluation tool will not be perfect for all cases and cannot eliminate all errors, the proposed system could help guide us to an improved outcome. The original goal of teacher evaluations was to improve teacher effectiveness and the proposed system takes into account where students enter the school year and focus on grow from there. Additionally there is room for opt out depending on a student’s background which I believe is a step in the right direction (Haertel, 2009).

Research done by Deno in 1985 suggests the importance of utilizing curriculum based measures as a means to drive instruction. Curriculum based measures are those directly tied to the curriculum taught such as end of unit tests. Deno states that the lack of confidence most teachers have in high stakes assessments results in a lack of value in the results. He goes on to argue that if teachers do not see the value in taking the assessments, they certainly won’t be okay with an effectiveness rating based on them. By utilizing assessments teachers believe in, teachers will be more willing to use this assessment data toward an evaluation on their teaching (Deno, 1985).

To further support this argument, a study conducted for many years by Sanders talks about investigating the factors considered to affect student learning. While a variety of things in education have changed, one thing that has remained the constant is the value of the teacher in student’s learning journey. Because teachers are so important in the academic journey, an evaluation of teacher’s has become even more necessary. However, the article goes on to talk about all of the factors legislators don’ think about with regards to education. This article argues that student data is not the greatest tool to use in measuring student success (Sanders, 1997).

Sanders goes on to talk about all of the factors that could negatively influence a student’s test score on any given day. He mentions home life and prior year’s preparation as the leading factors influencing a snapshot in the educational journey. As teachers, we see this every day. While many times these factors are out of our control and take place outside of the school day, the effects last throughout the school day. Imagine a child who is ill, has parents going through a divorce, recently lost a pet, moved to a new school, or didn’t get breakfast. All of these scenarios happen at some point for our students and it would be a poor representation of their true academic picture if the data used from this particular day was used to evaluate their teacher.

While many of the articles I found were not in favor of the use of standardized testing as part of the evaluation, a study from 2000 done by Darling-Hammond supported the use of student data. This research was based on data from a 50-state survey. Their study looked at whether teacher quality and school inputs were related to student performance on high stakes tests. They found that when investments were made into quality teacher’s student performance improved. Quality relates to teacher preparedness, prior training, years of effective ratings, and effective instruction. Further data analysis indicated that measures of teacher preparation and certification are by far the strongest correlates of student achievement in reading and mathematics. This was true both before and after controlling for student poverty and English language learner status (Darling-Hammond, 2000).

An additional article done on teacher evaluations discusses the movement to improve the education system suggests that teacher evaluations are the primary mechanism for improved outcomes in a K-12 school system. They argue that the number one way factor in improving student performance is teacher quality and therefore feel focusing on student test scores is necessary. The article goes on to quote an Economic Policy institute study saying, ‘In general, American public schools do a poor job in developing and evaluating teacher.’ While I disagree with this statements, I do believe this represents the thoughts of the general public in America (Lynn, 2015).

The greatest argument to the quote above was an article I read by the Radical Teacher from 2012. This article was written from a very opinionated perspective of a former teacher who recently left the field. He writes in defense of teachers focusing on teacher morale, student engagement, and other areas negatively affected by high stakes testing. Additionally, he criticizes the current evaluation system calling it ‘madness’ because of the pressure put on teachers with regards to student data. He goes on to talk about the role of the government and criticizes the number of policies and mandated curriculums being pushed down from people who have very little knowledge about public education. While this was a very opinionated article, I appreciated the perspective from someone who has worked under the education reform (Radical Teacher, 2012).

**SUPPORTING PERSPECTIVES**

Educating children is a profession unlike any other. Only those who have taught themselves or who work in a school building really understand the preparation, decision making, and energy needed any given day. Just as the government does not enter hospitals to mandate an assessment of a surgeon’s knee replacement work, they should not enter a school and require an assessment of an educator’s effectiveness. In an attempt to hold schools to a higher standard, our government has slowly ruined our world of education. If something does not change, our nation will greatly suffer from a lack of teachers going into the profession as well as a lack of well-prepared students.

The increased role of government in public schools has been a big change for all parties involve, and in my opinion, a change for the worst. Research done in June 2014 stated that the corporate education reform is the biggest threat to education in today’s society. While outsiders find the lack of teacher effectiveness a point of contention, this research sites a loss of resources, time, and support in the areas needed in K-12 education as the largest issue. Antush writes that educators are preparing curriculums that focus on high stakes testing rather than educating well rounded and responsible citizens who are ready to leave their mark on the world (Antush, 2014).

Antush goes on to talk about the New York City’s evaluation system for public educators. Several factors weigh into their teacher evaluation such as student engagement, instruction, professionalism, teacher goals, progress toward achieving those goals, and student data. Data makes up 20% of the evaluation for teachers total effectiveness rating. What I found intriguing about NYC’s evaluation system was the fact that districts or in some cases specific schools chose the assessment data they used to show growth or proficiency for their students (Antush, 2014).

I don’t think anyone would argue that data isn’t important. I believe it is how we are currently using student data that is the issue. That is why I really liked NYC’s use of data in their teacher evaluations. Using this data as less than a quarter of the overall score was encouraging but what I appreciated even more was the fact that each district or school chose what data they used. Allowing those working with students to choose data that is continuous and ongoing, rather than a snapshot in time makes so much more sense and is likely a much more accurate picture of the growth made. Additionally, the assessments given are also grade level and developmentally appropriate which is not always the case with standardized tests.

A study from Canada supports this focusing on how standardized tests could be used to promote accountability among schools and teachers. Research done in 2001 found that the only way to use high stakes testing scores effectively is to test more frequently with shorter more focused assessments. Additionally the found that students need to feel a greater buy in significance for students with regards to their testing outcomes (Casas, 2001).

Until students, the ones who are actually taking the assessments, feel connected to the data they will feel very little responsibility to do well. As the system stands currently, state testing is something that ‘happens to them’ once per year and does not directly affect them, in their minds, because the scores do not go toward a report card grade, promotion, or retention. Teachers are criticized for preparing students for these tests or teaching to the test however, the fate of their effectiveness is in student’s hands and in a majority of cases students do not see the value in performing their best on these tests.

**CONCLUSION AND IMPLICATIONS**

In conclusion, teacher effectiveness is most accurate when based on a variety of factors. Research does not support a direct correlation between student achievement scores on standardized tests and teacher effectiveness. While many of these bills were passed to help our students achieve at a higher level, most working in education would agree that they have only harmed our students.

For the first time in many years, Michigan is experiencing a shortage in teachers. Many college students are deciding not to go into the field of education because of everything happening in today’s schools. Additionally charter schools, online programs, and homeschooling are becoming more popular as parents are losing faith in our schools as well. By placing a great emphasis on standardized testing data and utilizing it for district report cards, funding, and teacher evaluations we are ruining our public schools and hurting our students.

As touched on earlier by the article written by the Radical Teacher, those working under the education reform are losing steam quickly. Teacher morale is at an all-time low and students are suffering. As a young person in education, I find it disheartening. We are working in the greatest field there is and have the potential to make a huge impact on each and every one of our student’s lives yet federal mandates are dulling the spark in many.

Rick DuFour, an educator for 34 years, now consults with schools, writes solution oriented books, and speaks motivationally to educators around the nation. He says that in working with schools today he is seeing the best educators we have ever had. He observes great practices and instructional methods and finds today’s teachers to be highly effective. However, the sad truth is his opinion is not reflected by those in positions of governmental power.

While we are in a current system in public education which focuses heavily on standardized testing despite the research, there is also reason for hope. As of December 3, 2015 the House of Representatives passed a bipartisan K-12 bill that will significantly shift public education across the nation. This bill will replace No Child Left Behind (2002), a bill that intensified the government’s role in the classroom by holding schools accountable based on state test scores.

While the bill still needs the Senate approval, the house was seen as the largest hurdle. Though the bill is not perfect, it would grossly dismantle the federal accountability system enacted by No Child Left Behind which penalized schools that did not make adequate academic progress based on their standardized test scores. By removing some of the federal government’s role in public education, we then place the power back into trained professional’s hands and allow them to make decisions based on what is best for their students.

Under the proposal, schools would still be required to complete state testing in third through eighth grade in math and reading. However, state governments would determine their methods for intervening with schools that are struggling and making decisions about weight to devote to standardized tests. States would set their own goals and timelines for progress with their public schools. Federal government would not be able to penalize schools for lack of progress or lower funding balances.

Overall, the passing of this bill lowers the emphasis on standardized testing. As schools begin to focus less on this type of testing it is my hope that student achievement will also become a smaller part of the teacher evaluation. While research around the most effective mechanism for evaluating educators is still fairly new, very little supports the use of test scores. Most research cited best practice at this time to be a more balanced approach taking multiple factors into consideration.

While the passing of this bill is not a total overhaul in our current educational system, it is a step in the right direction. As an educator, I am extremely excited by the movement away from a focus on high stakes testing. As the focus for students moves away from testing, teacher will be able to focus less on this is as well and more on teaching the whole child and preparing our students to be citizens first. It is my hope that this bill is the first of many steps in a direction to remove the current accountability system and allow districts to make decisions for themselves.

**REFERENCES**

Antush, J.C. (2014, June). *Labor and Ed Deform: The Degradation of Teachers’ Work*

*through Standardized Testing and the New York City Evaluation System.* Retreived from Summons GVSU Library Database <http://search.proquest.com.ezproxy.gvsu.edu/docview/1537948813/fulltextPDF?accountid=39473>

Casas, F.R., Meaghan, D.E. (2001). *Renewing the Debate Over the Use of*

*Standardized Testing in the Evaluation of Learning and Teaching.* Retrieved from Summons GVSU Library Database [http://download.springer.com.ezproxy.gvsu.edu/static/pdf/4/art%253A10.1023%252FA%253A1011946403091.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1023%2FA%3A1011946403091&token2=exp=1443527759~acl=%2Fstatic%2Fpdf%2F4%2Fart%25253A10.1023%25252FA%25253A1011946403091.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1023%252FA%253A1011946403091\*~hmac=f0f4a21854838859768c160d5fcff538abb703089a6949b34bc63cd513efb1d8](http://download.springer.com.ezproxy.gvsu.edu/static/pdf/4/art%253A10.1023%252FA%253A1011946403091.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1023%2FA%3A1011946403091&token2=exp=1443527759~acl=%2Fstatic%2Fpdf%2F4%2Fart%25253A10.1023%25252FA%25253A1011946403091.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1023%252FA%253A1011946403091*~hmac=f0f4a21854838859768c160d5fcff538abb703089a6949b34bc63cd513efb1d8)

Darling-Hammond, L. (2000). *Teacher Quality: Student Achievement.* Retrieved from

Google Scholar <http://epaa.asu.edu/ojs/index.php/epaa/article/view/392>

Darling-Hammond, L. (2012) *Evaluating teacher Evaluations.*  Retrieved from Google

Scholar [http://www.jstor.org/stable/41497541?seq=1#page\_scan\_tab\_contents](http://www.jstor.org/stable/41497541?seq=1" \l "page_scan_tab_contents)

Deno, S.L. (1985) *Curriculum Based Measurement: The Emerging Alternative.*

Retrieved from Google Scholar <http://ecx.sagepub.com/content/52/3/219.short>

Haertel, Edward. (2009). Journal of Literacy Research: A publication of the Literacy

Research

Association June 1, 2009 41: 151-195. Retrieved from Google Scholar

<http://epa.sagepub.com/content/8/1/45.short>

Lynn, A.R. (2015). *Teacher Evaluations Based on Student Testing: Missing an*

*Opportunity for True Education Reform.* Retrieved from Summons GVSU Library Database <http://search.proquest.com.ezproxy.gvsu.edu/docview/1462775493?pq-origsite=summon>

Michigan Department of Education. (2015). Retrieved from

<http://search.michigan.gov/search?affiliate=mi->&m=true&query=teacher%20evaulation

Radical Teacher. (2012). *Teacher Evaluation and Student Testing.* Retrieved from

Summons GVSU Library Database [http://go.galegroup.com.ezproxy.gvsu.edu/ps/i.do?p=ITOF&u=lom\_gvalleysu&id=GALE|A306753197&v=2.1&it=r&sid=summon&userGroup=lom\_gvalleysu&authCount=1](http://go.galegroup.com.ezproxy.gvsu.edu/ps/i.do?p=ITOF&u=lom_gvalleysu&id=GALE%7CA306753197&v=2.1&it=r&sid=summon&userGroup=lom_gvalleysu&authCount=1)

Sanders, W.L. (1997). *Teacher & Classroom Context: Effects on Student*

*Achievement: Implications for Teacher Evaluation*  Retrieved from Google Scholar <http://link.springer.com/article/10.1023/A:1007999204543#page-1>