THE AFFECT OF STANDARDS BASED GRADING ON STUDENT ACHIEVEMENT

By

KATRINA YOAKUM

Submitted to

Educational Leadership Faculty

Northwest Missouri State University Missouri

Department of Educational Leadership

College of Education and Human Services

Maryville, MO 64468

Submitted in Fulfillment for the Requirements for
61-683 Research Paper

Fall 2013

August 24, 2014
Standards Based Grading 2

ABSTRACT

This study was completed to find if there is a significant different in student achievement on the End of Course Exam (EOCE) due to a school’s use of Standards Based Grading. Pinpointing the exact standards that are necessary for student success should affect how students perform on the state standardized test. Findings of this study show that in one high school, there has been no change in student achievement despite the time and energy devoted in professional development and restructuring the grading system on the EOCE. The reasoning behind this is not determined in this study, but may be attributed to student work ethic and teacher strategies used within the classroom. After compiling and review the findings of this study, current research and literature, and the statistical data from the state, it is found that there is no significant difference in student achievement on the EOCE based on a school’s grading system. Student achievement is similar regardless of the grading system in use.
INTRODUCTION

Background, issues and concerns.

Standards-based grading is a system in which students are given rubrics with specific standards in which they are assessed. Arbitrary means of grading, i.e. extra credit, providing tissue, getting progress reports signed, etc. are ineffective measurement tools because they do not assess what students know or do not know. Standards-based grading focuses on specific student achievement by giving them a variety of formative assessments as a checkpoint for learning. Students will then be given a summative assessment of their learning. The concerns with standards-based grading are that students are not “graded” on turning in homework, so they do not complete homework, which does not provide them with the practice necessary to be successful. Furthermore, in the high school setting, standards-based grading is more rigorous and strict, which affects student GPA and their ability to attain scholarships because their grading scale is different from others in their surrounding area.

Practice under investigation.

Standards based grading has been fully implemented in all core subject areas at a Midwest area high school since the 2011-2012 school year and has branched out to nearly all classes being standards-based for the 2012-2013 school year.

School policy to be informed by study.

All schools within the school district should utilize standards-based grading to improve student achievement.

Conceptual underpinning.

It is thought that students will perform better if they know explicitly what they are to learn and are given the opportunity to practice those skills repeatedly. Standards based grading
allows for students to have a clear set of expectations which will increase their achievement by giving them a clear set of standards in which they are to master throughout the course of a school year or unit.

Statement of the problem.

The problem is to determine if standards-based grading increases student achievement.

Purpose of the study.

The purpose of the study is to determine the effect of standards based grading on student achievement in regards to the End of Course Exam.

Research questions.

RQ#1: Is there a significant difference in student achievement as measured by End of Course Exam scores between standards based grading and traditional grading?

Null hypothesis.

There is no significant difference in student achievement as measured by End of Course Exam scores between standards based grading and traditional grading.

Anticipated benefits of the study.

There has been a lot of debate and resistance to change about the implementation of Standards Based Grading within our school district over the last four years. Our students are concerned that their GPA is being affected negatively and that they are not learning in any more depth than they were on a traditional grading scale. Investigating this in further detail will either prove or disprove the amount of time, energy, and research the teachers, administrators, and district as a whole are putting into the implementation of standards based grading.
Definition of terms.

EOCE- End of Course Exam- test given in the state of Missouri at the end of certain courses created in 2009 (Algebra I, Geometry, English I, English II, Biology, Government)

SBG-Standards Based Grading- involves measuring students' proficiency on well-defined course objectives

AYP- Adequate Yearly Progress- The No Child Left Behind Act, instituted in 2000, sets certain goals for school districts to achieve to show student performance. One factor is test scores on the EOC tests in high school and the MAP test in elementary school.

DESE- Department of Elementary and Secondary Education

Differentiated Instruction- changing instruction to fit needs of different groups of students so every student is able to master the skills and objectives associated with the course objectives.

NCLB-No Child Left Behind-Policy created by George W. Bush and his administration in 2001 to increase academic achievement for all students in the United States

Summary.

The problem with standards-based grading is that it does not affect student achievement. The purpose of the study is to determine the effect of standards based-grading on student achievement in regards to the End of Course Exam. There has been a lot of debate and resistance to change about the implementation of Standards-Based Grading within our school district over the last four years. Our students are concerned that their GPA is being affected negatively and that they are not learning in any more depth than they were on a traditional grading scale. Investigating this in further detail will either prove or disprove the amount of time, energy, and research the teachers, administrators, and district as a whole are putting into the implementation of standards-based grading.
REVIEW OF LITERATURE

In current educational practice, grading and assessment have become the focus in many K-12 classrooms. Formative and summative assessments have taken center stage in classrooms across the nation, which has directly impacted the way teachers think about grading. The grading systems can be broken into two groups: traditional and non-traditional (Marzano, 2010). Traditional grades are those that include participation, late work, extra credit, and other non-academic assignments reflecting life skills. Non-traditional grades, however, are knowledge grades reflecting content knowledge, and local, state, and national standards (Gordon, 2010).

This shift in educational practice relies heavily on student achievement due in part to George W. Bush’s No Child Left Behind Act (NCLB). According to The No Child Left Behind Act of 2001, its purpose is “to ensure that all children have a fair, equal, and significant opportunity to obtain a high-quality education and reach, at a minimum, proficiency on challenging State academic achievement standards and state academic assessments” (U.S. Department of Education, 2001 p. 15). Because of NCLB, states have been encouraged to develop specific learning standards that will meet the needs of students and improve student achievement on state assessments. In the past, American schools have shown low standards and weak incentives making NCLB hold students, staff, and schools more accountable for learning. The education departments for each state have developed content standards for core subject-areas, tests assessing the content to all students, publishing school results, and holding individual students and schools accountable for student achievement. This modification is referred to as standards-based reform (Bishop, 2002).

Although there is no universal definition of standards-based reform (SBR), most discussions of SBR include the following features: academic expectations for students, alignment
of key elements of the educational system to promote those expectations, the use of assessments to monitor performance and student achievement, decentralization of responsibility for decisions relating to curriculum and instruction to schools, support and technical assistance to foster improvement of educational services, and accountability provisions that reward or sanction schools or students on the basis of measured performance (Hamilton, Stecher, & Yuan, 2008).

NCLB was created as a means to hold school districts accountable for increasing student achievement as set by the standards, and by the 2013-2014 school year all students would perform at the proficient or advanced level. For schools achieving this goal, they would achieve what is known as adequate yearly progress (AYP) (U.S Department of Education, 2001). These goals, though unattainable, set high expectations for states, school districts, teachers, and students, which has created discussions on how to improve student achievement in an attempt to show progress and meet AYP. The way in which a state, school district, and students show improvement, is through completion and mastery of the End of Course Exam (EOCE).

According to Bishop, “Since they [EOCEs] assess the content of specific courses, the teacher/s of that course (or course sequence) will inevitably feel responsible for how well their students do on the exam” (2002, p. 11). He goes on to emphasize that the grade received on the EOCE should count as a course grade to become more integrated as part of the expectation or “classroom culture.” He says that by aligning instruction and assessment, accountability is ultimately enhanced. Advocates of this idea agree that it will benefit student achievement as a whole because teachers will set higher standards, creating more attentive students who are more likely to complete demanding work (Bishop, 2002).

However, George W. Bush’s NCLB Act has since been replaced with President Barack Obama’s educational initiative known as Race to the Top, which poses a new set of challenges
for educational officials and affiliates. The U.S. Department of Education was given 5 billion dollars to promote educational reform (Strauss, 2014). The purpose of the Race to the Top Fund is “to encourage and reward States that are creating the conditions for education innovation and reform; achieving significant improvement in student outcomes, including making substantial gains in student achievement, closing achievement gaps, improving high school graduation rates, and ensuring student preparation for success in college and careers” while implementing ambitious plans in four core education reform areas: schools have to evaluate teachers by the increase or decrease of their students’ test scores; increase the number of privately managed charter schools; adopt the common core state standards “turnaround” low-performing schools; and collect data on each individual student (U.S. Department of Education, 2009, p. 2).

Due to the increasing pressure placed upon states, schools, teachers, and students to achieve at higher levels, school districts have taken it upon themselves to reevaluate their current practices. For many schools, this has resulted in a change to grading and assessing. Such schools have taken to shifting away from traditional grading scales in lieu of non-traditional methods that reference student achievement to specific topics within each subject-area. This idea first developed with the conceptualization of a performance standard. Marzano credits Shirley M. Malcom for her definition of a performance standard dating back to 1993, which states, “how good is good enough.” Student performance and achievement have been broken into four categories since then. Those categories are: advanced, proficient, basic, and below basic (Marzano, 2010, p. 17).

According to the U.S. Department of Education, performance standards, sometimes called indicators, define excellence (1996). Setting standards involves defining the "essential" aspects of each subject and writing rigorous standards that speak directly to the concerns of
teachers and parents. Once standards are written, groups including educators and citizens statewide, create a plan to publish, review, and implement them. Many states even develop a strategy tracking each state’s progress toward implementation. The last step in the process is to inform the public about the process and the definitions they used in developing said standards. Standards developers include teachers and community leaders from disadvantaged, remote, or multilingual communities on their committees to create the best standards possible. They also turn to other educational leaders and committees to help determine what a good, quality standard should contain (U.S. Department of Education, 1996).

Educational standards have been around since the early 1990s, but they were not consistent. Each state had its own idea of “proficiency,” which led to the creation of the Common Core State Standards (CCSS) in 2009 (Common Core State Standard Initiative). These standards are part of Obama’s Race to the Top initiative and released to the public in 2010. Almost immediately, 45 of the 50 states, including the District of Columbia, had signed on to join Obama’s Race to the Top in lieu of NCLB (Strauss, 2014). The CCSS were developed and completed in two categories. The first category is College and Career Readiness, which dictates what students are supposed to know upon graduation from high school. The second category is the K-12 standards which break down the expectations of students by grade level and subject area (Common Core State Standard Initiative).

It is these standards that school districts have begun to implement and align their daily curriculum to. It is also these standards, which have aided in the development of the non-traditional grading system, known as Standards Based Grading. One high school breaks down what SBG is on their school website stating, “[Standards Based Grading] measures your student’s mastery of the essential standards for a class. At the beginning of every unit, the
teacher breaks down the standards for the unit into smaller objectives and criteria using a detailed rubric. During the unit, the student is assessed to see if they truly know the material using a variety of assessments, such as traditional pencil-and-paper tests, projects, discussions, or reports. The class grade will be based on all of the evidence the teacher collects demonstrating mastery of the essential standards” (Standards-Based Grading Frequently Asked Questions, pg. 1). Schools are breaking down each content area based on the CCSS, creating rubrics to assess their teaching/student learning, and giving frequent assessments to check student progress. All of this is being done in order to prove student proficiency on the state’s standardized test.

There has been research done that shows student achievement will improve based on focusing on student standards. For example, after *A Nation at Risk* was published, an attempt to improve academic achievement was made by making graduation requirements more rigorous, which was the first wave of reform, but it didn't have much of an effect. Therefore, the standards movement holds the greatest hope (Scherer, 2001).

Scrifiny (2008), states in her article “Seven Reasons for Standards-Based Grading” the differences made in the ways student earn letter grades. She said that when she was first asked to determine the difference between an A, B, C, D, or F, that she could not come up with major differences between them, and if she did, they were subtle. She then turned to SBG as a means to differentiate and communicate to students and parents their level of learning and understanding. “An *A* means the student has completed proficient work on all course objectives and advanced work on some objectives. A *B* means the student has completed proficient work on all course objectives. A *C* means the student has completed proficient work on the most important objectives, although not on all objectives. The student can continue to the next course. A *D* means the student has completed proficient work on at least one-half of the course objectives but
is missing some important objectives and is at significant risk of failing the next course in the sequence. The student should repeat the course if it is a prerequisite for another course. An $F$ means the student has completed proficient work on fewer than one-half of the course objectives and cannot successfully complete the next course in sequence” (p. 70). This grading scale would prove students getting an A or B in a given class would be able to score at the proficient range on the EOCE, which makes the case for SBG improving student achievement.
RESEARCH METHODS

Research design.

A quantitative study was conducted to see if there was an increase in achievement on the EOCE due to implementation of Standards Based Grading. The independent variables being tested were scores on the EOCE prior to SBG and after its implementation, while the dependent variable tested were the EOCE scores. If the difference is found significant in scores based on the grading system, teachers and administrators will know the grading system has no bearing on student achievement.

Study group description.

The students in the study attend a Midwest high school, which has students in grades 9-12 and has approximately 819 students. 37.6% of students qualify for free and reduced lunch. 88.5% of students are white, 4.6% are Hispanic, 1.6% are black, .9% are Asian, and .4% are Indian.

Data collection and instrumentation.

Archived data from DESE was collected to identify raw scores of students on the EOCE in Communication Arts and Math from the 2008-2009, 2009-2010, 2011-2012, and 2012-2013 school years.

Statistical analysis methods.

A t-test was conducted to find if there is a significant difference in student achievement on the EOCE because of Standards Based Grading. The source was broken into the following categories: before Standards Based Grading and after full implementation of Standards Based Grading, which is why data from the 2010-2011 school year are omitted. This was a transitional year for the district. The mean, mean D, t-test, df, and p-value were concluded from this test.
The Alpha level was set at 0.25 to test the null hypothesis: There is no difference in student achievement on the EOCE scores between the use of Standards Based Grading and traditional grading.
FINDINGS

A t-test was conducted to decipher whether there was a difference in performance on the Communication Arts EOC test between 2009-2010 when Standards Based Grading was not being used and between 2012-2013 after the grading system had been implemented. The following tables, graphs, and charts will depict the organized findings based on the statistical raw data found on the Missouri DESE website in 2014.

Figure 1

t-Test Analysis Results for 2009-2010 and 2012-2013 Communication Arts End of Course Exam Scores Before and After Standards Based Grading Implementation

<table>
<thead>
<tr>
<th>Source</th>
<th>Mean</th>
<th>Mean D</th>
<th>t-Test</th>
<th>df</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before SBG (n=25)</td>
<td>71.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After SBG (n=25)</td>
<td>65.70</td>
<td>6.15</td>
<td>1.87</td>
<td>2</td>
<td>.101</td>
</tr>
</tbody>
</table>

Note: Significant when p<=0.25

One school was selected to look at their individual EOC scores before they began SBG and after it was put in place. The data collected from the Missouri DESE website contains the percentage of students that scored in the top two tiers, proficient and advanced, of the grading scale. The mean of the student scores before Standards based grading was 71.85 and the mean of student scores after its enactment was 65.70. The Mean D, or difference between the two groups, was 6.15. The t-test result was 1.87 and the df was 2. The null hypothesis states that there is not a significant difference in Missouri End of Course exam scores based on the difference in the grading system used. This null hypothesis was rejected because the p-value, .101, is lower than the alpha level, 0.25. This shows that the grading system used in schools
does affect student achievement on standardized tests. In the subject-area of Communication
Arts, the traditional grading system was more effective regarding student achievement.

Figure 2

The mean of the students’ Communication Arts End of Course proficient and advanced
test scores while during a traditional grading scale in the years of 2009 and 2010 were 71.85.
This means that slightly more than 70% of students in the state of Missouri scored in the top two
tiers of the scoring rubric for this standardized test. The pie chart also shows that the number of
students reaching the top two tiers in 2012 and 2013 after the new grading scale was put into
place was lower with a mean of 65.70.
There are four categories in which a student can score on the End of Course exam based on the number of questions the student answered correctly: below basic, basic, proficient, and advanced. The bottom two categories that do not meet the state requirement for achievement are below basic and basic. The top two tiers students attempt to reach are proficient and advanced, which meet the state requirements for success. This chart shows how students performed in those top two tiers before standards based grading and after. 71.85% of students scored in the top two tiers, meeting the state standards utilizing a traditional grading scale. 65.70% of students scored in the top two tiers, meeting the state requirements, after SBG was introduced.
Figure 4

**t-Test Analysis Results for 2009-2010 and 2012-2013 Communication Arts End of Course Exam Scores Before and After Standards Based Grading Implementation**

<table>
<thead>
<tr>
<th>Source</th>
<th>Mean</th>
<th>Mean D</th>
<th>t-Test</th>
<th>df</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before SBG (n=25)</td>
<td>36.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After SBG (n=25)</td>
<td>53.45</td>
<td>17.15</td>
<td>13.9</td>
<td>2</td>
<td>.0051</td>
</tr>
</tbody>
</table>

Note: Significant when p≤0.25

One school was selected to look at their individual EOC scores before they began SBG and after it was put in place. The data collected from the Missouri DESE website contains the percentage of students that scored in the top two tiers, proficient and advanced, of the grading scale on the Math EOCE. The mean of the student scores before Standards based grading was 36.30 and the mean of student scores after its enactment was 53.45. The Mean D, or difference between the two groups, was 17.15. The t-test result was 13.9 and the df was 2. The null hypothesis states that there is not a significant difference in Missouri End of Course exam scores based on the difference in the grading system used. This null hypothesis was rejected because the p-value, .0051, is lower than the alpha level, 0.25. This shows that the grading system used in schools does affect student achievement on standardized tests. In the subject-area of Math, the standards based system of grading was more effective regarding student achievement.
The mean of the students’ Communication Arts End of Course proficient and advanced test scores while during a traditional grading scale in the years of 2009 and 2010 were 36.30. This means that a little more than 35% of students in the state of Missouri scored in the top two tiers of the scoring rubric for this standardized test. The pie chart also shows that the number of students reaching the top two tiers in 2012 and 2013 after the new grading scale was put into place was higher with a mean of 53.45.
There are four categories in which a student can score on the End of Course exam based on the number of questions the student answered correctly: below basic, basic, proficient, and advanced. The bottom two categories that do not meet the state requirement for achievement are below basic and basic. The top two tiers students attempt to reach are proficient and advanced, which meet the state requirements for success. This chart shows how students performed in those top two tiers before standards based grading and after. 36.30% of students scored in the top two tiers, meeting the state standards utilizing a traditional grading scale. 53.45% of students scored in the top two tiers, meeting the state requirements, after SBG was introduced.
The End of Course exam was introduced in 2009 for English 2 at the high school level. Results from the Missouri DESE website show that in 2009, 72.6% of students met the state standards scoring in the proficient and advanced range on the communication arts assessment. In 2010, the scores slightly decreased to 71.1% meeting this goal. Then in 2012, the scores continued to drop to a score of 68.9% before reaching an all-time low score in 2013 of 62.5% of students scoring in the proficient and advanced range.

The End of Course exam was also introduced in 2009 for students taking Algebra I in high school. Results from the Missouri DESE website show that in 2009, 37.4% of students met the state standards scoring in the proficient and advanced range on the math portion of the assessment. In 2010, the scores decreased to 35.2% of students scoring proficient and advanced. Then, in 2012 the scores drastically increased 54.0% to before leveling off and slightly decreasing to 52.9% in 2013.
All of the findings compiled answered the research question: “Is there a significant difference in student achievement as measured by End of Course Exam scores between standards-based grading and traditional grading?” Figures 1-3 reported there was a significant difference in performance using traditional grading versus that of standards based grading on the Communication Arts End of Course exam. The data gathered shows that while the scores were fairly consistent, students performed better under the traditional grading system as opposed to the new standards based system. Figures 4-6 concluded again that there was a significant difference in using a traditional grading scale versus that of a standards based one using the math End of Course exam scores. This particular data shows that students performed at a higher rate after standards based grading was implemented. Figure 7 shows the trends with the Communication Arts and Math End of Course exams before and after standards based grading since the implementation of the assessment. In 2009, the first year the test was proctored students performed their best on the communication arts assessment scoring at 72.6% proficient and advanced. In 2010, the scores slightly decreased to 71.1% of students meeting this goal. Then in 2012, when standards based grading was fully implemented, the scores continued to drop to a score of 68.9% before reaching an all-time low score in 2013 of 62.5% of students scoring in the proficient and advanced range. However, in 2009, 37.4% of students met the state standards scoring in the proficient and advanced range on the math portion of the assessment. In 2010, the scores decreased to 35.2 % of students scoring proficient and advanced. Then, in 2012, after the implementation of standards based grading, the scores drastically increased to 54.0% before leveling off and slightly decreasing to 52.9% of students scoring in the proficient and advanced ranges in 2013.
CONCLUSIONS AND RECOMMENDATIONS

The outcomes reported from this study show that students are currently performing better on the Communication Arts End of Course exam under the traditional grading system while the findings show that students performed better on the Math End of Course exam under the new grading system. Regardless, the data shows that there is a significant difference in student achievement based on the grading system being utilized. The t-test result was 1.87 and the p-value was .101 on the communication arts test. The t-test result for math was 13.9 and the p-value was .0051. Both are lower than the alpha level that was set at 0.25, therefore, the null hypothesis tested is rejected with confidence. This shows that the grading system used in schools does affect student achievement on standardized tests.

The conceptual underpinning that students will perform better if they know explicitly what they are to learn and are given the opportunity to practice those skills repeatedly is supported using the math data, yet is rejected based on the data collected using the communication arts scores. While standards based grading allows for students to have a clear set of expectations which will increase their achievement by giving them a clear set of standards in which they are to master throughout the course of a school year or unit is seen beneficial in math classes, it is unclear why this did not prove true of the communication arts test scores and further examination is needed to pinpoint the reasoning for this.

While math tends to be more linear and makes sense to have set topics to be assessed, it is not clear whether or not communication arts is as clear cut. More studies may need to be conducted included the brevity in which content is covered in communication arts since the assessment covers many aspects including reading, writing, grammar, and usage. Furthermore, studies may need to be conducted to see which of the above aspects are being assessed. Also,
communication arts tends to lend itself more to text-to-world and text-to-self applications in which students make inferences about their surroundings and their ability to relate and live within that world using literature as a medium to translate those ideas. These are not tangible, clear ideas that can be assessed, but focus more on the individual rather than a certain skill base whereas math is skill specific.

Questions have also formulated about how the other subject areas have performed on the EOCE before and after the implementation of SBG. Is there a similar correlation between the results determined on the math t-test and what would be found using the science scores? Given the similarities between math and science, one may assume that, but further assessment would be necessary before forming such an opinion. The same question could be asked about the correlation between the communication arts scores and those of social studies. All scores should be analyzed and interpreted for trends to get to the heart of how students can best achieve and meet the standards within their school.

Professional development needs to occur across the state to ensure teachers are aware of best practices for teaching the given standards and what is most effective for student productivity. While there are many researchers who support standards based grading, the evidence is clear that it is not a full-proof plan to ensure student success. Changes and adaptations need to be made based on the students within the building, individual learning styles, and the types of questions asked on the assessment being used to measure student performance. Teachers need time to research and practice methods that would be best for their students and be allowed to use their professional judgment for the betterment of all.
REFERENCES

Bishop, John H. (2002). What should be the federal role in supporting and shaping development of state accountability systems for secondary school achievement?.


