Summer 7-2018

An Investigation of the Relationship Between Standards-Based Grading and End of Course Assessment Scores

Steven W. Smith
Lindenwood University

An Investigation of the Relationship Between Standards-Based Grading and End of Course Assessment Scores

by

Steven W. Smith

A Dissertation submitted to the Education Faculty of Lindenwood University

in partial fulfillment of the requirements for the

degree of

Doctor of Education

School of Education

An Investigation of the Relationship Between

Standards-Based Grading and End of Course Assessment Scores

by

Steven W. Smith

This dissertation has been approved in partial fulfillment of the requirements for the degree of

Doctor of Education

at Lindenwood University by the School of Education

Dr. John Long, Dissertation Chair	7.18.18 Date
Dr. Kevin Winslow, Committee Member	7/13/18 Date
Dr. James Allison, Committee Member	7/13/2018 Date

Declaration of Originality

I do hereby declare and attest to the fact that this is an original study based solely upon my own scholarly work here at Lindenwood University and that I have not submitted it for any other college or university course or degree here or elsewhere.

Full Legal Name: Steven William Smith

Signature:  Date: 7-13-18

Acknowledgements

I would like to thank the professors of the Doctoral Program in the School of Education at Lindenwood University for their support in helping me get through all of the coursework required for this degree. I would like to thank my committee members, Dr. Kevin Winslow and Dr. James Allison, for taking the time to critique my dissertation and provide needed feedback; and I would like to single out Dr. John Long, without whom I would not have made it through the entirety of the process. Dr. Long provided timely feedback and motivation throughout the course of the study, and for that I thank him. I would also like to thank the City of St. Charles School District for allowing me to utilize their quantitative and qualitative data for my study. Lastly, I would like to thank my wife, Amie, and my kids, who provided me with the support and motivation necessary to complete this gargantuan task. Without that support and their understanding, I would not have been able to finish this.

Abstract

The purpose of this study was to determine if there was a correlation between the implementation of standards-based grading (SBG) and a rise in student achievement as indicated by the evaluation of end-of-course exam data from the four core subject areas in secondary schools - English, Algebra, Government, and Biology. This mixed methods study focused on the collected data of 6,000 test scores, split evenly from tests taken prior to the launch of SBG and after the launch of SBG, as well as focusing on varying perceptions of SBG from both teachers and members of the community as a whole. Quantitative data consisted of test scores aggregated by the Missouri Department of Elementary and Secondary Education (MODESE) and distributed to school districts. Qualitative data were secondary in nature and taken from two separate surveys administered by the Cooperating School District to teachers and parents concerning how they felt about the implementation of SBG. These data were aggregated and analyzed by using coding techniques for qualitative data to determine the study outcomes. The quantitative data indicated that there was no statistically significant increase in test scores over the five years researched for this study. The qualitative data indicated that perceptions of SBG were frequently negative in nature, although there were varying degrees of negativity. This result came from both the teacher and parent responses. According to these qualitative data, parents and teachers alike viewed SBG as diminishing the foundation of the education the students were receiving, while at the same time inadequately preparing them for the post-secondary world in that too many chances were given for them to succeed under SBG.

Table of Contents

List of Tables

Chapter One

Background

As a former high school social studies teacher, with a primary focus on teaching government (Civics), the researcher was able to see how students best learned and assessed themselves over an extended period of time. Part of this observation was centered on the Missouri End-of-Course Exam, or EOC, that was given in Civics. When looking at those EOC scores each semester, the researcher would consistently evaluate the course to see what could be done to improve instruction going forward, in an effort to improve the EOC scores and improve student learning and comprehension. As this same issue was going on in other EOC-assessed classes (Biology, English, and Math), it subsequently became a district issue, which turned into a district initiative. The initiative was called Standards-Based Grading, or SBG.

SBG was, in the simplest of terms, a new way of grading. The traditional way, which focused on an 'A' through 'F' scale, had been in place for quite some time in schools across the nation, and in many instances was still in place, at the time of this writing, because no one had put much thought into a better option. The traditional model focused primarily on a total-points system where assignments, quizzes, tests, and extra credit were added up, and divided by the total points to give a certain percentage. That percentage then correlated to an arbitrary grade, with no real indication of what exactly the student learned or had not learned. Rather, it just indicated that a particular student had passed a class. This traditional model also did not seem to take into account progress made by the student over the course of the semester. For example, if a student did poorly at the beginning of the semester, but then was able to improve as the semester went on,

that student's total points earned would still be divided by the total points possible, and the end result would be a lower grade, due to the deficiencies at the beginning of the semester. In addition, with the traditional model of grading, many subjective factors could be factored into a student's grade that ultimately played no role in indicating what that child learned over the course of the semester or year, as the case may be. These subjective factors included, but were not limited to, the following: attendance, effort, extra credit, and attitude. These subjective factors could then end up influencing a student's grade positively or negatively.

The theory behind SBG was that it removed the barriers found in the traditional method, which allowed for the development of student learning and higher levels of retention, as well as painting a more accurate picture for the students, parents, and teachers as to what the student had actually learned or not learned. SBG, on the other hand, graded on scales of one through four, with a four being the most advanced and a one indicating that a student needed to relearn that particular standard. The overarching goal of SBG was to, "Better communicate what each student knows and is able to do according to district and state content standards and separately assess the influence of positive and consistent work habits on student learning" (Heflebower, Hoegh, & Warrick, 2014, p. 95). SBG also potentially eliminated the need for those subjective aspects of grading previously mentioned in an effort to keep a student's academic success focused primarily on academics and not extraneous measures.

Problem

The problem was determining, beyond theory, whether SBG actually worked better than the traditional grading scale, and if so, to what extent.

Purpose of the Study

The purpose of this study was to investigate the possible relationship between the implementation of SBG and the achievement measured by data on EOCs in four core subject areas: Civics, Biology, Algebra I, and English I in a Midwest secondary setting. Five years of data, 2011 through 2016, were collected for the study. However, for purposes of analysis, the scores of the last year of non-SBG grading from 2013 were compared to the last year of available data in 2016. The data in 2016 was after two full years of SBG implementation. The idea behind SBG was rooted in the assumption that the traditional model of grading, otherwise known as the A-F scale, was an antiquated system of grading that did not accurately portray what a student learned and retained. This assumption led to the development of a system of grading called Standards-Based Grading (SBG) that allowed students and parents alike to see the areas a child might be struggling in and work within those specific areas to help that student improve. SBG, in theory, would also take out the differences between teachers and like classes, so that what a student learned in one class was the same that a student would learn in another class, and would be assessed in the same way, affording different students in different classrooms the same opportunities regardless of instructor, and subsequently 'leveling the playing field.'

Hypothesis and Research Questions

This was a mixed methods study, focusing primarily on the quantitative aspect, but also including the qualitative aspect to a degree. The hypotheses for the quantitative aspect of the study were as follows:

H1: The implementation of standards-based grading will lead to an increase in student achievement as measured by an increase in end-of-course test scores in four core subject areas between 2011 and 2016.

The research questions for the qualitative aspect of the study were as follows:

RQ1: How did the implementation of standards-based grading at the secondary level impact the perceptions of learning from the school district community?

RQ2: How did the implementation of standards-based grading at the secondary level impact how secondary staff felt about the change in grading systems?

Importance of the Study

The then-current gap in knowledge relating to SBG was that there did not seem to be any research data that indicated how effective SBG was in increasing the levels of student achievement, particularly at the secondary level. There was research that showed how to implement SBG at the secondary level, as Scriffiny (2008) focused her research on why grades should have meaning at the secondary level, particularly as students were readying themselves for a post-high school career. That being said, the large gap was showing data that proved that the implementation of SBG not only provided a more structured way of grading, but that the implementation of SBG actually did lead to an increase in student achievement. Most of the research as it related to the impact of SBG on student achievement was centered on the elementary population and not the secondary population of students.

Definition of Terms

Common assessments – were created by a team of like teachers (whether same school, or district-wide departments) to ensure that all students in all similar classes were assessed on the same course materials and in the same fashion (Heflebower et al., 2014).

Formative assessments – were used to periodically monitor and adjust instruction to improve learning for students. These were not counted for credit in a class, but rather as practice. Examples of these were homework, pre-tests, and observations. Teachers could provide helpful feedback on these (Heflebower et al., 2014, p. 34).

Proficiency scales – were scales that were standardized at the district level for each course offered and were broadly written. Each scale indicated a separate standard that the student must pass in order to pass the class. Each scale was based on a numerical system of one through four, with a four being advanced, a three being proficient, a two being basic, and a one being below basic. A score of one, indicating that not even basic learning was accomplished, led to a failing grade for that scale. These were what students needed to know at the end of the class (Heflebower et al., 2014).

Standard – statements that defined what students should, "Know, understand, and be able to do for each course" (Heflebower et al., 2014, p. 96).

Standards-Based Grading (SBG) – this method of grading helped to improve student achievement by focusing on the following four critical questions: "What do students need to know and be able to do; how will we know that they have learned it; what will we do when they haven't learned it; what will we do when they already know it?" (Heflebower et al., 2014, p. 96).

Summative assessment – otherwise known as assessment of learning, these were the summary tests of each standard that determined what a student knew and would solely determine his/her grade in the class. These were more formal in nature, as compared to formative assessments, and were usually given toward the end of any given unit. Examples include end-of-unit exams, research papers, and final exams (Heflebower et al., 2014).

Limitations

This study was based solely on the fluctuations in EOC scores in four core areas over a five-year timeframe. Limiting the study to just these scores eliminated some other data points that could have been utilized, such as ACT scores over the same timeframe. This study did not take into account significant demographic shifts in student population that happened over the five-year span, including race and student population. Although the data collected for this study were all secondary data collected by the Missouri Department of Elementary and Secondary Education (DESE, 2017) and the City of St. Charles School District, therefore being valid data, the interpretation of that data by the researcher could potentially be a limitation, given the researcher's inexperience in conducting these types of studies.

In addition, this study did not take into account qualitative studies about the traditional grading system in the studied district. Therefore, one would have to make assumptions on whether the traditional way of grading was something that stakeholders and teachers wanted changed to begin with. Further, the study did not take into account what those shortcomings of the traditional grading system would be if a survey on that grading system were presented to the community and teachers with input back from them

on what they believed were the solutions to the shortcomings of the traditional way of grading. Had this been conducted, it could potentially serve to integrate better with the research done on standards-based grading and help better find an amicable solution to the problems with both the traditional model and the SBG model of grading.

This study also only took into account the transition years, those years in which the district was transitioning from the traditional model to the SBG model at the secondary level. The problem with this was that this was the time period when everyone was trying to figure out the new method, implement the new method, work with colleagues to come up with best practices for the new method, and then develop their classes essentially from scratch. During this period, many questions were going to be asked, and there may not have been an equal number of answers, leaving many to work the old method of grading into the new method simply as a means to be compliant with the directive. As a result, during this transition time not much may have actually changed within the classroom itself. So, a limitation was the timing of the study. Perhaps had the study been done after the dust had settled, the quantitative and qualitative aspects would have been different and shown a larger percentage of people advocating for the change, as well as potentially showing a statistically significant increase in student achievement in those classes having implemented standards-based grading.

Assumptions

This study assumed that all EOC data collected by the Missouri Department of Elementary and Secondary Education in the core subject areas of Civics, Algebra, English, and Biology were accurate and were reported as such. The study also assumed that the survey provided to the staff and parents of the City of St. Charles School District

was answered in an honest manner by all involved, to the best of their ability in an effort to give honest feedback about SBG**.**

Summary

The researcher, a former secondary government teacher and then-current secondary administrator, conducted this study to investigate the relationship between standards-based grading and student achievement, focusing primarily on end of course exams as the quantitative aspect of the study and community/parent surveys as the qualitative aspect of the study. As the initial reactions from both teachers and community members were decidedly negative as it related to the implementation of SBG, this researcher wanted to determine whether their reservations with SBG were rooted in fact or fiction and explain why.

Chapter Two

History of Grading and Standards-Based Grading

According to Brookhart et al. (2016), grading "refers to the symbols assigned to individual pieces of student work or to composite measures of student performance on student report cards" (p. 804). This simple definition of grading left many educators to wonder, over the last 100+ years previous to this writing, how to best assess students while still leaving the teachers with complete autonomy over how they assess within their own classrooms, which had the potential to leave room for more ambiguous rationale over why a grade was the way it was, while at the same time leaving much room for inconsistencies in grading even amongst like classes within the same building. Within this context, some of the earliest educational research originated with grading as the basis for the research, in large part to help determine its purpose and to help establish consistency (Brookhart et al., 2016, p. 804). According to Brookhart et al. (2016), the study of grading was deemed important, primarily due to how central grades were within the context of the educational history of all students. These grades were largely to be viewed as what a particular child had earned in a class and/or for their achievement (Brookhart, 1993, p. 139) and had a significant impact and influence on schools and the students within them (Pattison, Grodsky, & Mueller, 2013). Further, as stated by Bowers (2010), grades also ended up as predictors of "important future educational consequences, such as dropping out of school" (Bowers, 2010, p. 3), and "applying and being admitted to college, and college success" (Atkinson & Geiser, 2009, p. 668). Further, Sawyer (2013) went on to state that grades had always been a "predictor of academic success in more open admissions higher education institutions" (p. 93).

Research on grading could be dated back to 1888, when one of the first investigations occurred. This research was largely conducted to try to determine how reliable grades actually were, as they were being assigned to students (Brookhart et al., 2016). Edgeworth published his research in the *Journal of the Royal Statistical Society* in 1888 and went on to identify three varying areas of error normally found in educator grading techniques; one being chance, the second described as "personal differences among graders regarding the whole exam and individual items on the exam" (1888, p. 600), and the last being that teachers were prone to accept the answers of the student as representative of his or her proficiency (Edgeworth, 1888). Brookhart et al. (2016) found that as far back as 1888, Edgeworth (1888) had been able to essentially point to the "educational consequences of unreliability in grading, especially in awarding diplomas, honors, and other qualifications to students" (p. 805) and was subsequently able to use this research to show that there was a need for greater reliability in our schools. In terms of grade reporting, through much of the 19th century, report cards and progress reports were given orally with very little standardization as it related to content. This changed toward the end of the 19th century, only to focus more on written descriptions of how students were doing based on areas that, at the time of this writing, would be considered soft skills, such as penmanship and reading (Guskey & Bailey, 2001, p. 4). As American society moved into the 20th century and secondary schools became more and more diverse, there appeared to be a need for schools to find a way to manage this more effectively; and so, while elementary schools largely continued on with the narrative descriptions of student performance, high schools sought a more objective approach as the narratives proved time consuming and inefficient. It was within this context of

American history that high schools started to head toward a grading system that favored percentages; effectively, one could argue, "Eliminating the specific communication of what students knew and could do" (Guskey & Bailey, 2001, p. 203).

During the course of the past century, not much had changed in terms of how students were graded in their coursework. According to Stephens (2010), many of our attitudes and opinions toward grading were ingrained with us in large part due to the ways we were graded when we were students, or what our personal beliefs were, or by the programs we enrolled in as undergraduates. Over the years, many educators had taken the traditional model of grading and simply combined what parts they liked, and along the way, may have taken part of standards-based grading and integrated it without realizing that they were doing that (Shippy, Washer, & Perrin, 2013). The grading system used in most schools as, of this writing, was a simple system based on the A through F model and was originally developed as a means to provide a mechanism that promoted children from one grade to the next, while not necessarily being reflective of student learning (Spencer, 2012). The traditional model of grading was a very flawed system and one in which almost anything could be used to help determine a student's grade (Hooper & Cowell, 2014). What this meant was that grades could be based on subjective things, such as attendance, effort, and formative assessments, etc. The list could go on and on and was only limited by the imagination of the teacher. This model evolved into a "sorting machine that allowed educators to rank students and establish classroom curves and hierarchies" (Spencer, 2012, p. 5). No matter the foundation that the original form of grading took on, be it grading on the curve, pass/fail models, or weighted grades, they all seemed to rely on the basic principle of adding up scores and

calculating a subsequent grade. This seemed contrary to what the objective of education should be, which was learning (O'Connor, 2007). According to O'Connor (2007), grading, "traditionally promoted a culture of point accumulation, not learning" and, "encouraged competition rather than collaboration" (p. 127). Secondary classroom instructor and grading researcher Tucker (2018) found that she had been "logging in more than 100 assignments in her online gradebook each semester and that her students had received points for everything from completing assignments to bringing in materials" (p. 84), which subsequently made her grades more a reflection of the ability of the student to organize rather than what their skill level actually was at the end of each unit or summative. Tucker (2018) found that she felt "uneasy . . . that some students were receiving As and not demonstrating proficiency while some students were receiving Cs simply because of missing assignments even though they had demonstrated mastery of the topic" (2018, p. 84). Kohn (2011) spoke to the deficits of this system by noting the following: 1) grades lessened the interest of the student, 2) grades were inherently going to create a preference for the easiest possible task, and 3) grades tended to reduce the quality of students' thinking. Essentially, grades became more important than the learning that was supposed to be taking place. In her experiences as an instructor, Tucker (2018) also developed three truths about learning that she had not realized prior to her transition to SBG, noting: 1) students needed different amounts of practice to master a skill, 2) grades should be an accurate reflection of a student's current ability, and 3) learning was ongoing (p. 85). O'Connor (2007) discussed several of what he considered to be the basic, general beliefs of grading as a whole, with a focus on the idea that grading was: a) complicated, and not just about adding numbers together, and b)

subjective (p. 129). The subjectivity of grading seemed to be, in the author's experience, one of the single biggest deficits of the traditional grading model in that it allowed too many choices for teachers. Intrinsic in the subjectivity of grading was bias, which O'Connor (2007) also discussed as a reason for the faults within the traditional model of grading (p. 129). This bias, or excessive autonomy, can potentially lead to classes that, in context and content, should be the exact same class taught by two different teachers. The reality of this, however, was that two different teachers could teach the same class and come up with completely different student outcomes, due to the way they set up their respective point totals and weights, not to mention extra credit opportunities. Because of this, educators and researchers started turning to standards-based grading as early as 1973, when Hirst (1973) explained his standards-based grading (SBG) approach ahead of his time in the following way:

> Its greatest promise seems to be its ability to place the student in a meaningful relationship with the domain of knowledge so that his progress can be measured based entirely on his individual efforts. What is more human than letting the learner know, in advance, what he is expected to know, under what conditions he will be expected to know, under what conditions he will be expected to demonstrate his knowledge, and the level or degree of acceptable performance he is to achieve. (pp. 50-51)

Stephens (2010) argued the reason educators had not changed over the years in terms of the grading system was that the model was rooted in a system used when they were students themselves, and it was difficult to break that cycle, particularly as it seemed to be reinforced by undergraduate degree programs throughout the country.

Having set standards that encouraged continuity and student growth, and took the focus off of individual point accumulation, these standards, according to Guskey (2009), represented the "goals of teaching and learning" and "Describe[d] what we want students to know and be able to do as a result of their experiences in school" (p. 11). It was in this context that SBG differentiated itself from the traditional grading model – as it set standards that were clear and attainable, and provided continuity between classes. Standards also allowed for students to know what exactly was expected of them.

Why Standards-Based Grading?

The information in the previous sections suggested a need to move away from the traditional model of grading and into a model that assessed what a student knew, but the path to across-the-board implementation of standards-based grading had been difficult, as it overcame several challenges along the way. Teachers reported difficulty completing the standards-based report card, due to its increased complexity, and parents resisted as well, noting that students seemed to lack a motivation to study when they knew they would be retested (Spencer, 2012). Parents indicated concerns with nonacademic factors, such as attendance and extra credit no longer being a component of the student's final score (Spencer, 2012). Spencer (2012) also made clear that many parents found the new report cards, which were broken down by standards in each class, overwhelming and difficult to read, preferring an easier to read summary of what their child's grade was.

That being said, Marzano and Heflebower (2011) contended that at the end of the day, awarding a straight letter grade to show how well a student understood a concept provided very little in the way of useful information about that student and his or her performance in the class. The contention here was that a child might have received an

overall letter grade of a B+ in a particular class only because he or she was well behaved in class, participated, and turned in all of his or her work on time, while another student might have received a D+, not because he or she did poorly on the summative tests, but instead did not turn in homework and did not participate in class discussions (Marzano & Heflebower, 2011). Essentially, grades needed to be meaningful, and under the traditional model it was generally impossible to tell how meaningful a grade was, in large part due to the difficulty in interpreting a particular grade. Grades, as they stood under that traditional model, did not necessarily provide valid and reliable grades. Under SBG, the goal would be to head more toward a system where grades were reliable, valid, and equitable across the board. Munoz and Guskey (2015) stated that, at the time, teachers reached a final grade in a class by simply merging various scores from exams, homework, papers, quizzes, reports, homework, evidence of punctuality, class participation, and effort, etc. The result of this was that teachers then turned to their online gradebooks, where they were able to set up a number of weights for each of their courses, specific to their courses, and there had been no guidance or professional development to support that teacher in how they came up with those categories, and what meaning the categories had. Therefore, the end grade ended up being something that was "impossible to interpret accurately or meaningfully" (Munoz & Guskey, 2015, p. 64). A supporter of standards-based grading would also contend that with the traditional grading model, what one teacher might determine to be an A in one class might be a C in that same class with another teacher, because different criteria and point situations were being utilized differently in their respective classrooms (Marzano & Heflebower, 2011). The single biggest factor when critiquing the traditional model of grading was that educators

did not have any idea as to what was actually measured to get that final letter grade in any particular class. While teachers tended to agree that one of those factors should be achievement, often times one found that grades were a total accumulation of everything from homework, to effort, to participation, to attitude, all of which seemed very subjective in nature. In terms of then communicating this grade with students and parents, the topic became very ambiguous, as it did not indicate what a student "knows and is able to do" (Hooper & Cowell, 2014, p. 60). Couple this with weighted scales, meaning each one of the categories could be weighted differently by teachers even in the same department at the same school, and one would have a cornucopia of potential student outcomes, none of which seemingly represented what he or she actually learned in the class (O'Connor, 1999, p. 135).

In rare cases, the implementation of standards-based grading even had the potential to hold up in court better than a traditional model would have done. A case in Arizona from May, 2002, showcased as much. As graduation was approaching, an English teacher at West Valley High School in Peoria, Arizona, Elizabeth Joice, received a letter from an attorney representing one of her students who was going to fail the course, subsequently denying this student the ability to graduate with her peers. As a result, according to the letter, the student had been "very sick, unable to sleep or eat, and has been forced to seek medical attention" (as cited in Taylor, 2003, pp. 58) and went on to state that the student had:

> Experienced severe emotional and physical distress over this…and that we know a student's grade or total score in a given class is based on a variety of factors, only some of which are test scores, which leads to the thought that there is a

> question of subjective vs objective grading, and that all information regarding your background, your employment records, all of your class records, past and present, dealing with this and other students become relevant, should litigation be necessary. We ask at this time that you take whatever action is necessary to correct this situation so that it can be settled amicably. Failing that, you will force us to institute litigation. (Taylor, 2003, pp. 58-59)

The teacher, Mrs. Joice, responded that the student had been given every opportunity to pass the class, but had chosen to not take advantage of any of the opportunities provided for her, and was regularly absent from the class. Mrs. Joice went on to add that the student had been caught plagiarizing and had low test scores throughout the course of the year, and that notices of these failures had been sent home throughout the year. She even spoke to the father during the course of the year to update him on his child. The result of all of this was that five hours prior her graduation, the student was allowed to retake a 50-question test she had previously failed and was able to then walk with her class (Taylor, 2003).

This case study proved relevant to this study in that it highlighted the need for accountability in grading as a means to avoid situations like these going forward. In this case, the lawyer for the student made the case by bringing forth the potential doubt of the subjective nature of grading, which was a policy of the district as a whole. According to Taylor (2003), the mother of the student questioned the variances in grading amongst the teachers at West Valley, and as a result, was subsequently able to make the case that, due to the subjectivity of grades at this high school, her daughter should have been allowed to graduate with her class. Subsequent to this case, the district changed many of their

policies regarding grading and concluded that a standardized method of grading would be appropriate moving forward, so as to keep teachers on the same page with respect to their grading procedures (Taylor, 2003).

This case study was a rare occurrence, but it brought to light many of the issues that educators had to face, as of this writing. Litigation, or the threat of litigation, often brought up policies that may be well intentioned, but flawed in implementation, and although the researcher did not agree with the basis of the lawsuit, it did bring up the concept of standards-based grading as a means to provide consistency and validity to courses and the grading in those courses. When grades were not arbitrarily developed as a result of one teacher's thoughts on the importance of one aspect of a class, parents, students, teachers and administrators were better able to determine what a student knew and whether, based on that simple requirement, that student was ready for graduation, or to move on beyond that class.

High stakes testing, such as the grading and testing described, had long been the subject of debate within education in that it gave students one opportunity to pass, and allowed no room for mistakes. The flawed nature of high stakes testing spoke to something that Falk (2002) identified when she researched how standards could support better learning. Falk (2002) noted, "High stakes testing causes harm," and surmised, "When high stakes consequences are attached to tests, they hold the potential for great harm" (p. 614). What this meant was that students were too often faced with standardized tests that only had a right or wrong answer, when they should be taking into account the ability of students to use higher order level thinking skills, or to solve problems, or even to take the knowledge learned and apply it to real world problems that

they may face (Falk, 2002). In other words, students needed to be assessed in a manner that replicated what they would run into in real life, and allow those students varying forms of assessment to show mastery, rather than limit their ability to show mastery to a multiple-choice test.

There were also lessons to learn from other districts that had implemented standards-based grading and had done so successfully. Districts should reflect on what other districts did correctly and incorrectly in their respective SBG implementations, so as to avoid many of the pitfalls that stymied them when they were going through it. One such district was the Solon Community School District in Iowa, which started the process in May 2012 (Townsley, 2017). The district stated that, although most of the teachers got on board, as did many of the parents, some were still skeptical and many had questions about standards-based grading. As a result, the district came away with three lessons learned as a result of their implementation: 1) Involve voices from higher education – this could go a long way in attempting to dispel the myth that SBG could somehow hamper students in life after high school or somehow affect them getting through the college admissions process. The concern was mainly that with the increased rigor of SBG classes, there was potential for Solon students to not get scholarships they would have otherwise received; 2) Traditional grading dominated the electronic grade book market – the district could switch to a standards-based grading model, but it did not necessarily mean that the software companies were going to adjust accordingly. The researcher ran into this with the district that was the subject of this study. In using Tyler SIS as the gradebook model, there had to be many lines of open communication between teachers in the district, the IT staff, and the development staff at SIS in order to process

and adjust any recommendations that teachers and administrators may have had. After all, as the Solon district found out, if a district switched grading policies to encourage more effective communication, the electronic gradebook must be able to digitize that concept; and 3) Keep the Board of Education (BOE) 'in the know' – the Solon district did a good job of keeping the BOE informed of all changes, no matter how minute, as it related to SBG, so they would not be caught off guard in their role as a representative of the community. They would need to be the ones to explain the rationale behind the change, so it proved beneficial to keep them informed and teach them about the principles of SBG (Townsley, 2017, pp. 27-28)

Redos/Retakes

One of the cornerstones of standards-based grading, and subsequently one of the most controversial, was the idea of allowing students to retake their summative tests until they reached a point where they could show they had attained the level of knowledge needed, in order to show they had passed a particular standard. Under the then-current model of grading utilized in most schools, retakes were generally not allowed; so, a student got the score they received the first time they took the test, with no opportunity to go back and reattempt that particular standard to show they had learned the material. Wormeli (2011) stated that this happened because teachers under the then-current model of grading justified this based on the assumption that they were preparing their students for the real world by holding them responsible the first time they took the assessment, and that generally, these were the same teachers that may give a zero on an assignment if not turned in on the due date, as that taught responsibility. His argument to this was that these practices tended to have the exact opposite effect, as they hindered how the students

matured and how they achieved in the classroom, and that without hope, students tended to disengage from classroom learning (Wormeli, 2011). Dueck (2011) went on to make the assertion that, although he was skeptical of retakes based on the fact that his courses were very content heavy, he lacked time to retest, and that he felt there was no way to authentically test on a retake, he eventually came around to the idea of retakes on summative assessments. He was able to come around based on Stiggins, Arter, Chappuis, and Chappuis' (2004) assessment model, which asked three questions of students that they should be able to answer: 1) where are we going, 2) where am I, and 3) how do I close the gap? Based on those, assessments could be remade to allow students to retest on specific learning targets, as well as allow the teacher to see if a poor grade was a result of lack of knowledge or simply bad questions, which then allowed the teacher to allow the student to essentially choose his or her own form of how they were assessed (Dueck, 2011).

Although there was considerable backlash against retakes due to the previously stated reasons, Wormeli (2011) argued that the backlash went contrary to what many argued in the first place – that teachers were not preparing our students for the real world by allowing retakes on assessments. In the real world, adults did very well in their chosen professions with a system of retakes set in place. Take, for example, an airline pilot who crashed a simulator during training, or a doctor who performed a surgery wrong on a cadaver the first time they performed, or a teacher who did not pass the certification test the first time. Wormeli (2011) stated that these professionals were not precluded from continuing on, but rather were provided helpful practice so they could continue to improve and hone their skills; so, when it came time for that pilot to fly a real

plane, or that doctor to perform surgery on a real patient, or that teacher to teach to real students, they would have learned through their mistakes and came out much better on the other end.

O'Connor (2002) stated, "Learning is incremental" (p. 139). What he meant was we, as teachers, should look to the most recent achievement evidence to determine what a child knew and did not know. Learning builds on itself, so a student may know something in December that they had a difficult time understanding in August; therefore, if that student could indicate to the instructor that he or she knew the material in December, why would we mark that student with the grade earned in August, rather than updating the grade to what they knew in December? This new evidence must therefore replace the old grade rather than simply supplement it or have it averaged in, which seemed to be the standard practice under the traditional model. According to O'Connor (2002), this meant that students needed to have several opportunities over the course of a semester to demonstrate how much they knew, and what they were able to accomplish as it related to each standard in the class (p. 139). To accomplish this, students should be given the chance to go back and re-do a particular assessment over a standard they did poorly on or be provided with a secondary make-up assessment that assessed over the same material but in a different manner than the first assessment. In order to ensure students did not take advantage of this, the student should be able to provide some measure of reinforcement over that material, such as working through a study guide, seeing a tutor, or after school help from the teacher prior to attempting a re-do.

Homework

Under the traditional model of grading utilized in schools across the country, homework played a factor in the final grade of any given class. However, how much of a role it played in classes was another story. Take, for example, two freshmen Civics classes taught at the same high school by different teachers - one may have counted homework as 40% of the final grade, while the other teacher may have only counted that same amount of homework as 10% of the final grade, leading to completely different outcomes for students in the same school taking the same class, but with different teachers. Fisher, Frey, and Pumpian (2011) argued that grades should reflect students' understanding of the content, and homework being included in these grades was misleading, as a student could do well in a class simply by turning in all homework and failing the tests, while another student in the same class with a different teacher might get As on all of the summative tests, but not do any homework; subsequently, ending up with a much lower grade than that student deserved for that class. The history of homework, and how it had evolved, certainly bore an explanation in an effort to better understand how it was originally viewed and how it could be changed to be more effective going forward. Prior to the mid-19tn century, grading in schools was not taking place. However, as time marched on, the number of high schools across the country grew and with that came a sense of responsibility to hold students accountable in their classes; so, teachers began to grade based on percentages as a means to rate how a particular student had performed in their classes (O'Connor, 2002, p. 127). No matter what system was used, whether it was a curved system, pass/fail system, or something else, they all seemed to have a general commonality amongst them – teachers would add up the score and

simply calculate a grade, but not necessarily take into account what all had been included in that grade and how much it had been weighted. As this had been the norm throughout much of the history of the United States, the focus of teachers had been on point accumulation rather than what the students were actually learning in class, which seemed counterproductive to the point of education in the first place (O'Connor, 2002, p. 127). Grading, it seemed, had been turned into a competitive sport where everything a student did in a class would be considered summative in nature, which brought about a culture in which the grade tended to be infinitely more important than whether the student actually learned anything in the class to begin with. And therein lies one of the biggest issues with the traditional model of grading - that everything a student does, counts. Every assignment, no matter its significance or whether it properly assessed anything, counted toward that student's end grade. It would be helpful at this juncture to view homework as practice, because that was what it was. Homework should be given as practice in much the same manner a sports team practices repeatedly to correct mistakes before the big game. If that team makes many mistakes in practice, and then they end up performing flawlessly in the big game because of that practice, then they did not get penalized for making mistakes in practice. Rather, they were rewarded and succeeded at what they set out to do – set an end goal, worked toward that goal, and accomplished it. This was much the same way many professionals were treated in the work place. Take, for instance, an airline pilot. That pilot was not thrown into the cockpit of a large commercial aircraft without having hundreds of hours in a flight simulator. Within this flight simulator, mistakes were tolerated so that the pilot could learn from them. The pilot corrected those mistakes by repeating the same simulator over and over again, and

then was finally ready to take the summative exam – flying the real plane. Pilots were not penalized for making mistakes in the simulator, rather they learned from them. Why were our students not treated the same way? Instead, students were graded (assessed) on everything they did in class – assignments, quizzes, tests, projects, and group projects, etc. The issue was that educators should hold students accountable in much the same way we hold ourselves in our professions accountable – by practicing and practicing until we get it right, rather than being penalized as we go along. So, much consideration needed to be given to the meaning of homework by each teacher – what was the purpose of the homework and why was it being given? As O'Connor (2002) stated, most homework was "practice, and so any assessment of it should be considered formative, and thus has no place in grades" (p. 138). Students should be afforded the opportunity to practice, be given in-depth feedback, and then practice again. This cycle was what would invariably help produce students that could retain information and learn in a class, rather than create someone who could manipulate his/her grade by figuring out, on a percentage basis, how well he/she needed to do on a given assignment in order to maintain a grade in a particular class. This did not show what a student had learned.

To better understand the culture of homework, it was good to look at several common beliefs about homework that permeated through our academic culture in the United States over the century or so preceding this writing. Vatterott (2009) identified five beliefs in her book, *Rethinking Homework*, beginning with the idea that the "role of school is to extend learning beyond the classroom" (p. 10). It had long been thought that keeping children busy when they were not at school kept them out of trouble and focused, making the teacher into someone who not only controlled a student's life within school,

but outside of school as well. The second belief she identified was, "Intellectual activity is intrinsically more valuable than nonintellectual activity" (2009, p. 10). This belief was targeted toward those who tended to believe that academics should take precedence over anything else – sports, clubs, and playing with friends, etc. In order to have a well-rounded child, all of these factors were important, and none were superior to the others. Homework should not be given in such amounts that it detrimentally hurt any development of social, emotional, or physical skills of a child. The third belief was that, "Homework teaches responsibility" (2009, p. 11). Vatterott (2009) indicated there was no research that supported this belief, and then continued further to say that students could be taught to be responsible in many different ways, such as involving them in the decision-making process regarding their own academic success and making them accountable for these results (p. 11). The fourth belief she discussed was that people tended to believe that, "Lots of homework is a sign of a rigorous curriculum" (2009, p. 12). The issue with this was there did not seem to be any indicator that meant more homework equated to more rigor, in the sense that more rigor meant a higher quality of learning. In fact, many of these assignments might have been 'busy work' and had no inherent value other than making the student do something for a particular class. This was a prime example of 'quality over quantity' and; unfortunately, quantity seemed to be the norm for most teachers in this country, as of this writing. The quality of any given assignment was lost on the teacher, as in many cases adequate time was not taken to develop and critique the assignment before it was given to the students. The fifth and final belief Vatterott focused on was, "Good teachers give homework, and good students do their homework" (2009, p. 13). This insinuated that teachers were bad if they did not

give a rigorous amount of homework and students were bad if they did not complete that same rigorous amount of homework, when in reality, that was not true at all. This goes back to the quality over quantity debate; and, unfortunately, teachers fell into the trap of increasing quantity out of an abundance of peer pressure from other teachers who had taught that way for the entirety of their careers.

More recently, researchers questioned the basic framework of how teachers grade, starting with the purpose of homework itself (Shippy et al., 2013). Scriffiny (2008) noted, "Homework does not accurately express student understanding" (p. 71) and stated that in his classroom "students who were clearly learning sometimes earned low grades because of missing work and conversely, some students actually learned very little but were good at playing school" (Scriffiny, 2008, p. 71). In their own research, Marzano and Heflebower (2011) reinforced this idea of homework in accurately representing the level of student understanding, seemingly showing that arbitrary homework should not be counted toward a student's final grade, but rather simply as practice.

Going back to the 1960s, when the Vietnam war was going on and students were, more and more, beginning to speak out about prior societal norms they were expected to live up to, homework became "seen as a symptom of too much pressure on students to achieve" (Vatterott, 2009, p. 6). It was during this time that the National Education Association (NEA), a union which represented a large majority of teachers nationwide, stated:

> Children in the early elementary school have no homework specifically assigned by the teacher, that limited amounts of homework – not more than an hour per day – be introduced during the upper elementary and junior high years, that

> homework be limited to four nights per week, and that in secondary school no more than one and a half hours a night be expected. (as cited in Wildman, 1968, p. 204)

This led to a movement through the late 1960s and into the 1970s of parents advocating that their children should be free from the constraints of school and able to enjoy their friends and play and/or relax in the evenings, which subsequently led to a decreased focus on homework into and throughout the 1970s (Vatterott, 2009, p. 6).

However, as the 1980s came into focus, the public opinion as it related to homework began to head back in the direction of the pre-1960s era, where a great emphasis was placed on homework (Vatterott, 2009, p. 6). *A Nation at Risk* was a governmental study published in 1983 that attempted to tie the failings of our public schools to the dip in the U.S. economy (as cited in Kralovec & Buell, 2000, p. 50). As a result of this conclusion, more standards were put in place, what was called an "intensification movement," where the only answer to the failings of education in our society was to have more of it, in the form of longer school days, longer school years, much more standardized testing, and much more homework, particularly for high school students (Vatterott, 2009, p. 7). As the country entered the 1990s, this trend continued, and more and more homework was assigned, under the assumption that it correlated to increased student achievement (Vatterott, 2009, p. 7).

As the 1990s wore on, public opinion again began to shift back toward the direction of less homework, as journalists and parents started to become more outspoken about how the increases in homework negatively impacted the life of the family (Ratnesar, 1999). When Ratnesar published "The Homework That Ate My Family" in

the January, 1999, issue of *Time* magazine, the homework debate effectively came full circle. In this, Ratnesar (1999) argued that homework had become an "intrusion on family tranquility and added one more stressor to an already overstressed life, especially for two-career families" (p. 57). In this same article, Ratnesar (1999) also cited a study that had been completed by the University of Michigan indicating that homework for students aged six to eight showed an increase in assigned rates of 50% over the period from 1981 to 1997 (p. 58). As of this writing, the debate over the inherent value of homework continued, with some comparing it to politics and religion, in the sense that it could be construed as a faith itself (Kralovec & Buell, 2000, p. 9).

In the culture of 2018, parents seemed to be not so concerned about the concept of completely eliminating homework, but rather finding a balance between school and childhood, as a loss of free time was linked to a loss of leisure time, higher stress, and an overall decrease in health (Vatterott, 2009, p. 22). According to Jensen (2000), parents seemed to be primarily concerned with the idea that students were "losing their childhood" and "don't have time to be kids" (p. 22). The concerns of these parents were also backed by research that supported the concept and importance of a childhood that indicated a need for downtime and rest in order to be able to achieve "peak learning efficiency" (2009, p. 22).

Homework also tended to bring into focus the issue of socioeconomics, in the sense that typically students from middle-to-upper-class families tended to have better home environments to complete homework in, while students in lower-class families tended to not have that same support at home, causing homework to create inequalities among students in classrooms where all things were supposed to be equal. In her book

Rethinking Homework, Vatterott (2009) identified several gaps which occurred when a student's home environment suffered. The first of these was what she referred to as the 'reading gap,' meaning that students in a lower family income bracket did not have access to books at home, and as a result were less likely to be reading at home. They were also much less likely to witness their parents reading for pleasure and attempting critical thinking activities. The second gap identified here was the 'conversation gap,' which was a gap in the amount of conversation a student had at home with their parents as they were in their critical years of growth. Vatterott (2009) noted, "By the age of 3, children of professional parents have a vocabulary twice as large as that of welfare children" (p. 36). The third and final gap noted was the 'health and housing gap,' which indicated that children in lower-class families lacked access to regular health care and had worse environments to call home. Subsequently, this led to students missing school for common health issues and created a poor work environment at home for students to be able to complete homework in (2009, p. 37). These inequalities at home translated into inequalities at school, and when homework was counted as part of a grade, it would put students of lower-class families at an inherent disadvantage from the start, while giving an advantage to students who came from better home environments. Limiting homework not only encouraged a more accurate portrayal of what a student had learned, but it also helped put all students onto an equal playing field academically, and potentially set students up for success.

Unfortunately, studying homework and how effective it was presented several problems, so it ended up being very tough for both the pro- or anti-homework groups to make any concrete assertion that their side was correct and the other wrong. For

educators, the primary goal of homework was to increase achievement and improve learning. Secondary goals of homework related to character education traits, such as responsibility, and hard work, etc. But for researchers, the question was this: How could you tell where the effect of classroom instruction ended and the effects of homework began? According to Vatterott (2009), educators did not know. She stated:

> We don't know how to tease out the effect of homework from prior learning or what occurred in the classroom. We don't know if the same child would have scored just as well on the test without doing the homework, or how much better the child scored because of doing the homework (p. 57).

That being said, the findings of much of the research done on homework could be broken down into generalities, though even these general assumptions should be viewed with a healthy dose of skepticism, given the number of variables that played into any research done on homework. These generalities, summed up by Vatterott (2009), were:

> 1) the amount of time spent doing homework is positively correlated with achievement; 2) homework appears to be more effective for older students than younger students; 3) as more variables are controlled for, the correlation between homework and achievement diminishes; and 4) at each grade level, there appears to be an optimum amount of homework. (pp. 59-61)

Based on these general findings, it still seemed as if the jury was largely out on how effective homework was and; if so, how much homework should be required of a student at each particular grade level? These questions attempted to be answered in a standards-based grading format by encouraging homework as practice, but not counting it in the grade, but then allowing that homework to help a student achieve a '4' in class, which

showed evidence of mastery of a particular standard. This way, the incentive still remained, while also not counting against the student, except in the potential to do poorly on a summative assessment in the class.

Feedback

With the traditional model of grading, teacher feedback was not always a required aspect. Often times, grades could be marked A, or B+, or F without any real explanation as to why a student received that grade. The student was left with no recourse in this situation – no option to learn from mistakes and improve, no option to talk to the teacher and find out how it could be done better, and no option to counter any feedback with the student point of view and show learning in a different manner than maybe the teacher had intended. For these reasons, the traditional model of grading failed students and parents, in that teacher feedback was absent from the process. No one really had a good idea of what exactly it was that a student learned or did not learn. Rather, we tended to just look at the end result – the letter grade – and not ask any questions. In many cases, teachers were not even taught how to grade and provide proper feedback in their undergraduate or graduate programs. The focus was on content area and lesson planning, with very little taught in the way of how to effectively grade. Feedback would fall into that category. With effective feedback came increased student achievement. The question was how to provide teachers with the tools they needed in order to change their ways and help provide that effective feedback within the classroom environment.

In "Designing a Comprehensive Approach to Classroom Assessment," Marzano (2007) identified a five-step process in designing an effective system of comprehensive classroom assessment with a focus on instructional feedback. Marzano's (2007) first

finding was, “Classroom assessment feedback should provide students with a clear picture of their progress on learning goals and how they might improve” (p. 103). Students needed effective feedback if they planned to increase their achievement level in any given subject. Marzano’s (2007) conclusion that feedback was necessary for increased student achievement stemmed from a study of 8,000 students by Hattie (1992), in which it was found, “The most powerful single modification that enhances achievement is feedback” (Hattie, 1992, p. 9). This was supported again by Hattie in 2007 with an updated version of the study from 1992 that came to the same conclusion about the effectiveness of feedback in increasing student achievement (as cited in Hattie & Temperly, 2007, p. 85). Marzano (2007) then went on to explain that feedback itself was not enough, that the type of feedback given mattered just as much when it came to increasing student achievement (p. 104). He referred to a meta-analysis by Bangert-Downs, Kulik, Kulik, and Morgan that looked at 40 various research studies on student achievement and classroom assessment that found,

> Simply telling students they were correct or incorrect in their answers had a negative effect on their learning, whereas explaining the correct answer and/or asking students to continue to refine their answers was associated with a gain in achievement of 20 percentile points. (as cited in Marzano, 2007, p. 104)

This spoke to what every teacher should know, but clearly did not – explain answers and give the reasons/rationale for whatever it was one was looking for. This helped guide students and helped them learn more efficiently. Students must have feedback when they did things in class incorrectly, so that they knew how to do them correctly going forward. Simply telling a student that they were correct or incorrect did not do anything at all in

terms of that student grasping and comprehending the material needed to understand for the class. Marzano's (2007) second finding was, "Feedback on classroom assessments should encourage students to improve" (p. 104). In his own research, Marzano (2007) developed two ways of encouraging feedback in an effort to provide constructive feedback without discouraging the student going forward.

> First, he stated, the feedback provided must be provided in such a way that it did not discourage the student, even if the score was low. One should focus on positives. Second, he believed that students had to understand and subsequently realize that the hard work they put in on the front end would result in better results on back end, increasing learning for that student. (p. 105)

Contrasts Between SBG and Traditional

Although the traditional model of grading had been around for decades, the duration of something did not always determine its effectiveness in achieving the outcomes it desired. Table 1 lays out the contrasts between the traditional model and SBG.

The table provides a great place to start to understand the variances between the two grading systems, traditional and standards-based, as administrators and teachers needed to not only know the information, but they also needed to know the 'why' behind the information. In other words, why were the changes necessary? Why did the traditional model not work? As the model stood, everything fell into varying categories in a teacher's gradebook. However, O'Connor (2007) stated, if grades were to be meaningful "they must be built on more than these methods" (p. 131). Rather, grades

should be based on established learning goals, or targets, set up as part of the curriculum and followed by every like teacher.

Table 1

Traditional Grading System vs. Standards-Based Grading System

Traditional Grading System	Standards-Based Grading System
1. System is based on assessment methods. One grade is given for each subject.	1. System is based on learning goals and performance standards. One grade is given for each learning goal.
2. Assessment are norm-referenced and based on a percentage system. Criteria are often unclear or assumed.	2. Standards are criterion-referenced and proficiency-based (using a limited number of levels to assess performance on a scale). Criteria and targets are known to all.
3. Use an uncertain mix of assessment of achievement, attitude, effort and behavior. Use penalties and extra credit. Include group scores.	3. Measure only achievement. No penalties or bonuses are given. Includes individual evidence only.
4. Score everything – regardless of purpose.	4. Use only summative assessments for grading purposes.
5. Include every score, regardless of when it was collected. Assessments record the average, not the best, work.	5. Emphasize the most recent evidence of learning when grading.
6. Calculate grades using the mean.	6. Use median, mode, and professional judgment to determine grades.
7. Assessments vary in quality. Some evidence comes only from teacher recollection.	7. Use only quality assessment, and carefully record data.
8. The teacher makes decisions about grading and announces these decisions to students.	8. Discuss all aspects of grading with students.

Note: Adapted with permission from O'Connor, K. (2002). *How to grade for learning: Linking grades to standards* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Then-currently, grades were also based on one letter grade, so if a student received an 'A' in a class, what exactly did that tell the student? Standards-based report

cards provided a "summary symbol for each learning goal to provide a profile of the student's achievement" (O'Connor, 2007, p. 133).

Table 2

Proficiency Scale for Legislative Power Standards in 9th Grade Civics

Strand: Branches of Government Topic: Legislative Branch Grade: 9		
Score 4.0	Student is able to meet all expectations for level 3 plus in-depth inferences and applications that go beyond what was taught. 3.5 In addition to score 3.0 performance, in-depth inferences and applications with partial success.	Sample Activities • Write essay detailing the process of making their own bill and how it becomes a law
Score 3.0	The student: • Students will be able to identify the requirements to be a member of Congress and understand how Congress works The student exhibits no major errors or omissions. 2.5 No major errors or omissions regarding 2.0 content and partial knowledge of the 3.0 content.	• Develop a scale detailing how a bill becomes a law
Score 2.0	There are no major errors or omissions regarding the simpler details and processes as the student: • recognizes or recalls specific terminology, such as: o Constituents, legislation, congressional powers, checks and balances, leadership positions o Recognize or recall the checks and balances built into the constitution 1.5 Partial knowledge of the 2.0 content, but major errors or omissions regarding the 3.0 content.	• List the individual checks & balances in our government
Score 1.0	With help has 2.0 0.5 With help, a partial understanding of the 2.0 content, but not the 3.0 content.	
Score 0.0	Even with help, no understanding or skill demonstrated.	

Note: Information from the researcher's own educational practices and experience.

Table 2 provides an example of one standard from a 9th grade government class taught by the researcher. This indicated the 'power standard,' which was the strand at the top of the chart, followed by the learning targets. The scale represented in Table 2 uses what O'Connor (2007) referred to as "Criterion-referenced performance standards" (p. 133). These standards were what helped determine a grade in a particular class rather than the traditional model, which focused on a percentage equaling a certain letter grade. The problem with this form of grading was that there tended to be no real meaning with the grades, because those letter grades were arbitrary in nature and there seemed to be no consistency among teachers. With standards-based grading, the performance standards referenced in Table 2 had "clear descriptions," which allowed for consistency and meaning to be applied to grades so that students, parents, and teachers could have a better grasp on what a student had specifically learned and what he or she had not learned (O'Connor, 2007, p. 133). In Table 2, the descriptors are numbered one through four and are described as follows:

- Level 4 – Advanced: the student has mastered the material and gone above and beyond what was expected; critical thinking was applied.
- Level 3 – Proficient: the student has performed well enough to know and understand the material taught; student has demonstrated a level of competency, but not mastery, of a particular topic.
- Level 2 – Basic: the student has a partial understanding of the topic taught but has still left some gaps in their knowledge. Student has to reference books and not achieved a level of knowledge to show complete competency in the topic.

- Level 1 – Below Basic: the student has a very limited knowledge of the skills taught and has below average expectations for the opportunities in class to help bring the grade up.

O'Connor (2007) noted that the aforementioned performance standards were "Criterion-referenced, not norm-referenced" (p. 135). What this meant was that grades were determined by how the student did and were an actual representation of what a student learned, rather than norm-referenced, which would involve some sort of curve and an assurance that at least some students in the class would end up with an A, regardless of whether they actually learned anything to earn themselves that A.

One of the biggest drawbacks to the traditional model of grading was that, at its core, educators were unable to find out exactly what a student knew. In other words, they did not know what went into a particular grade. There could be several factors, including achievement, effort, participation, homework, attitude, tests, quizzes, projects, and extra credit; and the list could go on. This begged the question, however, of what a grade should be based on. Should it be based on all of the above, or just on achievement? O'Connor (2007) would argue that it should be achievement based, and the other factors, while important, should not be considered in the final grade (p. 136).

With SBG, students had the opportunity get involved with the assessment process and had some input into how they were able to demonstrate proficiency in a specific standard. As it related to differentiated learning and assessment, this provided a unique opportunity for students to show what they had learned in a manner in which they felt most comfortable, rather than taking a catch-all summative assessment provided by the teacher for all to take, regardless of learning style or ability. In this researcher's

experience, allowing the students to help conceive of ways in which they could be assessed grew interest in the subject and allowed students to take ownership of their learning; subsequently, they chose to put more effort in and the knowledge was retained on a more permanent basis. According to O'Connor (2007), "Higher levels of achievement occur when students are involved in their learning" (p. 143). Based on this, it stood to reason that if students were aware of how their grades would be determined, and if they were somehow involved in the assessment process, then student achievement was going to rise as a result; because, one could assume, they had a stake in the outcome.

Report Cards and Reporting

Reports cards were a staple of any communication piece in schools in 2018. They showed how a student was doing in class, and generally would indicate behavioral reports as well, as indicated by 'effort' and 'conduct.' For most of the past half century, as of this writing, these types of report cards had become the standard for any household with school-aged children. They served as a communication tool well before the advent of the internet, emails, and online grade reporting software. Through these, teachers could report home as to how a student was doing up to that point in the semester. However, as time marched on and other aspects of education began to change, the report card remained, in large part, the same. With the exception of parents today being able to log in to a district website to see the most current assignments and grades as reported by teachers, the actual format and what they indicated to the parents and students remained the same. As SBG started to make waves through education, it naturally led to a conversation on how grades were communicated and how students knew what they learned and what they had not. The answer to this was that students, on the whole, only

knew a letter grade. They were largely unable to state what exactly they learned based on definitive standards, nor did they know how they could bring a grade up. SBG, it was hoped, would be the answer to the question of how schools could effectively report how students were performing in class and how they could best communicate that with parents and students.

The traditional grading system was normally faulted with two main issues: curriculum and communication. The curriculum issue was seemingly solved by establishing standards related to specific content. The communication issue was an issue that continued to challenge educators. According to Guskey (2004), trying to establish a report card that based itself on standards was a challenge, due in large part to the challenge of identifying the learning goals and then deciphering specific performance criteria. Once this step was complete, each level must be benchmarked, which was generally where the four levels of mastery were developed (Guskey, 2004). The researcher had experience with this step in that when SBG was implemented in the researcher's school district, he was tasked with developing these benchmarked labels for the classes he taught. This proved difficult to accomplish, due to the forced collaboration that SBG brings to the table in that educators must find common ground with other teachers in their department and across buildings as they were tasked with developing the most important aspects of the class and agreeing on what needed to be put in and what needed to be taken out.

One of the biggest hurdles to overcome as it related to reporting on SBG was ensuring that parents understood the language expressed in the report cards, and if they did not understand the language, Guskey (2004) stated that educators needed to take on

the responsibility of knowing that it needed to be corrected and ensuring that the message was "clear and comprehensible" for the people for whom it was intended (p. 327). It was within this light that Guskey (2004) developed four guidelines when developing a standards-based report card, as follows:

1) Avoid comparative language
    a) This helps the parents understand how their child is doing not in the class, but rather how they are doing with respect to the expectations for a particular level.
2) Provide examples based on student work
    a) Teachers should provide examples for each level of work that they expect to meet the expectations for that level. These examples can help further enhance the teacher expectations for the parents so that the parents can help where needed
3) Distinguish between "Levels of Understanding" and "Frequency of Display."
    a) Don't use words like "often" or "occasionally" when describing various aspects of a child's performance.
4) Be consistent
    a) Parents, by and large, have a tendency to translate everything into a letter grade, as that is how grades were reported when they were in school. As that changes with SBG, parents may get confused with new terminology and how that relates. Teachers should be consistent in their use of terminology to describe student performance and be descriptive with them so as to avoid confusion. (p. 328)

With report cards, there were generally two distinct types. The first type was the

traditional model which indicated simply what the child received in a particular class as indicated by a letter, and then there may have been some indicator of effort and behavior. The second type was the standards-based report card, which showed each standard in a class and indicated to the student and parent the progress the student took to get to the present level of mastery, which then allowed that student to know what he or she did well, and which areas needed to be studied more and then taken again. In the first, traditional report card, there was not much to take away as a parent or student. If there were any remarks from the teacher, they were generally short and read like 'pleasure to have in class' or 'does not use class time wisely.' In other words, there was not much to go on in determining why a student's grade was what it was, and then a detailed way for the student to bring that grade up and learn what the teacher wanted them to learn. A standards-based report card, on the other hand, was generally going to be very detailed with student specific comments from the instructor. Further, the actual meanings of the grades were gone over in detail, so that the parent had a much better understanding of what their child had to do in order to bring that standard up (O'Connor, 2017). These grades and standards that led to the grades also showed evidence of learning over time from various sources other than exams – this included teacher observations, conversations, and student-directed assessments, just as a few examples. According to O'Connor (2017), teachers that implemented SBG and the reporting that went along with it said it, "contributes to a learning culture, in place of the traditional grading/point accumulation culture, and that students become self-directed learners who have a much more positive attitude about school and learning" (p. 24).

The researcher, as a parent and an educator, understood the need for more explicit and specific standards for assessing, as it remained increasingly important to have a firm understanding of how the child performed in any given class and then had the ability to look at specifically what that student needed to improve upon in order to achieve mastery. The traditional, old-style report card did not allow for this. O'Connor (2007) went so far as to develop four conditions that must be met for effective grading:

1) Grades must be accurate because they support important decisions based on grades.
2) Grades must be consistent.
3) Grades must be meaningful.
4) The process that leads to the determination of grades should support learning, not just the accumulation of points. (p. 24)

As has been stated, the grading systems of old did not support O'Connor's conditions, as that scale did not accurately represent a mastery of learning goals.

One of the most significant factors the researcher believed affected teacher grading systems was that teachers did not generally receive classroom instruction during their undergraduate or graduate classes, as related specifically to grading and reporting, and once out in the workforce, there was minimal guidance from state and local authorities, more than likely due to the fact that they themselves had not received adequate guidelines on how to effectively assess and grade students. As O'Connor (2017) stated, standards-based grading and reporting went a long way in improving the lines of communication between parents, students and teacher, it required teachers to follow standard guidelines that focused on "shared practices," and from the perspective

of teachers, it honored the idea that teachers were professionals and could grade based more on that professional judgment, rather than simple numerical values (p. 24).

Assessments

Assessments, as society thought of them at the time of this writing were simply tests, typically lumped into a category like, unit tests, chapter tests, or some other end-of-section test that tested a student on what was just covered. A student took a test over Chapter 1, that child received a grade, the grade went in the gradebook, and that was it – the student got one chance, and the impact of the one chance would be with that student, for better or for worse, until the class was over. The tests were kept secret prior to the test date, and students were typically left to their best educational guesses about what might be or what might not be on the test. Needless to say, much rode on the outcome of these tests, and pressure built. If a student did poorly on one test, there was potential for that student to then give up, knowing that a grade might not be recoverable to the extent they wanted to recover it, and they may have focused their efforts elsewhere. Unfortunately, this scenario played itself out too often in our schools, at the time of this writing, in large part due to the traditional view of assessments ingrained in our schools – you (the student) study, you (the student) take it, you (the student) deal with the results, and if you (the student) wanted to do better, you (the student) should have studied harder, or longer, or both. The question that then presented itself was: How did that test the overall knowledge of the student when they left the class, and what if the student began to understand a concept later than his or her classmates? This could be answered by beginning to understand various forms of assessments and their role in the classroom, and

how they could effectively change student outcomes in classes and work within the constructs of standards-based grading.

Assessments in some form or another were around since the dawn of education. In the most basic form, an instructor was simply trying to know what the student knew and have the student retain that information. There seemed to be strong support of the fact that assessments were necessary; however, there seemed to be widely differing ideas on what type of assessment, coupled with effective instruction, reached this goal. A teacher was left with several choices when looking at what types of assessment he or she might use within their classrooms. In *Formative Assessment and Standards-Based Grading*, Marzano (2010) discussed three types of classroom assessment and three uses of classroom assessment. The three types were 1) obtrusive, 2) unobtrusive, and 3) student generated, and the three uses were 1) formative scores, 2) summative scores, and 3) instructional feedback (Marzano, 2010, p. 23). Obtrusive assessments were just that – obtrusive to the classroom. These were assessments that impeded the normal classroom instructional time, and more often than not looked like the old-fashioned pencil and paper tests most people grew up with. Unobtrusive assessments were the flip side to obtrusive assessments, in that they did not impede upon the instructional time at all, and in many instances, the students may not have even known they were being assessed. If, as a teacher, one saw someone doing something, and in doing that, the student displayed some proficiency in something that was discussed in class (e.g.: a student showing a friend where they lived on a map and the teacher observed this as part of a geography lesson on map reading), then this would be considered an unobtrusive assessment. The next option was a student-generated assessment, which meant that the student essentially decided

how they were going to be assessed, be it by an oral discussion, project, answering questions in the back of the room, or any other number of things a student could have done to show that he or she had become proficient in, concerning the topic at hand.

As for how those assessments were used in class, reference back to Marzano (2010) and how he classified this: formative, summative, and instructional feedback. In this, a teacher would use multiple formative assessments over the course of a semester to determine what the students' summative score would be. According to Marzano (2010), this final number would not be an average of the formatives, but rather would be a look at the students' "pattern of responses over time," due in large part to the fact that averaging would only make sense if "no learning had occurred from assessment to assessment or if assessments measure very different things" (p. 28). In other words, the formative assessments utilized over the course of the entirety of the semester would then indicate to the teacher what summative grade that student earned at the end of the semester, taking away the weight and importance of end-of-semester final exams.

The last use of assessments was for instructional feedback. Instructional feedback could come out of all of the previously discussed types of assessments, and these assessments may or may not have been scored (Marzano, 2010, p. 31). These assessments used for instructional feedback should not be utilized as a means to track the students in a formative fashion; rather, they were to be used to show how the students were progressing (Marzano, 2010, p. 31).

Guskey (2007) argued in *Using Assessments to Improve Teaching and Learning*, that because assessments provided such high quality feedback for both students and teachers, they should not mark the end of the learning, but rather should be followed up

with what he referred to as “High quality corrective instruction,” which was simply going back, determining where the student had trouble and then help remedy that trouble that was indicated on the first assessment (p. 21).

While SBG did not advocate for the elimination of formative assessments, it did advocate that formative assessments should not count toward the final grade, as that final grade indicated mastery, as formative assessments were merely practice and should not be held against a student when they were simply learning and progressing toward mastery of a subject. That being said, homework was still important. Grodner and Rupp (2013) found that homework still played a very important role in student learning, particularly for those students who started a class performing poorly. While this study focused primarily on a freshman economics course nationwide at post-secondary institutions, the principles of it were sound and could be backed up by several others, that all definitively suggested that homework did positively influence student achievement in whatever course they were in, secondary or post-secondary. Based on this information, one could reasonably assume that teachers would be more likely to assign homework in their courses, SBG or otherwise. However, Grodner and Rupp (2013) found that only about half of those teachers surveyed actually utilized homework and when asked why, over two-thirds of those that did not assign homework did not do so because they did not feel as if they had enough time to go through it and assess it. In the course of their research, Grodner and Rupp (2013) found three main results:

1) Students who were required to do homework had significantly higher retention rates.

2) Requiring homework improved test scores, especially so for students who initially performed poorly in the class.
3) They found that homework submission had a large positive effect on test performance. (p. 106)

These results indicated that homework, or formative assessments, did have a clear and significant impact on the outcomes for students, particularly as related to knowledge retention, summative improvement, and overall improvement in the class. SBG did not eliminate homework, as was previously stated, but seeing as it tended to improve student achievement, the data would lend itself to making a strong argument that formative assessments should have counted for credit, maybe as another standard in and of itself, where mastery was the completion of all homework, and then down from there.

Teachers seemed to be concerned that with the lack of 'teeth' given to formative assessments under SBG; students would generally opt to not complete the practice, blow it off, and then just try to take the summative test and take it again if they failed. However, given that educators knew that formative assessment had value in student achievement, and knew that practice should not be penalized, how were the two reconciled? According to Deddeh, Main, and Fulkerson (2010), the two could work within each other to improve student achievement.

In developing eight steps to meaningful grading, Deddeh et al. (2010) were able to show that teachers could assign homework, assess it, record it, and provide some bite to it, all within the confines of achieving mastery based on summative standards in SBG. This researcher was part of the same issue when implementation of SBG came to the high school where he taught, as many teachers were concerned that they would see a

noticeable drop in student completion rates on formative assessments as SBG was implemented. One of the core features of SBG was, "Formative assessments are weighted at zero percent: practice without penalties" (Deddeh, Main, & Fulkerson, 2010, p. 55) and parents and students alike seemed to take some time in getting accustomed to the idea that homework did not 'count' into their grade, but rather counted in that it allowed students to make mistakes, learn from those mistakes, and that the learning was essentially a 'work in progress' that would finish at the end of the course, where mastery would be tested summatively. The researcher found one of the biggest misconceptions with SBG and formative assessments was that if it did not count for anything in terms of a grade, that students would not do it. The researcher found, based on his own classroom experience, that when homework was assigned, it was completed just as it would have been under the traditional grading system. Under the traditional model, in a class of approximately 30 students, the researcher would generally have about three to five students not complete an assignment or would only partially complete it. Under SBG, the researcher found that in the same-sized classes, only about three to five per class would not complete it. This was also true for other teachers within the researcher's department; so, it would seem to be the rule rather than the exception. Further, Deddeh et al. (2010) found, "Students become internally driven to learn the material rather than externally motivated by a homework grade" (2010, p. 56). Knowing that the assigning of homework did lead to higher levels of student achievement and knowing that SBG had a solution to the idea of not weighting formative assessments, it seemed that the concern about the lack of formative assessment teeth in SBG was invalid, as research (formal and informal) showed that those concerns were not anything to be concerned about.

Student Achievement

Over the past few decades, there was a rising tide of people who opposed the amount of standardized testing that took place in the schools of 2018. On the contrary, that rising tide also created a group that believed in the value of standardized testing and wanted to see it continue. The problem was, up until the advent and broad implementation of standards-based grading, there was no way to keep a hand in both pockets, so to speak. Teachers continually administered both summative and formative forms of assessment of a nature similar to standardized testing, meaning that standardized tests and classroom tests had similar multiple-choice formats in many instances. Historically, the issue was that although teachers were administering these tests statewide and nationwide, there was no consistency at all in what was taught within the classrooms, much less in how those students were assessed on what they learned, leaving no good way to assess large groups of students in a statewide or nationwide classroom assessment that aligned with a particular class (other than College Board Advanced Placement courses, which did align and test nationwide). What this meant was an 'A' in one school in one region was going to be completely different than what an 'A' meant at another school in another region. Based on this, it would be inappropriate to have individually assessed assignments from all types of teachers utilized for the same purposes of most standardized tests nationwide, primarily because there would be so much inconsistency in those marks.

SBG could allow for a solution to this problem and could subsequently allow for the combination of standardized tests and teacher-made assignments to assess in what areas the students achieved mastery. In other words, the strengths of both systems would

be able to be kept and utilized to measure mastery of a topic, as well as allow for the ability to use that same data to assess how schools were doing and how students were doing nationally. This was not to say that it would be easy by any stretch of the imagination, but possible. In a study by Schneider, Feldman, and French (2016), the researchers recognized that the ability to develop common assessments across a broad swath of schools would be difficult, given vast differences in demographics; further, this type of assessment would require teachers to be trained better than they were, as of this writing, as it pertained to assessment construction; and finally, there would be the politics, and that would range from teachers' perceptions of why their professionalism was being questioned (even if it was not) to the perceived loss of autonomy as it related to assessment practices. Schneider et al. (2016) developed four guidelines for schools, districts, and states that could potentially be interested in attempting to implement these discussed ideas:

1) Collaborate
    a. Use technology and bring all stakeholders together.
2) Start somewhere
    a. Start staff on professional readings related to SBG; try field-testing ideas to see how they work in small settings before larger implementation.
3) Go slow
    a. SBG is a time consuming process and the goal should not be to place an undue burden on teachers in accomplishing this task quickly.
4) Don't stop halfway

a. Once you start the process, follow through with everything (Schneider, Feldman, & French, 2016, p. 63).

In an interview conducted by the Editor in Chief of *Educational Leadership*, Scherer (2001), Marzano indicated that classroom assessments could be utilized to track student achievement if "kept track of appropriately and scored appropriately" and contended that down the road, there would be "enough research to ensure that standards-based classroom tests can be at least as precise as external tests" as the two types of feedback – external assessments and internal classroom assessments – would "balance each other and will lessen the need for a single high-stakes test (as cited in Scherer, 2001, p. 16). Marzano contended that the use of both internal and external assessments would serve to decrease what possible chances there could have been in making incorrect assumptions about student achievement, and would rather serve to provide a more detailed report of how much a student was achieving in class (as cited in Scherer, 2001). Overall, the implementation of SBG had the potential to increase student achievement and used that data to showcase how students were doing at the local, state, and national levels.

Standards-Based Grading in Higher Education

As previously mentioned, one of the concerns with implementing standards-based grading was the relationship this type of grading system would have with post-secondary education, and whether the two (secondary and post-secondary) could work constructively together to determine what was best for students. There were two sides to this unique conflict, though. One train of thought was that school districts utilizing SBG would actually make secondary education more difficult for the students; therefore, decreasing the students' ability to apply for and win various scholarships. Based on this

argument, there was also the thought that class rankings in those same secondary schools would be more reflective of a more rigorous curriculum; therefore, decreasing a student's grade point average and lessening the chances that student would have in gaining acceptance to post-secondary education. The flip side to this argument was that SBG could water down the then-current level of academic achievement in secondary schools, and that by giving students second and third chances and not integrating several character education pieces into a grade (effort, participation), a student was not going to be nearly as prepared as they would have been had the traditional model of grading been maintained through their secondary education.

First and foremost, research at the post-secondary level with respect to grading indicated there was a similar disconnect between teachers and students as it related to how grades were attained in that students generally seemed to believe that grades were derived from using uniform, established standards while faculty tended to report that these perceived standards were just that – perceived (Tippin, Lafreniere, & Page, 2012). This initial qualitative aspect of the research suggested that, for the most part, there continued to be much subjectivity involved in the grading at the post-secondary levels, which could lead to misunderstandings and misinterpretations, as well as conflict, between students and staff. Further, the research showed that although both students and staff believed a final grade should be a combination of effort and mastery, the degree to which they thought those two separate areas should be weighted varied greatly, with students allowing 38% to effort and 62% to performance, while staff would give 17.2% to effort and 82.8% to mastery (Tippin et al., 2012). This researcher believed that this indicated there was a disconnect between students and staff at the post-secondary level in

much the same manner as there was a disconnect between students and staff at the secondary level.

This disconnect, it was argued, could be bridged by the implementation of some form of standards-based grading at the post-secondary level. Although SBG seemed to have taken the K-12 world by storm through the nation as a whole, this same mentality of aligning grades to standards and assessments went unnoticed in the post-secondary world. Students in the post-secondary world had just as much of a right as the students in the secondary world did to have a consistent answer to the question, 'How am I doing in this class?' as opposed to the question, 'What is my grade in this class?' These two questions were similar, but inherently different, as when a student asked how they were doing, they implied that they wanted to know how much they were learning and how they had done on various standards, where the student who asked how his or her grade was simply wanted to know a definitive letter grade. The traditional grading system, in 2018, in college classes did not seem to provide adequate detail to answer the detailed question of how a particular student was performing (Buckmiller, Peters, & Kruse, 2017).

Buckmiller, Peters, and Kruse (2017) noted, "Well-defined systems of grading are rare in higher education, give the variance in instructors' assessment practices, but to the extent that common practices exist, they have been conducted to date without a meaningful body of research to support them" (p. 151). This researcher would contend that at the post-secondary level, grade reporting and assessing were more divergent from each other than at the secondary level. Many classes would grade on just a few assessments, with no possibility for make-ups, while many others would take almost everything a student did into account when determining the grade a student received.

Almost no classes offered any sort of standardized option across departments or classes, instead leaving that decision up to the professor based on their own experiences, which in large part meant they graded their classes based on how they were graded when they were in the post-secondary world – with complete autonomy to do whatever they wanted to do with regard to grades. Marzano (2000) stated that, as far as grading was concerned as a whole, professors and teachers alike had taken to "incorporating behavioral factors in assigning supposedly academic grades, arbitrarily weighting assessments, and merging diverse knowledge and skills into single assessment scores" (p. 42). Given this strategy of grading at the post-secondary level, these practices proved difficult to defend as they were used in "academic institutions founded on principles of empirical research" (Buckmiller et al., 2017, p. 152). In other words, it was surprising to see that universities who had generally prided themselves on seeking out data to verify and verify again, did not seem to have much to say about how effective their staff was at increasing student achievement and growing those students. This had been so prevalent a thought over the years that in 1999, O'Connor went so far as to "advocate for a general examination of grading practices, without which grades may continue to serve as inadequate reports of inaccurate judgments of the extent to which students have attained undefined levels of mastery" (p. 3).

Researchers wanted to find out whether standards-based grading would work at the post-secondary level and wanted to identify the perceptions of students in a college SBG class and whether those students thought it was a fair way to be assessed. Initially, the researchers found that students, on the whole, were not accustomed to SBG and were apprehensive at first, indicating to the researchers that their entire educational careers

were based on accumulated points and that they would have to adjust their thinking going forward, as well as conforming to a new learning style, which scared many of them (Buckmiller et al., 2017). Students were allowed to have the freedom to choose the projects they wanted to do, which was a nice break from what they were traditionally accustomed to. This was one of the hallmarks of SBG, in that students essentially had the flexibility to identify how they best learned and demonstrated that to the teacher or professor in a manner that showed mastery of the standard. The researchers found that the post-secondary students understood that the goal of grading should be to "improve student outcomes by changing the way instructors communicate and students demonstrate progress" (Hanover Research, 2015, p. 3). Further, they recognized that any grading system they were to be subject to should not hide the level of content mastery with arbitrary marks on attendance, effort, or any other factor that may not have an exact bearing on whether that student achieved mastery or not (Scriffiny, 2008).

The researchers indicated that SBG studies at the college level were few and far between, but the ones that existed did tend to show positive results as it related to SBG in upper level courses (Buckmiller et al., 2017). The researchers came away from their study able to detail the positive aspects of implementation of SBG at the college level, and noted that, as Reeves (2004) stated, "If our objective in assessment and grading is the improvement of teaching and learning, then there is an obligation on the part of instructors to provide information that is accurate, meaningful, and relevant to student success," (p. 37) and that going forward, both secondary and post-secondary schools must understand that one single letter grade had to showcase a plethora of information, such as achievement, effort, participation, and behavior, and that it became increasingly

difficult to understand the significance of what that letter grade meant, and why it was significant (O'Connor, 1999).

Essentially, colleges were starting to realize the advantages of standardized grading across their campuses and how students could benefit from this. Although the flow of innovation from secondary to post-secondary schools, as it related to assessment and mastery may have been slow, as more and more research developed that linked standards to mastery, the flow and cooperation between these two entities would increase. This meant that students coming from the secondary world into the post-secondary world should increasingly have a smooth transition, and issues like scholarships and class rank should have started to fade away.

Further, post-secondary institutions recognized that high schools played a large role in determining college admissions; therefore, they had a vested role in trying to ensure that secondary schools met standards that could be applicable to many different post-secondary institutions. These post-secondary schools realized that they placed a premium on a student's GPA and how they performed on the ACT, SAT, or both. However, they also realized that, as stated earlier, an 'A' at one high school might have a significantly different meaning than an 'A' at another high school. Rauschenberg (2014) described this phenomenon as, "differential grading" (p. 7) and, as described by Godfrey (2011), this occurred when "students in courses with the same content and curriculum receive inconsistent grades across teachers, schools, or districts," which happened due to several factors, including variances in teacher grading standards, grading policies of the district, the behavior of the student, and the quality of teacher, to name a few (p. 1). Camara (1998) contended that as high school teachers had such a large influence in

developing their own grade distributions, factors other than student achievement might play a role in how grades were determined, intentionally or unintentionally. As a result, teachers might "assign a student's grade to reflect effort, persistence, a personal relationship, or a desire to increase a student's chances for college admission or scholarship" (Rauschenberg, 2014, p. 4). As a result of this, and the non-standards aspect of many high schools, as it related to grading, college admissions might be giving preference to students based on factors they were unaware of, subsequently making the subjective approach to high school grading an issue when students were applying for post-secondary schooling. In this respect, colleges were beginning to work with secondary schools in developing standards that were rigorous and standard, as a means to remove the differential grading aspect of high schools and ensure that college admissions were fairer.

Standards-Based Grading in Special Education

As a part of the Individuals with Disabilities Education Act (IDEA) of 1997 and 2004, there was a legal requirement that Individualized Educational Placement (IEP) teams "plan and document how progress will be monitored and communicated for students with disabilities (Individuals with Disabilities Education Improvement Act [IDEA], 1997, 2004). However, despite this requirement, 2018 data suggested that there was less compliance with this progress monitoring aspect of the IEP than there was with any other IEP component (Etscheidt, 2006).

Based on this information, it made sense that the implementation of standards-based grading would be beneficial not only to those regular education students that it would be exposed to in their regular education classrooms, but that it also would make

the most sense as it related to the implementation of IEPs and the ability of everyone involved in an IEP to track a student's progress.

As of this writing, there had been a significant increase in both the number of students with disabilities that were being put in regular education classrooms, but also a significant upward shift in the amount of time those students were spending in those same regular education classrooms (Handler, 2003). While this was not a bad thing, as research indicated the multiple positives of involving students with disabilities in the regular education classroom (Waldron & McLesky, 1998), this change did present a unique problem for those teachers in how they reported out on those students. This was where standards-based grading fit in and could play a significant role in the assessment and reporting of how those IEP students were doing in their regular education classrooms, which allowed them to maintain that level of challenge, but also allowed for them to be assessed on standards that were not necessarily the same standards as the students in the class without IEPs would have had to master. There were also legal ramifications of this as well, as IEPs must "enable the child to achieve passing marks and advance from grade to grade" (Board of Education v. Rowley, 1982, p. 1). What that meant was that a failing grade for students on an IEP could indicate that the district did not provide appropriate educational services for that student. SBG allowed this to happen so that districts could have more of a firm footing when it came to working with their IEP students on how to best be successful. Further, as there had been little guidance from the world of special education on how general education teachers were supposed to grade IEP students, regular education teachers resorted to making accommodations on a case-by-case basis (Jung & Guskey, 2007). In this vacuum, IEP students were given

instruction and feedback from teachers who, for the most part, had limited knowledge of how an IEP functioned, or what progress goals were to be made as a result of the IEP.

As districts moved away from the traditional grading scales of the past and moved toward a grading scale based on standards, there were some problems that needed addressing for smooth integration. According to Jung and Guskey (2007), one of those inherent problems was that when districts moved away from traditional models of grading, they also moved away from what would be referred to as a "general overall assessment of learning" (e.g. How did the student perform in government?) and move toward what would be called a much more "detailed description of a student's performance on a discrete set of skills (e.g. How thoroughly did the student master the ability to differentiate between the three branches of government?) (p. 48). In response to this, districts developed a "modified standards-based grading model" that encompassed all students and was all-inclusive (Jung & Guskey, 2007, p. 48). This model was what Jung and Guskey (2007) referred to as the All Inclusive Model, and it was designed specifically to work within the parameters of a standards-based grading system as well as integrate smoothly with the reporting system developed for SBG, while at the same time meeting all of the legal requirements for reporting IEPs. The model had five steps as indicated below:

1) Determine whether accommodations or modifications are needed
2) Establish standards for modified areas
3) Determine the need for additional goals
4) Apply fair and equitable grading practices to appropriate standards
5) Communicate the meaning of the grades. (Jung & Guskey, 2007, p. 50)

These five steps had gone a long way in not only ensuring that standards-based grading procedures were implemented and done so correctly, but they had also helped ensure that students with IEPs were given the chances they needed to succeed while allowing the district to stay in compliance with laws related to IEPs. At the end of this process, the students with disabilities, as well as their families, were able to define information in a manner in which they were able to interpret accurately and then be able to use that information as effectively as possible as it related to the success of the student involved.

Standards-Based Grading Criticism

With all of the research that seemingly supported the implementation of standards-based grading in secondary schools over the past several years recent to this writing, researchers were surprised with the amount of pushback SBG received as it was been rolled out in district after district across the United States. Researchers wanted to discover why that would be the case, particularly when so much of their prior research had done so much to debunk the myth of the traditional grading scales and the impact it had had over secondary students over the past century. As with any policy change in a school district, it had always been vitally important to have the understanding and support of all stakeholders in the process, be it teachers, students, or parents. The district must have sought out the support of those groups before plowing forward with any substantive policy change, particularly as it related to grading policy. Brookhart et al. (2016) found that many teachers in districts with SBG actually did support these reforms in grading, but also found that in these same buildings, there were teachers that reported the utilization of effort, improvement, and/or motivation in determining the level of academic

achievement a particular student achieved. This aspect of the research conducted flew directly in the face of the basic ideal of SBG, which was that those factors should not factor into the final grade at all (Brookhart et al., 2016, p. 829). Brookhart et al. (2016) also found that teachers generally did not practice the use of common assessments; they tended not to follow minimum grading policies; they almost always took issue with the idea of accepting late work without imposing some form of penalty for the late work; and teachers as a whole had generally not been fond of the practice of allowing students to retest on summatives and using their most recent score on that summative as their final score, replacing the original summative exam score (p. 829).

Further, Peters and Buckmiller (2014) found that one of the largest concerns voiced by parents in districts where SBG had been implemented was the "perceived threat that SBG initiatives might pose to their children's post-secondary opportunities" (p. 153). This perceived threat showed up in the research for this study as well, indicating a very real concern from stakeholders that SBG implementation might be counter-productive to the post-secondary opportunities provided for their children. This concern is addressed by this researcher later in the study.

Summary of Standards-Based Grading

Education had, by nature, continually changed as society changed. Schools became more inclusive and became more encouraging. Schools adapted to differentiate their instructional practices to meet the demands of an increasingly diverse student population. Schools had taken on the role of a second parent, and in many instances could almost be considered the first parent. However, education lagged behind in adjusting the grading system to meet the ever-changing needs of students and families.

Like Munoz and Guskey (2015) said, grade reporting should strive to be reliable, valid, and fair, and anything short of this was unacceptable. Grading and the reporting of those grades should have always had an apparent purpose, followed up with an in-depth knowledge of the criteria utilized, as this went a long way in measuring up to the goals of reliability, validity, and fairness (Munoz & Guskey, 2015). This researcher had been around education long enough to understand that change did not come easy, but also around long enough to realize that change was needed as it related to how educators assessed and reported on our students. Educators should be at a point where they were able to look to grading and reporting more as a road map, so to speak, using explicit details to serve in providing students the progress they were promised when achieving their learning goals. As educators continually worked to improve grading techniques in education, the gap between then-current levels of learning and achievement and the desired outcomes of that learning should close. As Munoz and Guskey (2015) stated, "Grading and reporting are other important tools for what matters most: improving student learning" (p. 68).

Chapter Three

Overview

The purpose of this mixed-methods study was to determine whether the implementation of standards-based grading would correlate to an increase in student achievement as measured by end-of-course exam scores in four core subject areas, as well as determine public and teacher perception of the implementation of SBG and how that would potentially play into the ultimate implementation. In this, the researcher wanted to know whether the district-wide implementation of standards-based grading correlated to an increase in student achievement, as measured by EOC scores before and after the implementation of SBG at the secondary level. Further, the researcher wanted to know how the teachers and community felt about the implementation of SBG, and whether their thoughts aligned themselves with the statistical outcomes of this study.

The Research Site

The research for this study was based on secondary data; however, the researcher believed it was worthwhile to elaborate on where the data came from, so the reader could have a better understanding of the population on which the study was based. The quantitative research for this study came from a Midwest middle-sized school district of roughly 5,000 students in K-12. The district had two high schools, grades 9 through12, with one having a population of around 850 and the other having a population of around 700. The district had a free and reduced lunch population of around 43%, as of 2016. In addition, the district demographic breakdown was as follows: 76.2% White, 5.3% Hispanic, 14.0% Black, and small sample sizes of Asian and Indian populations. The District had a four-year graduation rate of 87.44% and a dropout rate of 2.4%. Forty-two

percent of the students went on to a four-year college or university, 26.2% enrolled in a two- year college, and 2.3% entered a post-secondary (technical) institution (Missouri Department of Elementary & Secondary Education, 2017).

The community of the district was predominantly of a blue collar, working class background that tended to support the district overwhelmingly on bond issues and initiatives brought forth for them to either vote on or participate in. The culture was one of a focus on continual increases in student achievement, as well as what strategies were best to meet that goal. The district was collaborative in nature, from the top on down, and feedback and shared visions were encouraged, if not mandated.

Null Hypotheses and Research Questions

The researcher used the following null hypotheses and questions to guide the research:

NH1: The implementation of standards-based grading will not lead to an increase in student achievement as measured by an increase in end-of-course test scores in four core subject areas between 2011 and 2016.

RQ1: How did the implementation of standards-based grading at the secondary level impact the perceptions of learning from the school district community?

RQ2: How did the implementation of standards-based grading at the secondary level impact how secondary staff felt about the change in grading systems?

Data Collection and Analysis Procedures

This study was based on secondary data provided by the Missouri Department of Elementary and Secondary Education (2017), as well as the City of St. Charles School District.

As this was a study based on secondary data, the researcher sought out the permission of the Superintendent of the City of St. Charles School District to utilize the district EOC data for the five years needed, as well as permission to utilize the results of a district survey given out to parents and staff in January of 2016. The researcher collected EOC scores by accessing the Missouri Department of Elementary and Secondary Education's (2017) database of EOC scores for the district during the timeframe noted. These data were then broken down into the four core areas studied in this research: Algebra I, English I, Biology, and U.S. Government. These data were compared to determine how scores fluctuated after the transition from a traditional model of grading to the standards-based model being used then-currently.

To determine the outcome of the quantitative aspect of this study, a *t*-test for difference in means was used, as it was best able to compare two different data sets. This *t*-test for difference in means would determine whether student achievement saw gains, losses, or no difference at all, based on a 95% confidence level.

Qualitatively, the researcher looked at the responses to the district surveys sent to parents and staff, coded that data, and then looked for any patterns that developed as a result. These data were then utilized to determine if any link existed between community perception and understanding of standards-based grading and the actual results of the implementation of SBG. In other words, the researcher set out to determine whether attitudes toward SBG reflected the actual results that SBG brought to the district, and if not, why did that discrepancy exist?

For this study, the researcher utilized secondary qualitative data gathered from the City of St. Charles School District, which was taken in the form of surveys sent to both

teachers and community members as it related to standards-based grading. Quantitative data used were gathered by the Missouri Department of Elementary and Secondary Education (2017) and disseminated to the City of St. Charles School District.

Random sampling was utilized for this study, as it related to the quantitative aspect in which 50 EOC scores from each of the four core subject areas were taken from a larger sampling totaling 6,000 scores.

Summary

In terms of methodology, the researcher conducted an analysis of de-identified secondary data, which were gathered from the Missouri Department of Elementary and Secondary Education's website for the purpose of assessing secondary students to determine their proficiency levels in the core subject areas. The survey data collected by the District related to both parents and staff were initially collected by the Office of the Superintendent of the study, for the purpose of understanding community perceptions of standards-based grading and how teachers felt it had impacted the achievement levels of the students they taught; and, as parents, the students they sent to school in that District.

Chapter Four

Introduction

Chapter Four details the results from this mixed methods study. This includes information gathered and analyzed that answers the one hypothesis and two research questions. First, the quantitative data and findings are discussed, utilizing tables and graphs to showcase the resulting data, along with explanations about the findings. The quantitative results looked at 50 random sample scores from each of the four subject areas, both prior to the implementation of SBG and after the implementation of SBG, focusing on 2013 (the last year of the traditional model of grading) and 2016 (the last year of data post-SBG implementation). The qualitative aspect of this study is then presented; this based on electronic surveys distributed to district parents and another set distributed to district staff after the implementation of standards-based grading. There were 343 parent responses on the surveys and 88 staff responses; and this accounted for the whole of the qualitative data.

Hypotheses and Research Questions

The researcher used the following null hypothesis and questions:

NH1: The implementation of standards-based grading will not lead to an increase in student achievement as measured by an increase in end-of-course test scores in four core subject areas between 2011 and 2016.

RQ1: How did the implementation of standards-based grading at the secondary level impact the perceptions of learning from the school district community?

RQ2: How did the implementation of standards-based grading at the secondary level impact how secondary staff felt about the change in grading systems?

Results

The quantitative data presented corresponded to one null hypothesis, which was taken from the initial hypothesis:

Null Hypothesis 1:

Null Hypothesis 1 (H01): The implementation of standards-based grading will not lead to an increase in student achievement as measured by an increase in end-of-course test scores in four core subject areas between 2011 and 2016.

For Null Hypothesis 1, the researcher ran a series of independent sample *t*-tests for difference in means, one for each of the content areas measured on the EOCs. For each test, the researcher used the randomization feature of Microsoft Excel to randomly select 50 EOC scores from 2013, the last year before SBG was introduced. He then randomly selected 50 EOC scores from 2016, the latest data available to the researcher. For each test, the researcher ran a preliminary test of variances to determine the appropriate form of the *t*-test to be used, since one form relied upon equal sample variances and the other relied on unequal variances. Since the hypothesis anticipated an increase in scores from 2013 to 2016, the *t*-tests were applied as left-tailed tests at the .05 significance level.

To test Hypothesis 1 in relation to the Algebra EOC, the researcher ran a *t*-test of independent means comparing Algebra EOC scores from 2013 to scores from 2016. The preliminary test of variances indicated that the variances were not equal. The analysis revealed that the scores from 2013 (M = 199.50, SD = 19.48) were not significantly lower than the scores from 2016 (M = 195.24, SD = 13.20); $t(49) = 1.280$, $p = 0.8967$; α

= 0.05. This indicated that the Algebra EOC scores did not increase from 2013 to 2016, and so the researcher failed to reject the null hypothesis for the Algebra EOC scores.

To test Hypothesis 1 in relation to the Biology EOC, the researcher ran a t-test of independent means comparing Biology EOC scores from 2013 to scores from 2016. The preliminary test of variances indicated that the variances were equal. The analysis revealed that the scores from 2013 (M = 209.16, SD = 20.54) were not significantly lower than the scores from 2016 (M = 203.52, SD = 20.61); $t(98) = 1.370$, $p = 0.9131$; $\alpha = 0.05$. This indicated that the Biology EOC scores did not increase from 2013 to 2016, and so the researcher failed to reject the null hypothesis for the Biology EOC scores.

To test Hypothesis 1 in relation to the English EOC, the researcher ran a *t*-test of independent means comparing English EOC scores from 2013 to scores from 2016. The preliminary test of variances indicated that the variances were equal. The analysis revealed that the scores from 2013 (M = 206.14, SD = 14.58) were not significantly lower than the scores from 2016 (M = 209.42, SD = 16.18); $t(98) = -1.065$, $p = 0.1448$; $\alpha = 0.05$. This indicated that the English EOC scores did not increase from 2013 to 2016, and so the researcher failed to reject the null hypothesis for the English EOC scores.

To test Hypothesis 1 in relation to the U.S. Government EOC, the researcher ran a *t*-test of independent means comparing U.S. Government EOC scores from 2013 to scores from 2016. The preliminary test of variances indicated that the variances were equal. The analysis revealed that the scores from 2013 (M = 196.00, SD = 16.19) were not significantly lower than the scores from 2016 (M = 199.12, SD = 17.08); $t(98) = -0.937$, $p = 0.1754$; $\alpha = 0.05$. This indicates that the U.S. Government EOC scores did

not increase from 2013 to 2016, and so the researcher failed to reject the null hypothesis for the U.S. Government EOC scores.

Table 3

Summary of Results of Hypothesis 1

	Pre-SBC (2013)		Post-SBG (2016)				
	N	M(SD)	N	M(SD)	d.f.	*t*-Score	*p*-Value
Algebra EOC	50	199.50 (19.48)	50	195.24 (19.48)	49	1.280	0.8967
Biology EOC	50	209.16 (20.54)	50	203.52 (20.54)	98	1.370	0.9131
English EOC	50	206.14 (14.58)	50	209.42 (14.58)	98	-1.065	0.1448
Government EOC	50	196.00 (16.19)	50	199.12 (16.19)	98	-0.937	0.1754

Research Questions

Qualitative data were collected from 343 parents and 88 staff members utilizing two different surveys related to SBG, but separate aspects of SBG as it related to their specific areas. These qualitative data were collected by the school district and was thus secondary data. The school district collected and used these data to determine perceptions of standards-based grading implementation within the school district. Survey responses were imported into a Google plug-in to aggregate the data and help determine clusters of information.

Research Question 1

The first research question for this study focused on the perception of SBG from stakeholders within the school district community and how that perception could impact the implementation of standards-based grading.

Research Question 1 is stated as:

> How did the implementation of standards-based grading at the secondary level impact the perceptions of learning from the school district community?

The first question in the parent survey asked what the parents' level of understanding of SBG at the high school level was. The results from this question are illustrated in Table 4.

Table 4

Parent Understanding of SBG

I am unfamiliar with SBG	37.6%
I have a basic understanding of SBG	35.3%
I have a very limited understanding of SBG	13.1%
I consider myself an expert in SBG and could clearly explain it to someone	10.8%
I am unfamiliar with SBG	3.2%

The results of this question aligned closely with what Peters, Kruse, Buckmiller, and Townsley (2017) said was one of the most problematic areas for the implementation of SBG, which was a, "Lack of understanding and/or support by community members" (p. 11). These researchers went on to state what was also prevalent in the district where this research was done – that, "Small but vocal groups often go to great lengths to contest its implementation" (Peters, Kruse, Buckmiller, & Townsley, 2017, p. 16). The results attained in this study only went to support results demonstrated elsewhere in previous research conducted on the implementation of SBG.

The second question in the parent survey asked what was liked about standards-based grading. After coding the qualitative data, there was one dominant response mentioned in 90 of the 242 responses to this question. That one theme was simply: nothing. Many other responses referenced something similar, like, 'Nothing, it does nothing to prepare them for college.' Many parents, although a smaller number than

above, also referenced the thought that standards-based grading would not properly prepare their students for college or for adult life in general, primarily based on the fact that they said students do not get second chances. Other parents did respond more favorably with answers where the common theme formed around liking retests, or that their unorganized but smart child did not get punished for being unorganized. These positive responses were outnumbered by the negative responses, however.

The third question on the survey was: Current policy prohibits a student from passing a class if they receive below a 2 on any standard (a 1 means they cannot do it, even with help). Do you agree with this practice? The results are illustrated in Table 5.

Table 5

Parent Agreement with SBG Policy

Yes	29.4%
No	46.9%
Not Sure	23.7%

This question was in response to the idea that students could too easily pass a class under standards-based grading. The basis of this question was that the district in this study decided to implement a part of SBG that would not allow a student to pass a class if they had fallen to a 2 or below on any one standard in a given class, essentially forcing the student to go back and work on that one standard, and in theory, retain that knowledge longer, before being allowed to pass a class. Therefore, by implementing this change to SBG, the rigor at the secondary level would increase and satisfy some of the complaints about SBG as it related to rigor and preparation for post-secondary instruction.

After coding this question, there seemed to be a little more wiggle room for SBG, as opposed to the responses on the previous question, but the vast majority of respondents seemed to conclude that even in this type of scenario, SBG would largely be ineffective, as it related to preparing a student for life after high school. One of the responses that was more prevalent than the others was the idea of averaging standard scores to come up with a final grade in a class, something which was frowned upon in the basic idea of standards-based grading.

The fourth question on the parent survey was: Factoring in homework into the overall calculation of the final grade is important. The results of responses to this prompt are listed in Table 6.

Table 6

Homework Inclusion in Grades

Yes	76.5%
No	13.4%
Not Sure	10.1%

One of the underlying principles of standards-based grading was that practice, or formative assessing, should not be counted against the student, as it was simply practice. The community of stakeholders seemed to take issue with this premise as they overwhelmingly supported counting homework as part of the grade at the secondary level

The fifth question on the parent survey stated: Knowing that all students do not learn at the same pace, should students have an opportunity to re-assess if they do not initially show understanding of the material? The results of this question are listed in Table 7.

Table 7

Reassessment

Yes	85%
No	15%

According to this, most stakeholders tended to agree that students should have an opportunity to reassess to improve learning and increase retention, which then led into the next question which stated, If students are allowed the opportunity to re-assess, should they be required to complete additional requirements in order to do so (i.e. study packet, complete late assignments, additional assignment, etc.)? Roughly the same number of stakeholders seemed to think that additional requirements to re-assess would be a reasonable expectation of re-assessment of students at the secondary level, indicating a small concession to the implementation of standards-based grading, where they otherwise seemed to dislike the strategies all together.

Research Question 2

The second research question focused on the impact that standards-based grading would have on secondary certified staff and how that would impact their overall feelings toward a change in the grading system. The idea behind the questions were to gauge whether staff perceptions about the change matched data to support or refute the claims. Research Question 2 is stated as:

> How did the implementation of standards-based grading at the secondary level impact how secondary staff felt about the change in grading systems?

The first question in the staff survey was: Would you support the addition of a homework standard immediately? The results are shown in Table 8.

Table 8

Homework Inclusion

Yes	72.7%
No	18.2%
No Opinion	9.1%

As indicated earlier in the parent surveys, one of the issues that parents and staff alike took issue with, as it related to standards-based grading, was the lack of homework in counting toward the final grade of the class. As indicated in the parent survey and in this survey, the general opinion seemed to be that homework should count toward the final grade in some way, shape, or form. In this survey for teachers, the idea was that a separate standard be made for homework, where all homework from the semester would be aggregated and a final score based on the 1 to 4 scale would then be attributed to this standard. Based on this, teachers seemed to believe that practice did need to count toward that final grade, because it was meaningful and worthwhile. This aligned with parent respondents who wanted to emphasize the same aspect.

The next relevant question in the staff survey was focused on what the goal of any standards-based learning/grading model should be. After coding this question, the most common theme throughout (as indicated by 85.1% of the respondents answering the question this way) was that the one overarching goal of any grading system should be to have 'clear and specific standards for learning.' This was followed closely (with 69% of respondents writing this) with an increased emphasis on ensuring that the 'final student grade reflects true learning.' Based on this, it certainly seemed that the overarching theme of any grading system should be that there needs to be very clear and specific standards for learning to show what students exactly needed to know and were expected

to know by the end of the class. While this question did not touch on the 'how' of accomplishing this, it did provide a basis for agreement between administration and staff as to how to get students to do better while keeping the emphasis where it needed to be – on student achievement.

The third question/statement of the survey asked the teachers to describe their biggest frustration with standards-based grading, as it was currently being implemented. After coding the answers to this, it seemed that the biggest concern among teachers (with 40.2% of respondents choosing this as their response) was that students would not do homework or other formative assignments. The next concern (with 13.8% of respondents discussing this aspect) was the difficult logistics of re-testing students. Based on this, it seemed clear that teachers were concerned about the formative aspect of the class not having any weight and consequently having the unintended consequence of teaching laziness to students and not preparing them for post-secondary education. The second largest concern, relating to the logistics of re-testing, indicated that teachers were concerned that, with a secondary teacher's student caseload, it may have been difficult to find time to re-test all of the students who may have needed to re-test. This logistical issue seemed to be a legitimate concern of the teachers, in that they were not sure how they would find time or space to re-test many students if it came to that, especially given that many of them had a student caseload of around 150.

The fourth question dealt with much the same issue as a similar question on the parent survey, in that it asked if the current policy that prohibited a student from passing the class if they received below a 2 on ANY standards, was a policy that the teacher agreed with or not. The results are illustrated in Table 9.

Table 9

SBC Pass Policy

Yes	28.7%
No	48.3%
Maybe	23%

The above survey results were similar to the results on the same question from the parent survey, almost down to the actual percentage point. It would seem that a majority of teachers and parents tended to agree that one standard should not fail a child for the entirety of the semester. It did not come up with a solution to this issue, however.

The next two questions on the teacher survey dealt predominantly with formative assessments and how that should look in a grading system. The first part asked teachers whether they agreed that allowing teachers to hold students accountable for formative assessments would greatly help SBG work at the high school level, and the second part asked teachers to agree or disagree on whether the best way to hold students accountable for formative assessments/homework was to have it count toward their final grade in some way. The results from both of these survey questions matched up almost identically, with exactly 80% of respondents agreeing with both statements. This was a clear indicator that formative assessing had been deemed important and necessary at the secondary level and that teachers, for the most part, tended to agree that standards-based grading needed to find a way to integrate formative assessments into the grading framework.

Summary

For this study, both qualitative and quantitative date were collected. There was one hypothesis in this study, which analyzed secondary data in the form of EOC scores

from the four core subjects tested at the high school level in Missouri – Biology, English, Algebra, and Government. One sample included data taken from the last year of the traditional grading system in 2013 and the other sample was taken from the last year of collected data – 2016 – after SBG had been implemented. A *t*-test was utilized to determine if there was a difference in means in each of the four subject areas. All of the *t*-tests came up with the same results, which was that the researcher failed to reject the null. The qualitative aspect of this study included two research questions. The first research question was based on survey results from community stakeholders. The second research question was based on survey results from district staff. The questions focused on the perception of SBG from both the community and staff and how that perception affected the implementation of SBG. The overwhelming result from both surveys and in answer to both research questions was that there was a significant negative perception of SBG and that the perception played a large role, as opposed to data driven facts, in shaping how smoothly the implementation went.

Chapter Five

Introduction

The purpose and intent of this research study was to gauge whether a district-wide change in grading would correlate to a rise in student achievement, as based on EOC scores. Further, the research was also to gauge how standards-based grading was perceived by both the teachers and the parents in the community and how their viewpoints would align with what the data indicated. In order to accomplish this, the researcher quantitatively utilized secondary EOC data provided publicly by the Missouri Department of Elementary and Secondary Education in a Midwest suburban school district with two high schools. These data were provided for the four core classes of Algebra, Government, Algebra I, and English II, and was spread out over a five-year span from 2011 to 2016, with 2013 being the last year the traditional model of grading was used. This produced thousands of results, but for each core subject area, there was a random sampling of 50 scores used from pre-SBG and post-SBG. As for the qualitative aspect of the study, the researcher utilized secondary data related to perceptions of the implementation of SBG. These two surveys – one for teachers and one for community stakeholders – were administered by the same school district, as mentioned from the quantitative aspect of the research. These surveys were sent out to parents and teachers after the full implementation of standards-based grading. The researcher took the qualitative data provided and coded it to develop an answer to the research questions.

Hypothesis and Research Questions

The researcher used the following hypothesis and research questions to guide the study.

H1: The implementation of standards-based grading will lead to an increase in student achievement as measured by an increase in end-of-course test scores in four core subject areas between 2011 and 2016.

RQ1: How did the implementation of standards-based grading at the secondary level impact the perceptions of learning from the school district community?

RQ2: How did the implementation of standards-based grading at the secondary level impact how secondary staff felt about the change in grading systems?

Quantitative Results

The researcher utilized one hypothesis as the basis for the quantitative aspect of this study. This consisted of testing the means of two samples that included a random sampling of 50 EOC scores from each of the four core subject areas using a *t*-test for difference in variance, followed by a *t*-test for difference in means. The confidence level for this study was 95%.

Hypothesis 1 (H1): The implementation of standards-based grading will lead to an increase in student achievement as measured by an increase in end-of-course test scores in four core subject areas between 2011 and 2016.

The researcher addressed this hypothesis by taking the data provided by MODESE and running a *t*-test for difference in means to determine the correlation. Upon completion of the *t*-test for difference in means, the researcher failed to reject the null; therefore, Hypothesis 1 was not supported by the data and no real statistically significant difference existed between the means of the two samples. This indicated that it was statistically irrelevant which grading system was used, as the results were largely the

same, at least at the 95% confidence level, which is the commonly used level in educational research.

Analysis of the Hypothesis

Based on the results of the *t*-test for difference in means conducted for Hypothesis 1, it seemed that the implementation of standards-based grading did not have a significant impact on student achievement. This lack of impact over the traditional grading system could be due to several different factors. The first factor could be that teachers had not adjusted their instruction to match the standards-based format. In other words, teachers may have technically been inputting grades into an SBG report card, but not changing anything else in their classes and converting from a traditional scale to an SBG scale to suit the needs and make it as equal as possible.

As with any significant change in a school environment, teachers did not take to it right away. As a result of this, the easy way to conform to administrative directives was to nod one's head in agreement and then convert as one saw fit in order to meet the perceived needs of the teacher and the classroom. The second factor that could have contributed to the lack of correlation could have been a lack of teacher understanding of how to articulate the standards they wanted in a format suitable for standards-based grading.

Upon implementation, there was much confusion among the staff about how to write standards, and what should and should not go under those standards. There was a period of trial and error as it related to the standards, where teachers would collaborate and develop, and then restructure for the following semester or following unit. So, in a sense, many teachers were working this new system 'on the fly' and not really taking the

time to access the research of how to work through these issues. In the early set-up of SBG, there was much confusion over how standards would be integrated in each class and how various strands would work under each standard. As a result of this confusion, combined with lots of information being thrown at the teachers at once, it was possible that not much changed at all in terms of instruction or in terms of class objectives in those early years of SBG. If this was going on, and teachers were going back and forth on how to set it up while continuing to teach in the same fashion they had been, it would make sense to see no change in standardized test scores over the time period illustrated in this study.

A third factor could be that teachers did not collaborate as they should on common formative and summative assessments, as well as a potential lack of collaboration on developing standards. One of the tenets of standards-based grading was that teachers who teach the same classes needed to have their classes structured in much the same way in order for SBG to work. In other words, if a child were to take Civics with Teacher A and transfer mid-semester to Teacher B, there should be continuity and the student should be able to jump into the new class without much acclimation. So, in order for that to happen, teachers needed to get together prior to and during the implementation of SBG to develop common assessments, both formative and summative, and work out any issues that may arise as a result.

Further, teachers were to collaborate on developing standards for their like classes and coming up with something agreeable to everyone. This proved difficult for several reasons, namely that proper time was not provided to accomplish this task on as regular a basis as needed. This forced teachers to work in a vacuum and collaborate via email,

which was not the best way to accomplish what needed to be accomplished. This collaboration also forced newer teachers and more experienced teachers to work together to develop standards for a class that they both taught. In many instances, this type of pairing did not work well, as the more experienced teacher often was against the change in the first place, as it would require all new lesson plans and assessments, etc., and the newer teacher would be all for the change as they were just coming out of college and more open to new ideas. This dynamic provided a pretty solid bone of contention among colleagues who should have been working together toward increasing student achievement, but were instead fighting for their own needs when going through this curriculum change. As a result of this, many of the deadlines set in place to help guide the elongated process of SBG implementation were missed and teachers reverted to what they were familiar with, which was the old way of grading.

The last factor that could have played a significant role in the lack of correlation between SBG implementation and student achievement increases, as measured by EOC scores, was teacher pushback. Many teachers, as indicated in the staff surveys still to be discussed, were not pleased with the theory of standards-based grading and pushed back hard against its implementation. Teachers, for many reasons, opposed it, and as a result, implementation was delayed to attempt and allow administrators to work through these differences with staff. This teacher pushback led to union involvement, which added another layer of complication and length to an already complicated and lengthy process. Some teachers simply retired as a result of this implementation. Those who stayed, in several cases, flatly refused to utilize SBG in their classrooms, setting up a situation where teachers in some classes were experimenting with SBG and other teachers in other

classrooms were not paying any attention to it at all, and going about teaching as if nothing was happening in any way that would affect the way they graded students.

Qualitative Analysis

For this study, qualitative data were collected in a secondary manner from surveys given to both parents and teachers in a small, Midwest suburban school district. The parent survey was distributed by this school district to all parents on the district email server, and it received 343 responses. The teacher survey was emailed out to every certified staff member in the district, and it received 88 responses. The questions asked on the two surveys were used to form answers to the two research questions named in this study.

Research Question 1

How did the implementation of standards-based grading at the secondary level impact the perceptions of learning from the school district community?

In addressing Research Question 1, the survey respondents overwhelmingly favored what essentially amounted to a repeal of standards-based grading in an effort to take the district back to the traditional model of grading that had always been around.

Parents overwhelmingly stated that they did not have a clearer understanding of their child's grade as a result of SBG. One hundred sixty-five respondents indicated as much on the survey responses. One of the stated goals of SBG was that parents have a much clearer understanding of the grade their child brings home; but clearly, that is perceived to not be the case. As one parent stated in response, 'It is a very frustrating system for parents to understand. When a parent looked online at his/her child's grades, it is so confusing.' SBG was not supposed to create more confusion, it was supposed to

clearly define objectives and accentuate what a student learned and what they had not learned, while making it easy for both the parent and teacher to understand. As stated above, and throughout this survey, that goal was not one that had yet to come to fruition with the implementation of SBG in this district.

When asked what aspects of SBG they would like to revise, parents indicated that they believed SBG should only be in elementary schools, as high schools should not be in the practice of allowing test retakes, as they believed this did not adequately prepare them for their post-secondary lives. Parents indicated that they did not feel SBG helped guide their children on the correct path if they were looking to be successful. As one parent respondent wrote,

> I do not see how it provides any value-added in higher grades (high school), and it is generally not received well by colleges, making it necessary to convert the results of this grading system back into standard grading, which is additional work and allows for additional opportunities for mistakes to be made.

This parent seemed to speak for many of the parents who responded and indicated in some manner that SBG at the high school level was going to negatively impact their child's ability to get into college, or be successful once in college, as they would have been accustomed to retaking tests and not having anything formative count for a grade. They believed this would discourage them from practicing while in high school, which could potentially lead to poor habits once they reached the post-secondary world.

Other parent respondents indicated that they did not believe there was consistency in how learning was being evaluated by teachers. Thirty-seven of the respondents indicated they, in some way, shape, or form, did not think that SBG was having the

intended outcomes, as it relates to consistency and continuity for a particular class across the board. In SBG, one of the goals was to have teachers who taught like-classes get on the same page. In other words, teachers should be giving the same formative and summative assessments, and the standards should be the same in the same course, no matter the teacher. This was to ensure continuity and to make sure students were getting the same education in one class as they would get in another teaching the same subject. Parents, however, did not seem to perceive this aspect of SBG as being all that successful, as indicated by the parent response, 'Inconsistency in how a final standard is evaluated from teacher to teacher and an inconsistent finals policy that could result in a failed standard despite having a three or higher on a summative standard.' This parent was referring to the part of SBG where if the student failed one standard, the student would fail the entirety of the course, regardless of whether the student had done well on the other standards. As a result, the student may end up only being able to retake that standard on the final exam, and if the student failed it there, that student would be out of luck, even if the student scored well on the rest of the standards and the rest of the exam. Parents did not seem to like this, because they perceived it as being weighted too heavily on one standard and not giving the students enough credit for doing well in the other aspects of the course.

Further, when asked what they like about standards-based grading, a sizable majority of respondents simply answered, 'Nothing.' These respondents did not give any thought to anything positive that may result from standards-based grading, but rather decided that enough was enough and it was time to deconstruct the SBG implementation before it was fully implemented. The antithesis to this is that some parents, albeit a small

minority, did see some value in SBG implementation and advocated for aspects, such as retaking tests, that it more accurately reflected what their child was learning, and that practice (the formative assessments) did not count toward the final grade, as they are just that – practice.

The parent responses on this survey overwhelmingly suggested that public perception of SBG was highly negative, in that they perceived the system as being too forgiving and not preparing their children adequately for whatever post-secondary area they planned to go into. Their reservations in regard to this seemed to be supported by the quantitative data in this study that showed no statistically significant increase in student achievement as a result of the implementation of SBG. As a result of this aspect of the qualitative study, it seemed reasonable to conclude that parent perceptions aligned with statistical data in showing that SBG may not be worth the hassle of implementation.

Research Question 2

How did the implementation of standards-based grading at the secondary level impact how secondary staff felt about the change in grading systems?

In addressing Research Question 2, the survey respondents (the teachers) seemed to overwhelmingly favor a repeal or reboot of standards-based grading, as it had been implemented.

Teachers were asked several questions as it related to SBG and how it had been implemented and developed at the high school level. One of the biggest concerns with the implementation of SBG was the lack of formative assessments being included in the final grade. This was a huge issue for teachers, as many felt that students would no longer work on homework, as the incentive to do the homework had fallen away. This

was illustrated by the fact that 40.2% of the respondents indicated that their biggest frustration thus far with SBG was that students would not do homework or other forms of practice. When one considered that the next largest concern only garnered 10.3% of the vote, one could begin to see what a big deal the homework issue had become to the teachers. The perception from teachers was that if SBG took away the teeth of the formative assessment, then why would a student be inclined to participate in the practice? And if that occurred, they would score poorly on the summative assessments, and then that would require a retake. If this cycle continued, it would potentially devolve into students continually retaking tests until they did well enough to pass, and then moving on, and no learning would have actually taken place. As indicated by one teacher respondent, 'Making formative assessments count in some way will be the biggest help in changing the current perception among students.' Another replied that, 'Formatives should count for something, even if it is a small percentage. For example, make formatives 10% of the grade and summatives the remaining 90% so that the homework has some teeth to it.' This was a pattern that was seen in many responses from the teachers. They were concerned about the lack of accountability on the part of the student and the potential for much more accountability on the part of the teacher, as a result of the implementation of SBG.

Another aspect that teachers seemed concerned about with SBG were difficulties with the gradebook, and the subsequent set up that students and parents would see at home. When SBG was implemented, the gradebook was still set up in the traditional format at the high school level. This made the communication aspect of SBG difficult, when it was supposed to be one of the selling points of SBG. As a result, teachers would

manipulate the gradebook to get grades in there, similar to fitting a square peg into a round hole. This created confusion and created a list of technical glitches that would need to be fixed by the software developer. This also created a need for a new report card, as the traditional high school report card was one sheet with each class, a letter grade, and then an effort mark. Under SBG, the new report cards would have to break down each class the student had taken into their respective standards, and then identify those standards on the report card. Therefore, a way to export the teacher data had to be set up, and then other districts that had already implemented SBG were researched to see what type of report card they were using and how effective it was. This concern was indicated by this teacher's response,

> It takes a very long time to enter grades and it is difficult to see a student's grade without clicking multiple buttons. To talk to each student in your class about their individual grades takes a very long time because you have to click through each objective to show them how they did on that objective and why.

This was a big deal because SBG was supposed to make communication clearer and more concrete, and based on teacher responses, it did not seem to be accomplishing this goal.

Implications

This researcher believed that the implications of this study may simply be that the district in question takes a different approach going forward, as it relates to the fine tuning of standards-based grading and how it will work with secondary teachers, parents, and students. Based on both the qualitative and quantitative results of this study, it could mean that the district needed to extend an olive branch to the staff and community and come to a meaningful middle ground on what could be done to best ensure high student

achievement, while at the same time continuing to ensure that like classes were taught the same with the same standards and same assessments to ensure continuity, while also ensuring that report cards were meaningful and indicated what a child mastered and what they still needed work on, rather than just giving the student a letter grade and being done with it. From the teacher perspective, the district needed to show that they were listening to their needs and were able and willing to work with them to adjust the grading system to fit what they saw as important to the success of the students, rather than basing the decision solely off of theory and not how it had been integrated at other similar schools.

If the researcher was in charge of the rollout of SBG, he would have sought out the input of the teachers and parents prior to implementation, rather than waiting until the school was a few years into the process. As far as standards-based grading as a whole is concerned, the researcher's recommendation would be to realize that SBG needed to be tweaked to serve the needs of each individual district, and that the national Marzano (2000, 2007, 2010) theories were just that: theories. Each district must take SBG and piece it together to suit the needs of their community and their teachers in order to reach the goals of higher student achievement and more informed teaching, as well as further continuity with instruction across classrooms. SBG implementation cannot be a blanket solution to every school district that it touches, because every school district is different and unique. The researcher would suggest taking the best parts of the traditional grading system (accountability, simplicity, responsibility) and blending it with the best parts of the SBG grading system (classroom continuity, meaningful grades, easy to read standards), so as to try and appease a much broader section of the teacher and parent population, while reinforcing the idea that although some things do need to change in

regard to grading, we can still take some older parts with us, as they have seemingly worked for years. The researcher believed that this approach would garner the most support from all parties involved. The big question for the researcher was: Does SBG improve the student achievement scores that each class receives, or is SBG just another name for something in education that will not change much? The researcher would advocate that almost any change is worth trying, and so long as there is no regression, positive steps should continue to be taken to ensure positive growth and higher levels of student achievement.

Personal Reflections

The researcher, during the part of his career in the classroom, recognized a need to work toward a more effective grading system that took out the 'randomness' of teachers developing their own grading scales and developing one that provided a model of consistency within classrooms. During the beginning of the researcher's career, the thought often crossed his mind that he was not qualified to develop his own grading system; and in fact, during his student teaching days questioned his cooperating teachers on why so little importance was given to what could arguably be considered the most important aspect of a classroom: the grading. The researcher did not think that undergraduate and graduate programs at colleges and universities adequately prepared students to open a gradebook and start tallying grades for students, which always struck him as odd given that these grades were what sticks with a student. So, the thought that a better grading system was needed was never really a question for the researcher. Rather, the question was always how does one go about developing and implementing a grading system that adequately does the things that the researcher thought a grading system

needed to do, namely provide a student with a grade reflective of what he or she had learned and retained in the class, and ensured that there would be consistency between like classes taught by different teachers. In other words, the researcher believed that teachers should not be given the autonomy to develop grading systems simply based on what they deemed important or not important, as a student could have one teacher attribute 40% of a final grade to homework and another teacher teaching the same class attribute 5% of that final grade to homework. If teachers could do this, what is the point? So, the researcher thought there must be a better way to implement an effective method of grading that would accomplish the above noted goals, as wells as increase the student achievement levels for students in classes that implemented these types of grading changes.

When the district that was the subject of this study began the changeover to standards-based grading, the researcher believed this was the correct course of action, as it accomplished the goal of streamlining classrooms so that students in like classes with different teachers would be graded the same and assessed the same, while eliminating the autonomy of teachers to essentially create a grading model based off what they deemed best for their own classroom. SBG eliminated the idea of extra credit, which did not really determine what a student had knowledge of; and therefore, impacted a grade positively while not correlating to any increase in student achievement. As a result, extra credit would pad grades and give a false sense of achievement to students who participated in it. The researcher believed that this was a positive change toward ensuring continuity between classes and instructors.

However, the researcher also believed that, as implemented, standards-based grading had some flaws as it was put into practice, namely there did not seem to be consequences for multiple retakes of summative assessments, and no teeth on the production of formative assessments for students. As a result of this, the qualitative results of this study came as no surprise to the researcher, as teachers seemed to key in on this aspect and focused on how little this was doing to prepare students for post-secondary life.

However, the researcher still believed that the net effect of SBG would be that students on the whole would see an overall increase in student achievement, as noted by standardized test scores, as the rigor of those classes increased dramatically as a result of the changeover to SBG. So, in that respect, the researcher was surprised to find that there seemed to be no discernable correlation between the implementation of standards-based grading and an increase in student achievement. Subsequently, the study showed that student achievement stayed stagnant over the course of the change from traditional to standards based, which then begs the question of 'why move forward with this change?'

The researcher found there was significant push back from the stakeholders in the community, as well as the teachers themselves, as neither group seemed to think that SBG was a good idea as implemented. The general consensus, as indicated by the qualitative aspect of this study, was that SBG as implemented did not adequately prepare students for post-secondary life and that it created more work for the instructor and less work for the student, in that formative assessments were essentially eliminated as students would not be required to complete them anymore. The researcher believed that this initial push back could have been avoided had there been more of an attempt to meet

these groups in the middle initially and listen to the teachers and stakeholders earlier in the process. This could have generated a productive conversation between the groups, and there could have been a standard for formative assessments, requirements for summative retakes, and clearer language, as it relates to report cards and how they were read.

However, on the whole, the researcher was surprised with the quantitative results of the study, which indicated that there was no statistical correlation between the implementation of standards-based grading and a rise in student achievement scores. Although the implementation of SBG accomplished the goal of streamlining like classes and report cards and providing more detailed information to the students and families, it did not provide a noted increase in any achievement data, which seemed to be the ultimate goal from the outset. It seems to stand to reason that if student achievement is not increased as a result, what is the purpose of the change to begin with?

The researcher will use the results of this study to inform decisions going forward that somehow integrate a controversial subject into the mix. Based on the qualitative results, the researcher would seek out the voices and opinions of the people the decision would affect most, namely teachers and community members. It seems extraordinarily important to make sure the administrator should seek out that feedback prior to any unilateral decision-making that would affect so many people in the school or the district. Education thrives on collaboration, not unilateral decisions, and if unilateral decisions are made, the people affected are not going to respond well. Going forward, the researcher believes that decisions need to be made in the best interest of the students, but not made in a vacuum. If, as a leader, the researcher believes something would be best for children

and their education, it would be important to explain that in a manner that shows an approach that is not combative, and still use research and data to support why one would think those decisions are decisions that need to be made. If, as a leader, the researcher could convince teachers and community members that his actions are the correct course of action, then it should be taken. If not, then it needs to be revisited and tweaked so that more buy-in can be achieved.

Further, as a leader, the researcher would use the quantitative data collected in this study to help determine where improvements could be made outside of the grading system used as a means to try and increase student achievement scores on summative assessments. The data could be distributed to teachers so they could collaborate and reflect on how to improve in certain areas, what worked, and what did not work, and then make adjustments going forward. The researcher believes that it would be wasteful to use the data collected in this research to only determine the success or lack thereof of a particular system of grading. It should also be used to help indicate where more can be done to help students succeed.

Summary

In summary, the researcher found that there did not seem to be a statistical correlation between the implementation of standards-based grading and a rise in student achievement scores. Further, based on the qualitative data, there also seemed to be a strong distaste for standards-based grading among the teachers and the community as a whole, indicating that the roll out for the new model of grading proved to be unsuccessful. The bottom line is that it did not appear as if SBG played any significant role in improving grades, and given the push back from the teachers and the community,

it would seem wise to stay away from a full-blown implementation of SBG, instead deferring to a version that meshes the two grading systems together in an effort to create a scale that shows what SBG shows while still reiterating the aspects of grading that the community wants and that the teachers want. Ultimately, these are the people that will be most affected by what is decided, so it stands to reason that a version would be put together that would essentially take the best of both to develop the best solution for secondary schools in this community.

The researcher chose this topic because grading was such an important part of learning. For many years, grading was an arbitrary practice. Not much thought was given to grading and grading scales, and teachers were free to grade in whatever manner they wished. Many would use a total points system, where the total number of points was divided by the total number of points earned to give a percentage. This system was all over the place; some teachers could count homework as 40%, others not at all. Some could count projects as 80%, some would count them as 25%. Some teachers would count tests as 50%, some would count them as 90%. Some teachers would count extra credit, others would not. The possibilities of the traditional points system were almost endless, and therefore no one class was alike in any way, shape, or form. This meant that for decades, students in the same school could have different teachers for the same class and yet somehow have completely different grades. The only indicator of this was a letter (A, B, C, D, or F) and that would have long term, lasting impacts on a student's grades. The researcher had always wondered why more emphasis was not placed on effective grading practices and various ways to grade when in the undergraduate portion of a teaching degree. If it made sense to spend time teaching about classroom

management, it seemed that it would also make sense to teach young teachers about grading and how to go about doing it. Unfortunately, that was not the case in most undergraduate programs in this country. So, the researcher was always left to wonder what the most effective way to grade would be. Standards-based grading offered an opportunity to develop this area of education in a way that had not been done before. However, there were shortcomings to the system, as there are with any new system. So, the researcher had his interest peaked in this once the district he works in (and the subject of this study) began to implement standards-based grading at the secondary level. As a result of the push back from this top-down implementation, many teachers became jaded, and many retired. Many thought SBG was caving to generational demands to ensure that everyone passed and graduated, no matter the cost. Many pushed back because they did not want to change what they had known for the entirety of their careers. Moreover, many just simply did not think that SBG was the right thing to do for the students in the school.

The researcher wanted to get to the bottom of these questions and concerns to see if there was any legitimacy to the claims on either side: the people that supported the implementation or the people who did not support the implementation. Much of the research on SBG that existed at the time of this writing dealt with SBG and how it existed in the elementary classroom, but not so much on the effect it had on the secondary classroom, and the students that inhabited that space. Further, the researcher wanted to identify what, if any, impact the implementation of standards-based grading would have on a student who was getting ready to head to a post-secondary institution and how his or her familiarity with SBG would translate to the more traditional form of grading at the

college level. The researcher was curious as to how well students would be prepared, and whether universities would accept some of the basic tenets of standards-based grading. The researcher wanted to find out if all of the trouble the district went to as a means to implement this new grading system was actually worth it, or if it was an exercise in futility. This simple question led to the research laid out in this study and that goes on to show that SBG did not, in fact, statistically contribute to an increase in student achievement, as based on EOC scores at the secondary level. The researcher wanted to provide hard evidence that either supported or did not support the claim that SBG was not working in the secondary schools, as there did not seem to be much that existed in the form of then-current research that answered those questions definitively. The researcher wanted to be able to go back to both sides of this debate and show that both sides had the best interests of the students in mind, but that there was room within both grading systems for some good and for some bad, and that a cross of the two systems would, more than likely, prove to be the answer for a majority of the people involved.

Recommendations to the District

Although the research completed here indicated that there seemed to be no correlation between the implementation of standards-based grading and a rise in student achievement, it would seem that more research needs to be completed on the impact it has. As the practices involved in the implementation process are refined, and as parents and teachers begin to see the inherent benefit of recognizing specific areas of a subject that would need remediation, the scores may start to rise. School districts, large and small, nationwide should continue to research the benefits of changing how they assess

and report students' achievement. As districts continue to research these methodologies, there may be unintended consequences that are beneficial to the success of the student.

Further, school districts should allow teachers to experiment with different variables as to how a particular grade is achieved. Allow some teachers to formatively assess and include those grades while others keep them out, and then encourage those varying groups to collaborate to see what works better for the students they teach and how far the community is willing to go to see that change. By doing this, individual school districts would be able to implement standards-based grading, but do so in a way that caters to the needs of their individual framework and what their community expects of them. This allows SBG to form a standard template to follow that could subsequently be adjusted to account for individual needs and wants, while also providing classroom autonomy to the instructor based on the needs and wants of that instructor.

That being said, educators and parents alike should not take the information presented in this study and discard it because of the above statements. On the contrary, the information and research provided in this study should go a long way in helping teachers, administrators, and parents work hand-in-hand in determining how they can make their respective districts successful while at the same time learning from the mistakes of the district that is the subject of this study. Others should be more cognizant of the effect that change might have on a large group of teachers and parents, particularly when one does not value their voice in the beginning stages of the implementation. If those stakeholders, parents, and teachers alike, do not feel as if they are being heard or do not feel as if their concerns about the grading change are being addressed, they are going to be more likely to resist that change. Consequently, districts utilizing SBG after this

study should incorporate the ideas, needs, and wants into the plan to the extent that those ideas, needs, and wants do not water down the grading scale as it was intended to be designed. Teachers should take their own sample sizes (within their classrooms) and find what variation of SBG works for their class. Teachers can do this formally or informally, utilizing conversations with students and parents, as well as electronic surveys and end-of-course reviews. Teachers can then use this feedback to mold a grading structure for their class based off the district template of what could be called 'district non-negotiables.'

Other districts may want to wait and see what future research provides in terms of assessing how SBG works to improve student achievement and determine if a longer timeframe might be necessary to really see the impact a switch like this might really have. These same districts may also want to experiment with various drafts of report cards and how best to communicate the assessments to the parents and students. Districts implementing SBG tread a fine line between putting too much and too little information on a report card. If the district ends up putting too much information on a report card, that district runs the risk of parents not reading through everything, as well as alienating the teachers who have helped bring the district to this point. If a district makes the reports cards more in-depth and thorough, that obviously requires more work for the teachers. The flip side to this is that the district also runs the risk of not putting enough information on a report card, and then people will end up wondering why the change was made in the first place. If the goal is to better communicate the knowledge a student has learned and retained, then the report card should reflect that without being too unbelievably lengthy. In addition to an easy-to-read report card, districts going forward

should communicate how formative and summative assessing works and how mid-term grades would be calculated, while focusing on the impact of test retakes and how they can have a significant impact on a grade. Essentially, under SBG, students and parents have to erase the old total point mindset where one would have a general idea of where one stood in terms of a letter grade at any given point during the semester. Under SBG, it is more difficult to quantify a grade, per se. Rather, under SBG, one would be able to let a student and parent know that the grade is fluid and that the student has tests to make up. Ensure that they understand that if the tests are not re-taken under the classroom re-take guidelines, that student will almost certainly fail the class because they did not demonstrate a full understanding of the material and were not able to retain that information. If this is detailed and explained early in the process, districts should not run into many, if any, issues as it relates to communication. Keep in mind that one of the primary reasons a district would change to SBG would be to better communicate progress to both the student and the parent as opposed to presenting that parent and child with an arbitrary grade. However, parents and students at the secondary level still expect and need to see a letter grade, primarily because of transcript issues relating to providing that transcript to colleges, universities, military, etc. So, the district is going to want to find a way to address the conversion of an SBG grade into a letter grade. However, while doing this, do not make it the primary focus, but rather the secondary focus, of the design. Put the standards on the first page of the report card, and the grades on the back. Have grades in a smaller font than the standards, or only have teachers turn in grades to the registrar for record keeping and only send home a report card with the standards attached. Have these standards be specific and concrete, and be sure the instructors do not write

insignificant words in the columns next to the grade. Rather, ensure that teachers are taking the time to write out thoughtful statements about the students so as to give the student and the parent a better insight as to how that student is doing in class, and what, if any, interventions need to be given

Recommendations for Future Research

The scope of this study spans five years and covers four different high school subjects. During those five years, the transition from traditional grading to standards-based grading occurred in full. What this means is that in year one of the data, secondary teachers could not have explained what SBG was, and by year five of the data, those same teachers had been inundated with professional development about SBG and had completely transferred to an SBG model of instruction. During this transition, as previously noted, there may have been a lot of corner cutting by teachers to get their old style of grading to fit into this new style of grading. In other words, times were choppy and not much was settled in the way of instruction. This brings about the thought that in the future, when other researchers and districts study standards-based grading, they may be wise to do this in a timeframe as far away as possible from that direct center time of change. Districts should contemplate, given the time to do so, collecting their pre-SBG data well before the implementation and talking occurs, and then collect their post-SBG data several years after the implementation so as to better see and understand the longer-term impacts that SBG would have on the students and community as a whole. A longer duration study could identify more factors that impact student achievement and help seek to improve those factors. A longer duration study is also going to provide more validity than a shorter one as it will close many of the gaps that could potentially exist in a shorter

duration study, particularly as it relates to the benefits of making a course consistent across all teachers who might teach it. A longer duration would show, or attempt to show, that one of the primary virtues of SBG is due in large part to the fact that a class is based on concrete standards, any teacher could come teach that class and have the same impact on student achievement. Prolonging a study would allow for more study into how continuity plays a role in increasing student achievement, something this study comes up short on.

Further research on the impact of standards-based grading could focus on the relationship secondary schools have with post-secondary schools and how SBG would translate to those post-secondary institutions. One of the most prominent complaints about SBG is that it does not adequately prepare a student for what college is like, or what the job force is like. The general belief, as it relates to college, is that the student gets one shot, and if the student blows it, the professor is not going to allow the student a retake. The other prominent belief is that in college, the students are largely responsible for their own learning, in particular where they have undergrad seminar classes that could number 200-300 students in one class. The lingering question remains, and it is one that this particular study did not address, of what to do as a means to make that transition from high school to post-high school as smooth as possible while at the same time preparing them for a life outside of high school. So, with that in mind, it would be to the benefit of districts and researchers going forward to study how colleges and universities are grading their students, and what they have changed over the years so as to better shape SBG implementation to reflect the expectations of local universities. High schools and colleges should be collaborating on the most effective means of monitoring student

progress and communicating that progress, while at the same time familiarizing those students with what life is going to be like once they leave the confines of their high school. Universities need to grapple with the notion that historically, how they grade (a total points system based largely on just a few summative assessments per semester) is seemingly inadequate when it comes to showing what students have mastered, if anything, in college. These colleges should be on board to find a means of grading that focuses not so much on a letter grade and less on knowledge and retention, but more so a grading system that encourages students to focus less on a particular letter and more on learning the knowledge that is being taught, and then retaining that knowledge to use in their chosen careers, having now been better prepared.

Another area that should be of benefit for future researchers is the area of standardized tests beyond the EOC exam that was studied in this research. Research could be continued as it relates to college entrance exams such as the ACT and SAT tests. It would seem beneficial to look into how the implementation of standards-based grading across the board would correlate, if at all, with a rise in these standardized tests, which could prove beneficial to college-bound students. Going even beyond this application, research could also be conducted to determine whether standards-based grading correlates to a rise in Advanced Placement (AP) scores aggregated from all applicable AP courses. As the College Board, that manages and scores the AP exams, starts to head toward a more ‘standards-based’ approach to the classes it audits, this could have potentially positive implications as there will be a large population to study based on before and after scores. The data this provides would be substantial given the depths to which the College Board goes to ensure continuity over time and security on tests

nationwide. This type of research would take into account many different school districts from many different demographically diverse backgrounds, allowing the researcher to gain a broader perspective on the impact the standards approach has taken relative to past results. A researcher in this type of study could also do a qualitative aspect by surveying the chosen readers for the AP exams, the teachers who come from around the entire nation to score the essays of the students who have taken the exam. Again, this brings diversity into the mix and provides a better opportunity to detail more accurate results as it relates to standardized tests rather than basing the study solely off of one school district and two high schools.

As this study did not focus on the special education aspect of SBG, educators that work in special education might be very curious to learn how the implementation of standards-based grading correlates to an increase in student achievement with students who have individualized educational plans (IEPs) and what, if any, benefit there would be to this particular population of students being involved in a class that structured itself off of SBG. Many IEPs are already focused around creating standards and setting goals for students, so this population of student should be somewhat familiar with this style of assessment, as SBG takes into account all knowledge learned rather than looking simply at a total point aspect. So, based on this, SBG would seem to be a good fit for our SPED populations. Future researchers may want to determine whether SBG does have the impact on SPED that one would think it would have, and also determine whether it would be more reflective of what those students are actually learning. This research could use IEP goals as qualitative data to determine the correlation, or standardized test results could also be used as to the quantitative aspect to show any correlation between SBG and

SPED. However, given the private nature of most IEPs, accessing these for any sort of qualitative aspect of a study could prove difficult without the inherent cooperation of a school district.

One of the other arguments in support of standards-based grading contends that SBG is a worthwhile venture due in large part to the fact that it brings more consistency to education, allowing teachers to collaborate and develop formative and summative assessments that are identical in like classes, so that a student taking a Biology class at one high school in one district would get the exact same curriculum and have the exact same expectations if that student were to move to the same class in the same school district, but with another teacher. The goal is to provide consistency so that students and parents have a better idea of what is expected of them in achieving academic success in their respective classes. One of the arguments against traditional grading is that it does not account for consistency between like classes as teachers can assign points as they wish and weight grades and categories as they wish. The argument against traditional grading would state that this type of grading gives too much autonomy to the instructor and provides little in the way of increasing student achievement as the grades are arbitrarily based off of a substantial number of variables that the individual classroom teacher controls. So, with respect to consistency, a research study in the future might want to determine what, if any, correlation exists between high academic achievement and consistency among classes. If so, this study could then be utilized as evidence of the positive nature of standards-based grading in lieu of the traditional model. While consistency sounds good as it relates to academic achievement, a researcher would have several different avenues to work with to research whether this assumption is true or not.

A future study could factor in other achievement scores, such as those of ACT results, SAT results, or overall class scoring before and after the implementation of standards-based grading so as to try and achieve a more accurate depiction of the outcomes. That same future study could also integrate a timeline of data that goes beyond the five years attributed to this study so as to catch a broader set of data rather than catching data that happens to come during the transition time frame where tension over the change is going to be at its worst and when there may be the most question about whether SBG implementation was actually taking place in the secondary classroom as it was intended. Further, a future study could also take into account a more substantive point of view from both the teachers and community members with different follow up questions as a means to really dig in and see what specifically bothers those two groups about the SBG implementation and then ask further follow up questions as to what could be done to close the gap between the old system and the SBG system, as a way to have more buy-in from both the staff and the community members.

As for changes to the methodology, the researcher believes that it is not necessary to utilize 6,000 varying pieces of data for the quantitative aspect of the study and would be better off taking smaller pieces of data from more varied sources so as to better see what effect SBG is having across the board, rather than just simply looking at EOC data from that transition period. As for the qualitative aspect of the study with the surveys, the researcher believes that this methodology could change in that the surveys could be geared toward more solution-based outcomes rather than just serving as what essentially amounted to a venue to list complaints about something they did not agree with. In this regard, the surveys did not seem to adequately address the need for change and did not

funnel answers to a conclusion that was productive. The questions could be geared toward forcing the person taking the survey to come to his or her own conclusion about what the best course of action for the students would be going forward; in other words, the survey should attempt to help draw the person to a logical conclusion that is more helpful rather than hurtful. Encourage the respondents to be productive and help be the solution rather than the problem.

The research for this study took place in a mid-sized school district in the Midwest with two high schools numbering roughly 1,500 students between the two of them. So, going forward, a researcher could certainly argue that the results would be better served by having multiple school districts represented with a large diversity of student backgrounds. This would be relevant in that a study based on those factors could give the researcher more ground to stand on, so to speak, when determining the outcome of the study. Going forward, a researcher could certainly look into researching how demographics plays a role in levels of student achievement and whether standards-based grading significantly impacts student achievement in lower performing versus higher performing schools. This would be an interesting study as it would not only look at the correlation between student achievement and SBG but would also look at whether the type of grading system utilized had any significant impact on grades and student outcomes in a school stricken with poverty versus one with a wealthier demographic. In other words, is the grading system utilized more important than the backgrounds of the students that are using it?

Conclusion

This mixed methods study investigated the relationship between standards-based grading implementation at the secondary level and increases in student achievement based on end-of-course exam scores and survey results about SBG from both parents and staff. Secondary data were collected from a middle-sized school district located in a suburban area in the Midwest with a diverse demographic. The secondary data examined were based on five years of EOC data from two high schools within the same district where the teachers had collaborated to develop the standards in like classes where EOC exams were the end result. Quantitative data were used with the district's permission and provided by them through the Missouri Department of Elementary and Secondary Education in a manner that kept the students' identification private. This secondary data only indicated to the researcher what the subject was, the grade level, and the outcome of the EOC for collection purposes. For the purposes of answering the one hypothesis for this study, a t-test of means was utilized. This test, based on the data that was input, indicated no statistically significant correlation between the implementation of standards-based grading and student achievement scores.

The qualitative data used for this study was also based on secondary data, taken from the same school district as noted above. This school district administered two surveys: one survey for parents and one survey for staff. The parent survey had 343 total responses and the staff survey had 88 total responses. This survey was conducted by the school district during the course of implementing SBG as a means to gauge community and staff interest and to see what they believed were the positives and negatives of SBG. This study then utilized that qualitative data to determine whether the general mood and

perception toward SBG matched the outcomes of the quantitative portion of the study. In other words, are the concerns of the parents and staff as indicated by the surveys aligning with the actual outcomes that are being produced in the classroom, so as to better assess how to work with the community and staff going forward through this transition process? After coding all of the qualitative data and applying it to the two research questions utilized in this study, it became apparent that an overwhelming majority of both parents and staff believed that implementing standards-based grading at the secondary level was a mistake, to say the least. Their responses questioned the wisdom of changing the grading scales, as well as how the change was occurring. Many respondents felt as if the change was forced on them with no input early in the process as to how they believed it would affect students in the classroom and how it could be worked to adequately prepare them for college.

Based on the results from both the qualitative and quantitative aspects of this study, it seems that the perception from staff and parents matches the results from the student achievement portion. Namely, this means that in large part, the parents' and teachers' concerns about standards-based grading are well founded, and do not seem to be rooted in a general distaste for change. The concerns of the community and staff match the quantitative results, which should allow for a more productive conversation going forward as it relates to any academic changes in the future. Essentially, this seems to be a case study in what not to do when attempting a large-scale change in a school district. Although the motivation of higher student achievement was well intentioned, the results of that well-intentioned action do not seem to warrant the expense, i.e., the equity that was lost with staff who felt like they were left out of the process as well as the equity

lost with the community as it relates to their trust in the academic success of their children.

However, with the above being duly noted, the intent of the district to try something to improve student achievement was an exercise that needed to happen, as it certainly ended up having the unintended consequence of starting a district-wide conversation on how to better implement policies, as at the end of the day, all parties want to see a focus on students and a focus on increasing their ability to learn. As this was the case with this district, district leadership learned (due in no small part to the surveys referenced in this study) that they needed to reevaluate their policy on SBG and rework some of the major sticking points that were mentioned in the surveys. Those issues were subsequently addressed, and a modified version of SBG was implemented at the secondary level that is still a work in process, but a well-received work in process that the majority of staff and parents had a role in creating.

Although the research of this study indicated no statistical correlation between the implementation of standards-based grading and a rise in student achievement, the researcher is still proud of the work conducted and believes that there is inherent value in many of the strategies promoted by having a standards-based curriculum, namely the idea of consistency among teachers and a more formal approach to having teachers learn more about assessment as it plays out in the real world. In follow-up conversations with members of the district after the initial data were collected, many were pleased with how the district corrected the perceived mistakes of the initial implementation and worked in a more collaborative approach with the community and staff.

Going forward, more research will need to be done on whether changing grading styles lends itself to a more effective means of increasing student achievement, but the researcher believes that any effort to correct the traditional grading system and not make it as subjective is a good thing. More research is bound to occur on a topic of such debate, and as districts continue to trial various applications of SBG, more research will be completed on what works and what does not work until we are able to refine it and create a much more objective approach to assessment that makes sense and can be viewed globally as the correct method of grading. At the time of this writing, educators are currently in the early stages of looking at how to change something that has been embedded in the country's educational conscience for over a century, and that type of history does not simply change overnight without a few figurative bumps in the road. This will be a collective effort on the part of undergraduate teaching programs, parents, teachers, administrators and the students over generations as they seek to bring the United States back to the forefront of global education.

References

Atkinson, R. C., & Geiser, S. (2009). Reflections on a century of college admissions tests. *Educational Researcher, 38*(9), 665-676.

Buckmiller, T., Peters, R., & Kruse, J. (2017). Questioning points and percentages: Standards-based grading in higher education. *College Teaching, 4*(65), 151-157.

Brookhart, S. (1993). Teachers' grading practices: meaning and values. *Journal of Educational Measurement, 30*(2), 123-142. doi:10.1111/j.1745-3984.1993.tb01070.x

Brookhart, S., Guskey, T., Bowers, A., McMillan, J., Smith, J., Smith, L., & . . . Welsh, M. (2016). A century of grading research. *Review of Educational Research, 86*(4), 803-848.

Board of Education v. Rowley. (1982). 458 U.S. 176. Retrieved from https://supreme.justia.com/cases/federal/us/458/176/

Bowers, A. J. (2010). Analyzing the longitudinal K-12 grading histories of entire cohorts of students: Grades, data driven decision making, dropping out and hierarchical cluster analysis. *Practical Assessment Research and Evaluation, 15*(7), 1-18.

Camara, W. (1998). *High school grading policies*. The College Board, Research Notes No. RN-04. Retrieved from https://research.collegeboard.org/sites/default/files/publications/2012/7/researchnote-1998-4-high-school-grading-policies.pdf

Deddeh, H., Main, E., & Fulkerson, S. (2010). Eight steps to meaningful grading. *Phi Delta Kappan, 91*(7), 53-58.

Dueck, M. (2011, November). How I broke my own rule and learned to give retests. *Educational Leadership, (3)*69, 72-75.

Edgeworth, F. Y. (1888). The statistics of examination. *Journal of the Royal Statistical Society, 51*(3), 599-635.

Etscheidt, S. (2006). Behavioral intervention plans: Pedagogical and legal analysis of issues. *Behavioral Disorders, 31*(2), 223-243.

Falk, B. (2002). Standards-based reforms: Problems and possibilities. *Phi Delta Kappan, 83*(8), 612-620.

Fisher, D., Frey, N., & Pumpian, I. (2011, November). No penalties for practice. *Educational Leadership, (3)*69, 46-51.

Godfrey, K. E. (2011). *Investigating grade inflation and non-equivalence*. The College Board, Research Report No. 2011-2. Retrieved from https://research.collegboard.org/sites/default/files/publications/2012/7/researchreport-2011-2-investigating-grade-inflation-non-equivalence.pdf

Grodner, A., & Rupp, N. (2013). The role of homework in student learning outcomes: Evidence from a field experiment. *Journal of Economic Education, 44*(2), 93-109.

Guskey, T. R. (2004). The communication challenge of standards-based reporting. *Phi Delta Kappan, 86*(4), 326-329.

Guskey, T. R. (2007). Using assessments to improve teaching and learning. In D. Reeves (Ed.), *Ahead of the curve: the power of assessment to transform teaching and learning* (pp. 14-29). Bloomington, IN: Solution Tree.

Guskey, T. R. (2009). *Practical solutions for serious problems in standards-based grading.* Thousand Oaks, CA: Corwin.

Guskey, T. R., & Bailey, J. M. (2001). *Developing grading and reporting systems for student learning*. Thousand Oaks, CA: Corwin.

Handler, B. R. (2003, April). *Special education practices: An evaluation of educational placement trends since the regular education initiative.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Hanover Research. 2015. *Best practices for Standards-Based Grading*. Washington, DC: Hanover Research.

Hattie, J. A. (1992). Measuring the effects of schooling. *Australian Journal of Education, 36*(1), 5-13.

Hattie J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Heflebower, T., Hoegh, J., & Warrick, P. (2014). *A school leader's guide to standards based grading*. Bloomington, IN: Marzano Research Laboratory.

Hirst, B. (1973). *A project for the development of a multi-state consortium for the production of performance objectives and criterion measures in occupational education*. Second Report No. CE009459. Retrieved from http://www.eric.ed.gov/PDFS/ED133515.pdf

Hooper, J., & Cowell, R. (2014). Standards based grading: history adjusted true score. *Educational Assessment, 19*(1), 58-76.

Individuals with Disabilities Education Act. (1997). 20 U.S.C § 1400 to 1491. Retrieved from https://www.gpo.gov/fdsys/granule/USCODE-1996-title20/USCODE-1996-title20-chap33-subchapI-sec1400

Individuals with Disabilities Education Improvement Act. (2004). 20 U.S.C § 1400 to 1482. Retrieved from https://www.gpo.gov/fdsys/granule/USCODE-2011-title20/USCODE-2011-title20-chap33-subchapI-sec1400

Kohn, A. (2011, November). The case against grades. *Educational Leadership, (3)*69, 28-33.

Kralovec, E., & Buell, J. (2000). *The end of homework: How homework disrupts families, overburdens children, and limits learning.* Boston, MA: Beacon Press.

Jensen, E. (2000). *Brain-based learning.* San Diego, CA: The Brain Store.

Jung, L. A., & Guskey, T. R. (2007). Standards-based grading and reporting: A model for Special Education. *Teaching Exceptional Children, 40*(1), 48-53.

Marzano, R. (2000). *Transforming classroom grading.* Alexandria, VA: ASCD.

Marzano, R. (2007). Designing a comprehensive approach to classroom assessment. In D. Reeves (Eds.), *Ahead of the curve: the power of assessment to transform teaching and learning* (pp. 102-125). Bloomington IN: Solution Tree.

Marzano, R. (2010). *Formative assessment and standards based grading.* Bloomington, IN: Marzano Research Laboratory.

Marzano, R., & Heflebower, T. (2011, November). Grades that show what students know. *Educational Leadership, (3)*69, 34-39.

Missouri Department of Elementary and Secondary Education. (2017). *St Charles R-VI School District Quick Facts* [data file & demographic data]. Retrieved from http://mcds.dese.mo.gov/quickfacts/sitepages/districtinfo.aspx?ID=__bk810003009300230000300930003000

Munoz, M. A., & Guskey, T. R. (2015). Standards-based grading and reporting will improve education. *Phi Delta Kappan, 96*(7), 64-68.

O'Connor, K. (1999). *The Mindful School: How to grade for learning.* Arlington Heights, IL: Skylight Professional Development.

O'Connor, K. (2002). *How to grade for learning: Linking grades to standards* (2nd ed.). Thousand Oaks, CA: Corwin Press.

O'Connor, K. (2007). The last frontier: Tackling the grading dilemma. In D. Reeves (Ed.), *Ahead of the curve: The power of assessment to transform teaching and learning* (pp. 127-145). Bloomington, IN: Solution Tree.

O'Connor, K. (2017, January). A case for standards-based grading and reporting: An expanded format report card shares progress on academic achievement and work habits beyond a numerical mark. *School Administrator, 74*(1), 24.

Pattison, E., Grodsky, E., & Mueller, C. (2013). Is the sky falling? Grade inflation and the signaling power of grades. *Educational Researcher, 42*(5), 259-265.

Peters, R., Kruse, J., Buckmiller, T., & Townsley, M. (2017). "It's just not fair!" Making sense of secondary students' resistance to a standards-based grading. *American Secondary Education, 45*(3), 9-28.

Peters, R., & Buckmiller, T. (2014). Our grades were broken: Overcoming barriers and challenges to implementing standards-based grading. *Journal of Educational Leadership in Action, 2*(2), n.p.

Ratnesar, R. (1999, January 25). The homework ate my family. *Time, 153*(3), 54-63.

Reeves, D. (2004). *Making standards work: How to implement standards-based assessments in the classroom, school and district.* Englewood, CO: Advanced Learning Press.

Rauschenberg, S. (2014). How consistent are course grades? An examination of differential grading. *Education Policy Analysis Archives, 22*(92), 1-38.

Sawyer, R. (2013). Beyond correlations: Usefulness of high school GPA and test scores in making college admissions decisions. *Applied Measurement in Education, 26*(2), 89-112.

Scherer, M. (2001). How and why standards can improve student achievement: A conversation with Robert J. Marzano. *Educational Leadership, 59*(1), 14-18.

Schneider, J., Feldman, J., & French, D. (2016). The best of both worlds. *Phi Delta Kappan, 98*(3), 60-67.

Scriffiny, P. L. (2008, October). Seven reasons for standards-based grading. *Educational Leadership, (2)*66, 70-74.

Shippy, N., Washer, B., & Perrin, B. (2013). Teaching with the end in mind: The role of standards-based grading. *Journal of Family & Consumer Sciences, 105*(2), 14-16.

Spencer, K. (2012, October). Standards based grading: New report cards aim to make mastery clear. *Harvard Education Letter, 28*(5), 3-6.

Stephens, E. (2010). *7th-12th Grade English/Language Arts teachers and their classroom grading practices: Investigating the use of standards-based grading in Nebraska's rural classrooms.* [Doctoral dissertation]. Retrieved from http://search.proquest.com/docreview/807680640

Stiggins, R. J., Arter, J., Chappuis, S., & Chappuis, J. (2004). *Classroom assessment for student learning: Doing it right, using it well.* Portland, OR: Assessment Training Institute.

Taylor, K. R. (2003). When students sue teachers over grades. *Education Digest, 68*(6), 58-63.

Tippin, G., Lafreniere, K., & Page, S. (2012). Student perception of academic grading: Personality, academic orientation, and effort. *Active Learning in Higher Education, 13*(1), 51-61.

Townsley, M. (2017). Three lessons for schools shifting their grading. *School Administrator, 74*(1), 27.

Tucker, C. (2018). Rethinking grading: Making mastery, not points, the reward for your students. *Educational Leadership, 75*(5), 84-85.

Vatterott, C. (2009). *Rethinking homework: Best practices that support diverse needs.* Alexandria, VA: ASCD.

Waldron, N. L., & McLesky, J. (1998). The effects of an inclusive school program on students with mild and severe learning disabilities. *Exceptional Children, 64*(3), 395-405.

Wildman, P. R. (1968). Homework pressures. *Peabody Journal of Education, 45*(4), 204.

Wormeli, R. (2011, November). Redos and retakes done right. *Educational Leadership, 69*(3), 22-26.

Vitae

Steven W. Smith

Professional Experience

Assistant Director, Lewis and Clark Career Center	2016 - Present
Administrative Intern, St. Charles West High School	2015 – 2016
Social Studies Teacher, St. Charles West High School	2007 – 2015
Social Studies Department Chair, St. Charles West High School	2013 – 2015
Building PD Committee, St. Charles West High School	2009 – 2015
St. Charles School District PD Committee	2009 - 2010
St. Charles West School Improvement Team	2009 – 2015
St. Charles School District Facility Assessment Committee	2012 - 2013

Education

Doctorate in Educational Administration, Lindenwood University	December 2018
Master of Arts in School Administration, Lindenwood University	December 2011
Master of Arts in Teaching, Lindenwood University	August 2006
Bachelor of Arts in Interdisciplinary Studies, University of Missouri – Columbia	December 2002

Professional Development Activities and Honors

First Year Teacher of the Year, Students' Choice	2007 – 2008
Marzano Standards Based Grading Committee	2011 – 2015
Marzano Standards Based Grading Workshop	2012
MASSP Aspiring Principals Workshop	February 2014
St. Charles West Commencement Keynote Speaker, Students' Choice	2012 & 2016