Investigating the Relationship of Standards-Based Grades vs. Traditional-Based Grades to Results of the Scholastic Math Inventory at the Middle School Level

Erin Lehman
Rapid City Area Schools

David De Jong
University of South Dakota

Mark Baron
Wayne State College

Grading is one of teachers' greatest challenges and most important professional responsibilities. Educators are unclear on whether standards-based grades or traditional-based grades do a better job of accurately reflecting what students have learned, so the purpose of this study was to understand the relationship between classroom grades and scores on the Scholastic Math Inventory (SMI) assessment. The individuals were sixth-, seventh-, and eighth-grade mathematics students from five different middle schools in the same district as they took the SMI assessment. There were about 500 students in the standards-based grading system and about 1,900 students in the traditional grading system.

Education Leadership Review of Doctoral Research, Vol. 6. Fall/Winter 2018
ISSN: 1532-0723

The purpose of assigning grades to student learning varies from educator to educator. Wormeli (2018) explains that allowing educators to decide how to assign grades to student learning may not be an accurate picture of what students know, understand, and be able to do. Teachers' beliefs in what to grade and to record are significant influences on what a student earns as a grade (Brookhart, 2017; Guskey, 2015; O'Connor 2011). O'Connor (2011) stated, "Teachers develop assessments based on their professional judgment of what is to be assessed and how–a subjective process" (p. 11). Grading practices at the middle school level must change in order to "meet the learning needs and desires of future generations of young adolescents, the core middle school practices must continue to grow and thrive" (Schaefer, Malu, & Yoon, 2016, p. 18).

Traditional grading systems utilize an A, B, C, D, or F or similar scales to denote student understanding of all content standards. In contrast, standards-based grading systems utilize a reporting system based on individual content standards (Guskey, 2015; O'Connor, 2011; Heflebower, Hoegh, & Warrick, 2014). Standards-based grading centers on specific learning standards or goals (Guskey, 2015; Hanover Research, 2014; O'Connor, 2011). "More and more educators are beginning to question traditional grading practices that were developed to sort students into learners and non-learners, not to support learning for all" (Brookhart, 2011, p. 10).

The grading and reporting system have changed throughout the years, most notably the traditional and standards-based grading system. The traditional grading practice tends to lump content with effort and behavior into one letter grade (Brookhart, 2011b). All too often, traditional grading practices in the United States are based on instructional and motivational principles that cause many students to give up in hopelessness and accept failure rather than driving them toward academic success (Stiggins, 2014).

When the standards-based educational reform began in the 1990s, the goal was to replace learning from basic facts to synthesis and application (Shepard, 2009). More recently, the Common Core State Standards for mathematics emphasized conceptual understanding as well as procedural skills and increased rigor of content (Briars & Foster, 2012). Content standards and common assessments improved consistency and coherency in curriculum and instruction, but grades and grading remained in the hands of the individual teacher (Wormeli, 2018).

To make grades more meaningful, issues related to both purpose and reporting format must change as well (Guskey, 2015). Grading often remained subjective. O'Connor (2009) stated,

> There is clearly no right answer or perfect grading plan, but for those who teach the same grade or course(s) in the same school and, ideally, in the same school district, it would not be unreasonable to expect that there would be some basic similarities or that discernible patterns would exist across their grading plans. (p. 33)

Grading within mathematics classrooms can be subjective and personal, which leads to inequalities in grading all across the United States (Guskey, 2015). Over the years, mathematics teachers have decided which criteria they should or should not report in terms of student learning: "For more than a century, grades have remained the primary indicator of how well students performed in school and the basis for making important decisions about students" (Guskey, 2015, p. 3). These traditional grading practices persist, resulting in ineffective communication about student achievement and can potentially undermine students' lifelong attitudes towards learning (Bourgeois & Boberg, 2016, p. 15).

Purpose of the Study

Educators are unclear on whether standards-based grades or traditional-based grades do a better job of accurately reflecting what students have learned, so the purpose of this study was to understand the relationship between classroom grades and scores on the Scholastic Math Inventory (SMI) assessment. We investigated the relationship that exists between a traditional or standards-based grading system and achievement on the SMI assessment.

Research Questions

Two research questions guided this study.

1. What is the relationship between SMI scores and end-of-year grades for sixth, seventh, and eighth grade students in a traditional grading system from four middle schools?
2. What is the relationship between SMI scores and end-of-year grades for sixth, seventh, and eighth grade students in a standards-based grading system from one middle school?

Significance of the Study

In order to determine the relationship of standards-based grades vs. traditional-based grades to results of the SMI, the researcher investigated the relationships of grades under a standards-based grading system and a traditional grading system to the SMI scores. The school district is unique in that fact they had one school with a standards-based grading system and the other four schools used a traditional grading system in the middle school mathematics program. This study's result could significantly persuade or dissuade the future implementation of a standards-based grading system at the middle school level in mathematics classrooms.

Literature Review

The current reality of education reflects an era of educational accountability, which drives a new way of viewing grading practices. Leaders of mathematics education must understand the standards-based grading reform, common core state standards, and mathematical accountability to make decisions that are best for the students learning in this era of educational accountability.

Standards-Based Grading Reform

Through the standards-based reform that took place during the 1980s and 1990s, there was a shift from basic skills to higher standards and assessments that required higher order thinking skills and complex performances (Brookhart, 2013). The 1983 *A Nation at Risk* report claimed United States students were falling behind their international counterparts. This study recommended that schools, colleges, and universities adopt more rigorous and measurable standards. It further recommended higher expectations for academic performance and student conduct, and that four-year colleges and universities raise their requirements for admission (NCEE, 1983). During this era of educational accountability based on assessing student achievement of standards became firmly entrenched in the public's mind, along with support for achievement testing and making comparisons (Brookhart, 2013). By the mid-1990s, most states had drafted a set of standards.

Mathematical Accountability

Research shows that students who are not successfully mastering mathematical concepts tend to demonstrate slow or inaccurate retrieval of basic mathematical facts, lean toward impulsivity when solving problems, and have difficulty forming mental representations of mathematical concepts or keeping information in working memory (Heflebower, Hoegh, & Warrick, 2014).

Students need both procedural and conceptual knowledge in order to learn and understand mathematics (NCTM, 2014). Knowledge of the procedures and formulas are critical to overall proficiency in mathematics, especially when individual students learn mathematical strategies (Hofman, Visser, Jansen, Marsman, & van der Maas, 2018). Also, exploration of the concepts through concrete experiments and manual manipulation is vital to students' overall understanding of the "why" in mathematics instruction. It is necessary to provide focused instruction that moves students from the concrete to the abstract and then to the application of the concept (Marzano et al., 2003).

Grading practices teachers use may also jeopardize the reliability of grades and weaken the link between grades and academic achievement (Welsh, D'Agostino, & Kaniskan, 2013). With the movement toward 21st Century Skills emphasizing creativity, critical thinking, and communication towards rigorous tasks, educators must change the way they historically have assessed. Fundamentally, more accurate grading practices must be adopted to directly capture more complex levels of achievement: the ability to solve non-routine problems, to analyze data and reason from evidence, to communicate effectively both orally and in writing, and to frame and conduct scientific investigations (Shepard, 2009).

Essence of Grading

Simon and Bellanca (1976) emphasized how grades are a key mechanism in the political processes of schooling, which differentially sorts students according to compliance in the form of academic performance and behavior. Grading originally determined which students continue to the next level, and eventually was a sorting mechanism that allowed educators to rank students and establish classroom curves and hierarchies (Brookhart, 2011a; Dilendik, 1978). Tyack and Tobin (1994) concurred,

> The graded elementary school–in which the curriculum is divided into yearlong batches, students are sorted according to academic proficiency and age, and individual teachers instruct them in self-contained classrooms–is now so familiar that it is hard to imagine a time when it did not exist or to conceive of alternatives. (p. 457)

Purpose of Grading

Purposes differ when it comes to grading. Parents' major focus is often on classroom grades, report cards, and honor roll (Reeves, 2011). Reeves (2011) described a study conducted by Fairfax County Public Schools that found that 89% of colleges use grades to compare applicants, 39% of colleges require a minimum grade-point average, and 33% of colleges require a minimum grade-point average for merit scholarships. When it comes to grading and reporting, stakeholders illustrate a wide variety of need and purpose. Wormeli (2006b) found six reasons for grading:

- To document student and teacher progress
- To provide feedback to the student and family, and the teacher
- To inform instructional decisions

- To motivate students
- To punish students
- To sort students

However, Guskey and Bailey (2010) explained why educators assign grades or marks on students' work:

- To communicate information about students' achievement in school to parents and others
- To provide information to students for self-evaluation
- To select, identify, or group students for certain educational paths or programs
- To provide an incentive for students to learn
- To evaluate the effectiveness of instructional programs
- To provide evidence of students' lack of effort or inappropriate responsibility

While educators may agree that all of these purposes may be legitimate, they seldom agree on which purpose is most important (Guskey, 2015). Guskey and Bailey (2010) suggested identifying who the necessary stakeholders are, otherwise, the communication of the assessments and grades will be unsuccessful in their attempt to meet any of the different purposes. O'Connor (2011) believed "the primary purpose of grades is to communicate about student achievement, with achievement being defined as performance measured against accepted published standards and learning outcomes" (p. 7).

Austin and McCann argued that when educators do not agree on the primary purpose of grades, they often try to address all of these purposes with a single reporting device, usually a report card, and end up achieving none very well (as cited in Guskey, 2015). O'Connor (2009) believed "purpose is like a compass–it provides direction" (p. 15). The basic problem with grades is they serve so many purposes, one letter or number symbol must carry many types of information (achievement, effort, behavior, etc.) in the grade which makes it very difficult to clearly understand what grades mean (O'Connor, 2009).

Allen (2005) suggested that validity is at the heart of effective grading, more specifically, the validity of the learning assessed and the validity of the communication of that assessment to others. Essentially, educators need to make sure assessments are reliable and valid, and that they communicate the results to necessary stakeholders. However, it is easy to overlook the multitude of meanings or purposes assigned to grades. Educators may have one purpose for grades, while parents may feel the grades have a different purpose, and students may ascribe yet another purpose (Brookhart, 2013; Stiggins, 2014). Without validity and a clear purpose, grading loses its usefulness.

Educators at various levels assess for many different reasons. At the instructional level teachers identify the needs of individual students, identify the needs of a class, group students, grade them, evaluate instruction, and evaluate themselves as teachers (Stiggins, 2014). However, assessments need to be reliable and valid (Marzano, 2006). Useful assessments provide teachers with the necessary data to understand which students are struggling in specific areas of the curriculum or which students need enrichment.

Students may learn many things in the classroom, but the primary objective is for students to learn academic content knowledge of a particular subject (Allen, 2005). The major reason for grades then, is to create a public record of the student's academic achievement that can accurately and effectively communicate to others the level of understanding of a subject a student has mastered (Guskey & Bailey, 2010). Wormeli (2006a) claimed,

> A grade represents a clear and accurate indicator of what a student knows and is able to do–mastery. With grades, we document the progress of students and our teaching, we provide feedback to students and their parents, and we make instructional decisions regarding the students. (p. 103)

Traditional Grading

Grades and report cards are a primary source of information about children's learning strengths, areas of struggle, and strategies to promote success at home (Guskey & Jung, 2009). Brookhart (2011b) explained conventional grading practices as one grade that sums up achievement in a subject and that one grade also often includes effort and behavior. Therefore, one letter grade representing achievement actually lumps in behavior and effort. This traditional grading system oftentimes includes averages of assessments, behavior, and other topics teachers include in the grade book (O'Connor, 2011). Instead of being a primary source of information, this system is not often an accurate representation of student knowledge. O'Connor (2009) states, "The focus of traditional grading practices is to sort, select, and justify" (p. 12). A traditional grading system of A, B, C, D, F scales has been dominant in most secondary and postsecondary schools. High schools tend to rely on letter/number grades for calculating Grade Point Average (GPA) and class rank, both of which impact college admission (Wormeli, 2018). However, traditional report cards that record only a single grade for each subject area seldom have detailed information regarding student progress and learning (Guskey & Jung, 2009).

Standards-based Grading

A standards-based report card centered on carefully articulated learning standards provides necessary stakeholders with the specific feedback required to ensure that improvement efforts are appropriately focused (O'Conner, 2017). Standards-based grading assesses students only on their academic performance and proficiency, not on any behavioral factors (Hanover Research, 2014). Marzano (2006) believed the most important purpose of grades is frequent, detailed feedback and, therefore, the best reference point must be specific objectives, standards, or other learning goals in which a standards-based system serves this purpose.

The primary goal of a standards-based system is for all students to "meet standards," that is, to be competent or proficient in every aspect of the curriculum (O'Connor, 2011, p. 2). Standards-based progress reports differ from traditional letter grade, percentage, narrative, or pass/fail report cards by requiring teachers to report student performance levels on specific educational goals instead of broad content areas (Welsh, D'Agostino, & Kaniskan, 2013). A standards-based report card allows teachers to report on nonacademic and academic elements separately (Guskey & Bailey, 2010; Iamarino, 2014). Furthermore, a standards-based report card breaks down each subject area or course into specific elements of learning (Guskey & Bailey, 2010).

A standards-based report card identifies the specific learning goals within the curriculum to ensure appropriate rigor. Today's standards and accountability movement, along with its counterpart, standards-based grading, leads the way to learning-focused grading (Brookhart, 2011a). Iamarino (2014) concludes, "A teacher using a standards-based system of evaluation is better able to determine a student's grade based on the single most important aspect of education–how well the student comprehends the content of the course" (p. 2).

A standards-based grading system also communicates more detailed information about student learning progress with regard to those goals to bring about higher levels of success. O'Connor (2011) believed "the primary purpose of grades is communication about achievement, with achievement being defined as performance measured against accepted published standards and learning outcomes" (p. 7). A standards-based grading system seeks to saturate grades with specific meanings that are easy for students, parents, and teachers to understand (Hanover Research, 2014).

This section examined the different views regarding the purpose and audience for grading. It described the traditional grading practices and standards-based grading practices. With this understanding of the essence of grading, the following section discusses existing grading practices in terms of product, process, progress, and positions.

Existing Grading Practices

This section explains the grading inequalities that show inaccurate measures of student performance. This section also describes existing grading practices in terms of product, process, progress, and positions.

Most classroom teachers today are unprepared to meet the increasingly complex assessment challenges they face in the classroom (Stiggins, 2014). Often, "grades are inferences, personal interpretations on the part of the teacher, not infallible truths about students' mastery" (Wormeli, 2006b, p. 95). The current trend is to place students in harm's way of the ongoing mis-measurement of their achievement in the classroom (Stiggins, 2014). Indeed, Marzano (2000) believed "grades are so imprecise they are almost meaningless" (p. 1). However, teachers often say they are striving to be as objective as possible in their assessment and grading (O'Connor, 2011). Objectivity is difficult to maintain.

Teachers draw from many different sources of evidence in determining a student's grade. Most reporting forms allow teachers to assign only one grade to each student for each subject area or course (Guskey, 2015). This reporting format compels teachers to merge scores from major exams, compositions, projects, and reports, along with evidence from homework, punctuality in turning in assignments, class participation, work habits, and effort (Guskey, 2015). "We err when we attach too much self-worth and celebration to so fleeting a moment, so inaccurate a tool, so subjective an overworked teacher's judgment" (Wormeli, 2006b, p. 95). The product is often "a hodgepodge grade that includes elements of achievement, attitude, effort, and behavior" (Guskey, 2015, p. 74).

Inequalities

McMillan, Myran, and Workman (2002) studied over 900 teachers in grades 3-5 and found that most elementary teachers used a multitude of factors in grading students. The assortment of factors included academic performance, behaviors, grade distributions, norm-referenced grade interpretations, and zeros. The following grading practices prevent grades from being accurate measures of students' performance (Hanover Research, 2011):

- Using a points system and averages
- Using zeros as a punishment
- Grading homework and other formative assignments
- Grading on a curve
- Allowing extra credit
- Grading for behavioral issues
- Incorporating teacher expectations and judgments into grades

Points-based grading puts the focus on numbers, rather than communication. Points are the source for the final grades and often no comprehensive system exists to determine the integrity of the methods utilized to determine the points (Iamarino, 2014). The range in scores is a tremendous source of error associated with the 100-point scale (Marzano, 2010). Averaging in a points system that values all assignments equally can create a situation where a few bad scores inaccurately skew

a student's final grade (Hanover Research, 2014). "Averaging grades, no matter the distance between the two or more scores, decreases accuracy" (O'Connor & Wormeli, 2011, p. 41).

When using the average of all scores throughout the semester, a formula that presumes that the learning early in the semester is as important as learning at the end of the semester discredits the theory of mastery over time (Marzano, 2010; O'Connor, 2009). Interestingly, when teachers and administrators have been students in graduate courses, research shows they routinely insist an evaluation on the basis of their understanding at the end of the semester rather than their work throughout the term (Reeves, 2008). "Percentage grades, despite their popularity, are the most difficult to justify or defend from a procedural, practical, or ethical perspective" (Guskey, 2015, p. 23).

O'Connor and Wormeli (2011) claim, "Recording a zero on a 100-point scale for a student's lack of work on an assessment not only falsifies the report of what he or she knows, but also immediately generates despair" (p. 41). When combined with the common practice of grade averaging, a single zero can have a devastating effect on a student's percentage grade. The atypical low score unfairly skews the overall grade (Guskey, 2015). Students readily see that receiving a single zero leaves them little chance for success or a higher grade because such an extreme score drastically skews the average (Guskey, 2015).

Many teachers see zeros as their ultimate grading weapon, using them to punish students for not making an adequate effort or failing to show appropriate responsibility (Guskey, 2015). Students get zeros for not meeting set deadlines, misbehaving in class, or refusing to heed the teacher's warnings (Guskey, 2015). A zero is seldom an accurate reflection of what a student has learned or is able to do (Guskey, 2015; O'Connor, 2017). Guskey (2004) stated, "No studies support the use of zeros or low grades as effective punishments" (p. 33).

Despite evidence that grading as punishment does not work (Wormeli, 2018) and the mathematical flaw in the use of the zero on a 100-point scale (Reeves, 2008), many teachers routinely maintain this policy in the mistaken belief that it will lead to improved student performance. Defenders of the zero claim that students need to have consequences for challenging the teacher's authority and failing to turn in work on time. Reeves (2008) added, "They're right, but the appropriate consequence is not a zero; it's *completing the work*—before, during, or after school, during study periods, at 'quiet tables' at lunch, or in other settings" (p. 86). If teachers do want improved student performance, then they must understand the difference between a harmful zero and students actually doing the work.

Giving a summative grade to homework is a common practice that distorts student learning. When teachers award points to students for meeting the basic expectations of turning assignments in on a regular basis, they often focus on meeting those rote requirements. They no longer think about learning; they have bought into a system that issues points in exchange for compliance (Iamarino, 2014). Wormeli (2006a) noted that awarding points for simply completing homework risks sending the wrong message to students: they can be successful without improving the quality of their work, if only they complete it and turn it in on time: "Homework is practice, never to be confused with absolute, final declarations of summative mastery" (p. 22).

Dueck (2014) believed that grading homework promotes busy work at the expense of intrinsic motivation and authentic learning and it could result in inflated grades as well as cheating. Homework should be a formative assessment that checks for understanding or that helps prepare students for summative assessments (Vatterott, 2015). If homework grades play into in a points system that assigns zeros for uncompleted assignments and calculates final grades through averaging, students who are capable could seriously damage their grades by failing to complete a number of homework assignments (Hanover Research, 2014).

Grading on a curve is another common practice that distorts the meaning of student grades. When teachers grade on a curve, they assign grades according to how students compare to their peers (Hanover Research, 2014). Basing students' grades on their relative standing among classmates can prompt resentment toward high-scoring students who inflate the curve and who cause other students to receive low grades (Guskey, 2015). Students must compete against one another for the few high grades awarded by their teachers. "Doing well does not mean learning excellently; it means outdoing your classmates" (Guskey, 2015, p. 51). Grading students by comparing their performance to one another distorts individual achievement (O'Connor, 2011).

The practice of giving students extra credit assignments also distorts grades away from being measures of performance and content or concept mastery (Hanover Research, 2014). Student achievement should not distort upward by the use of extra credit or bonus points. "Grades are supposed to be measures of achievement, so it is appropriate that students have "extra" opportunities to improve their grades, but these opportunities must involve demonstration of the knowledge and skills in the standards" (O'Connor, 2009, p. 104). Even if the completion of an extra assignment results in extra credit, it still skews the meaning of a student's grade because it rewards them for extra effort as opposed to achieving proficiency (Hanover Research, 2014).

Finally, the last component that inaccurately measures student performance is incorporating teacher expectations or judgment into grades. For example, the use of the "semester killer"–the single project, test, lab, paper, or other assignment that will make or break students. This practice puts 18 weeks of work at risk based on a project that might, at most, have consumed four weeks of the semester (Reeves, 2008). When a grade is supposed to report students' mastery at the end of that process, it is unethical and inaccurate to include earlier failed attempts. It is also imprecise to rely solely on single-sitting assessments for the most accurate report of what students know and can do. Instead, we look for evidence over time (O'Connor & Wormeli, 2011).

Methods

This study utilized a non-experimental, causal-comparative, ex-post facto research design. This quantitative study sought to correlate achievement on the Scholastic Math Inventory assessment as reported by a traditional or standards-based grading system implemented by the Heartland Area Schools. This study took advantage of a natural experiment in that the school district included some middle schools which implemented traditional-based grading and one middle school which implemented standards-based grading. All of the schools took the SMI assessment. The middle school, which implemented standards-based grading, was a Title I school with almost 79% of its students being minority and 98% receiving free and reduced lunch. The other four middle schools in this study are not Title I schools and these schools do not have the minority population as high as the Title I school nor do they have as many students receiving free and reduced lunch.

The use of pseudonyms helps to protect the identity of the school district in this study. The population consisted of students from Heartland Area Schools in a state in the Midwest United States. The individuals were sixth-, seventh-, and eighth-grade mathematics students from five different middle schools as they took the Scholastic Math Inventory assessment. One group of students learned in a standards-based grading system whereas the other group of students learned in a traditional grading system. There were about 500 students in the standards-based grading system and about 1,900 students in the traditional grading system.

Because this study compared end-of-year grades to Scholastic Math Inventory (SMI) scores, the researcher used ex-post facto data. Additionally, the researcher analyzed report cards to investigate relationships between traditional grading scores and SMI scores and relationships

between standards-based grading scores and SMI scores.

The administration and professional staff at Heartland Area Schools devised a grading system for evaluating and recording student progress. Four of the middle schools utilize a traditional grading system set by Heartland Area Schools in which an A, B, C, D, or F denote student understanding:

- A–outstanding work
- B–better than average work
- C–average work
- D and F–below average work

For the purpose of this study, letter grades convert to numbers as follows: A = 4, B = 3, C = 2, D = 1, and F = 0.

One middle school uses the standards-based grading system in which 4, 3, 2, or 1 denote levels of understanding based on specific proficiency levels developed by the team of mathematics teachers at that school:

- 4–Exceeds/Thorough
- 3–Proficient/Adequate
- 2–Partial
- 1–Minimal

Mathematics Assessment

Scholastic Math Inventory (SMI) assessment data and end-of-year report card grades for mathematics for students in grades 6-8 served as the data for this study. Scholastic Math Inventory is a computer-based adaptive assessment that measures students' readiness for instruction and tracks progress towards algebra readiness (Scholastic Inc., 2014). Leadership from organizations that included the National Mathematics Panel, the National Council of Teachers of Mathematics, and the Common Core State Standards Initiative contributed to the SMI. The SMI rates with the highest marks for reliability and validity by The National Center for Response to Instruction. The information from the assessment indicates the level at which students are ready to learn. This Framework provides a unified frame of reference across mathematics by organizing skills and concepts info functional, hierarchical relationships (Scholastic Inc., 2014). Students in grades 6-8 take this assessment a minimum of three times per year. The more often students take the assessment, the more they are able to demonstrate their mathematical understanding.

Results

The results of each research question included tabular results with narrative descriptions of salient findings. Pearson correlations were determined by means of SPSS, Version 22 using SMI and student end-of-course grades (traditional or standards-based). Pearson product moment correlations aided with the investigation of the relationship between SMI scores and end-of-year grades for sixth-, seventh-, and eighth-grade students in a traditional grading system (research question one). Research question 2, exploring the relationship between SMI scores and end-of-year grades sixth-, seventh-, and eighth-grade students in a standards-based grading system, also employed Pearson product moment correlations using scaled scores. The following guidelines guided the interpretation the Pearson's correlation coefficient: Weak: 0.1 to 0.3, Moderate: 0.3 to 0.5, and Strong: 0.5 to 1.0 (Cohen, Manion, & Morrison, 2013).

Relationship Between SMI Scores for Sixth-, Seventh-, and Eighth-grade Students in a Traditional Grading System from Four Middle Schools

In order to answer research question one, the researchers analyzed data from students in four middle schools who use traditional letter grades. The data consisted of the end-of-year letter grades for all mathematics students in grades 6, 7, and 8. Table 1 summarizes the results. The correlation between the Scholastic Math Inventory and traditionally assessed end-of-course mathematics grades in grades 6, 7, and 8 for all students was r (1892) = 0.355, p = 0.000, with an R^2 of 0.126 (13% of variance shared). These results indicate a moderate correlative relationship between the end-of-course grades from a traditional grading system in mathematics to the Scholastic Math Inventory.

Table 1
Relationship between SMI Scores and End-of-Year Grades for Sixth-, Seventh-, and Eighth-Grade Grade Students in a Traditional Grading System

Student Group	*n*	Pearson *r*	R^2	Significance
All Students	1892	0.355	0.126	0.000*
6th Grade Students	753	0.348	0.121	0.000*
7th Grade Students	595	0.397	0.158	0.000*
8th Grade Students	544	0.405	0.164	0.000*

*significant correlation at .05

Relationship Between SMI scores and End-of-year Grades for Sixth-, Seventh-, and Eighth-Grade Students in a Standards-based Grading System from One Middle School

In order to answer research question one, the researchers analyzed data from students in one middle school who utilized standards-based grading. The data consisted of the end-of-year standards-based grading scores for all mathematics students in grades 6, 7, and 8. Table 2 summarizes the results. The correlation between the Scholastic Math Inventory and students' mathematics standards-based scores in grades 6, 7, and 8 for all students was r (377) = 0.392, p = 0.000, with an R^2 of 0.154 (15% of variance shared). These results indicate a moderate correlative relationship between the end-of-course grades from a standards-based grading system in mathematics to the Scholastic Math Inventory. Table 2 displays the breakdown by grade level for the standards-based grading system.

When looking specifically at end-of-course grades from the 6th and 8th grade students, the results indicate a strong correlative relationship between the end-of-course grades from a sixth-grade and an eighth-grade standards-based grading system in mathematics and the SMI. These results indicate a moderate correlative relationship between the end-of-course grades from a seventh-grade standards-based grading system in mathematics to the Scholastic Math Inventory.

Table 2
Relationship between SMI Scores and End-of-Year Grades for Sixth-, Seventh-, and Eighth-Grade Students in a Standards-Based Grading System

Student Group	*n*	Pearson *r*	R^2	Significance
All Students	377	.392	.154	.000*
6th Grade Students	151	.607	.368	.000*
7th Grade Students	123	.465	.215	.000*
8th Grade Students	103	.576	.332	.000*

*significant correlation at .05

Discussion

There are positive correlations between standards-based grading and SMI results throughout the middle school grade levels. In fact, all standards-based grades in this study correlated more highly to SMI than corresponding traditional grades. This study contributes evidence to suggest that standards-based reporting provides accurate information regarding student learning as a measure for student achievement, which encourages support for a recommendation for all educators to utilize standards-based grading in school districts.

The subjectivity of a traditional grading system happens when teachers include both content and effort. In one school, a score of a "C" might mean something entirely different from teacher to teacher. Wormeli (2006a) stresses how traditional grades are not an accurate description of what a student knows and is able to do and this was the case with the traditional grading system at Heartland Area Schools. Grades are influencing areas of advanced placement as well as remedial classes. However, the inconsistency of grading in a traditional sense might make a parent wonder if their child is getting their mathematical needs met based on a subjective letter grade given by their child's teacher.

When using a standards-based grading system, teachers determine student learning based on objective goals and/or standards. Predetermined proficiency levels are set before students complete the assessment. When teachers use standards-based grading they separate academic and nonacademic components. Teachers can more accurately communicate achievement and learning to students, to parents, and to other educators. This consistency supports the goal of mathematical educators to support the needs of each individual child at their specific readiness level. When teachers use the same criteria to assign grades, students may receive a consistent message about the expectations. Once a standards-based grading system is in place it will not matter which teacher a student has because every teacher will have clear scoring criteria and expectations based on specific learning targets or goals.

Implications for Practice

The results of this study provide evidence that a standards-based grading system, as opposed to a traditional-based grading system, is more closely aligned with the results of the Scholastic Math

Inventory standardized test. This finding is important for leaders of K-12 mathematical instruction, for leaders of higher education mathematical instruction, and for policymakers of mathematical instruction because it adds to the argument that a standards-based approach to grading should replace the traditional approach to grading.

Leaders of mathematical instruction in K-12 education understand that this is an era of educational accountability, which is driving a new way of viewing how teachers in K-12 education are grading their students. When leaders can articulate that a standards-based approach is more closely aligned with the Scholastic Math Inventory test, it may build the confidence of teachers as they traverse through the difficult work of transitioning from a traditional-based approach of grading towards a standards-based approach to grading.

Leaders of mathematical instruction in higher education must continue to stay abreast of the studies that analyze the pros and the cons of traditional-based grading systems to standards-based grading systems. This is vital to the process of training the next generation of mathematics teachers as well as helping current mathematics teachers to understand how a shift towards standards-based grading aligns better with the Scholastic Math Inventory.

Future Research Recommendations

Grading has been the same for more than a century. Traditionally, teachers have combined content and learning behaviors into one letter grade. School districts need to identify what teachers are grading and reporting on report cards. Mathematical standards will show what a student knows and understands which is why professional development needs to focus on the purpose of grading and reporting. Through research, a shared understanding needs to be a priority with all teachers across school districts.

To further explore standards-based grading and traditional grading, we recommend that a study with a larger population across a wider geographic region utilizing standards-based grading, which makes the study more generalizable. This study included mathematics students in grades 6-8, which is a limiting aspect of the study. A similar study should include students in all grade levels and in other content areas besides mathematics. Finally, the data collected was from one school year, as additional years of data would add depth to future studies.

Conclusion

Policymakers of mathematical instruction must continue to develop policies steeped in evidence-based research. It is essential for policymakers to understand which type of grading approach more closely aligns with valid and reliable standardized tests. If traditional-based grading systems persist, leaders in mathematical instruction will continue to ineffectively communicate about student achievement, which is misrepresenting their learning. In conclusion, this study provides support for a recommendation for leaders of mathematics education at all levels to utilize standards-based grading in school districts.

References

Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning. *Clearing House*, *78*(5), 218-223.

Boaler, J., (2015). *What's math got to do with it?* Penguin Books: New York.

Bourgeois, S. J., & John Eric Boberg, J. E. (2016). High-Achieving, Cognitively Disengaged Middle Level Mathematics Students: A Self-Determination Theory Perspective, *Research in Middle Level Education Online*, 39:9, 1-18, DOI: 10.1080/19404476.2016.1236230

Brookhart, S. M. (2011a). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues & Practice*, *30*(1), 3-12.

Brookhart, S. M. (2011b). Starting the conversation about grading. *Educational Leadership*, *69*(3), 10-14.

Brookhart, S. M. (2013). The public understanding of assessment in educational reform in the United States. *Oxford Review of Education*, *39*(1), 52-71.

Brookhart, S. M. (2017). *How to Use Grading to Improve Learning*. ASCD.

Cohen, L., Manion L., & Morrison K. (2013). *Research methods in education.* London: Routledge.

Dilendik, J. R. (1978). Assumptions underlying criterion referenced assessments are educationally sound. *Education*, *99*(1), 89-96.

Dueck, M. (2014). *Grading smarter not harder: assessment strategies that motivate kids and help them learn.* Alexandria, VA: Association for Supervision and Curriculum Development.

DuFour, R., DuFour, R., Eaker, R. & Many, T. (2010). *Learning by doing: A handbook for professional learning communities at work.* Bloomington, IN: Solution Tree Press.

Guskey, T. R. (2004). Are zeros your ultimate weapon? *Education Digest*, *70*(3), 31-35.

Guskey, T. R., & Jung, L. A. (2009). Grading and reporting in a standards-based environment: Implications for students with special needs. *Theory into Practice*, *48*(1), 53-62.

Guskey, T. R., & Bailey, J. M. (2010). *Developing standards-based report cards*. Thousand Oaks, CA: Corwin Press.

Guskey, T. R., Swan, G. M., & Jung, L. A. (2011). Grades that mean something. *Kappan Magazine, 93*(2), 52-57.

Guskey, T. R. (2015). *On your mark: Challenging the conventions of grading and reporting*. Bloomington, IN: Solution Tree Press.

Hanover Research. (2011). *Effective grading practices in the middle and high school environments*. Washington, DC: Hanover Research.

Hanover Research. (2014). The impact of formative assessment and learning intentions on student achievement. Washington, DC: Hanover Research.

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

Hofman, A. D., Visser, I., Jansen, B. R., Marsman, M., & van der Maas, H. L. (2018). Fast and slow strategies in multiplication. *Learning and Individual Differences*, *68*, 30-40.

Iamarino, D. L. (2014). The benefits of standards-based grading: A critical evaluation of modern grading practices. *Current Issues in Education*, *17*(2), 1-10.

Iverson, M. (2014). *Relationships of the Iowa assessment results to traditional and standards-based English and math end-of-course grades*. University of South Dakota, Vermillion, SD.

Marzano, R. J. (2000). *Transforming classroom grading*. Alexandria, VA: Association for

Supervision and Curriculum Development.
Marzano, R. J. (2006). *Classroom assessment & grading that works*. Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).
Marzano, R. J., Pickering, D. J., Arredondo, G. J., Blackburn, R. S., Brandt, C. A., Moffett, D. E., Paynter, D. E., Pollock, J. E., & Whisler, J. S. (1997). *Dimensions of learning: Trainer's manual*. Alexandria, VA: Association for Supervision and Curriculum Development.
McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *Journal of Educational Research*, *95*(4), 203.
National Commission of Excellence in Education (1983). A Nation at risk: The imperative for educational reform. Washington, US Government Printing Office. Retrieved http://www.2ed.gov/pubs/NatAtRisk/index.html
National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all.* Reston, VA: NCTM.
National Governors Association. (2010). Center for Best Practices, Council of Chief State School Officers (CCSSO). *Common Core State Standards.* Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers.
O'Connor, K. (2009). *How to grade for learning: K-12*. Thousand Oaks, CA: Corwin.
O'Connor, K. (2011). *A repair kit for grading: 15 fixes for broken grades*. Boston, MA: Pearson Education.
O'Connor, K. (2017). *How to grade for learning: Linking grades to standards*. Corwin Press.
O'Connor, K., & Wormeli, R. (2011). Reporting student learning. *Educational Leadership*, *69*(3), 40-44.
Reeves, D. (2008). Leading to change: Effective grading practices. *Association for Supervision and Curriculum Development*, *65*(5), 85-87.
Reeves, D. (2011). *Elements of grading: A guide to effective practice*. Bloomington, IN: Solution Tree Press.
Schaefer, M. B., Malu, K. F., & Yoon, B. (2016). An Historical Overview of the Middle School Movement, 1963–2015, RMLE Online, 39:5, 1-27, DOI: 10.1080/19404476.2016.1165036
Scholastic Math Inventory. (2011). *Scholastic Software Manual*. Retrieved from http://edproductsupport.scholastic.com/content/techsupport/smi/manuals/SMI_SM_2_0_1.pdf
Shepard, L. A. (2009). Commentary: Evaluating the validity of formative and interim assessment. *Educational Measurement: Issues & Practice*, *28*(3), 32-37.
Sieling, C. J. (2013). *Standards-based grading in mathematics: Effects on student achievement and attitude*. Southwest Minnesota State University, Marshall, MN.
Simon, S. B. & Bellanca, J. B. (1976). *Degrading the grading myths: A premier of alternatives to grades and marks.* Washington, DC: Association for Supervision and Curriculum Development.
Stiggins, R. (2014). *Revolutionize assessment*. Thousand Oaks, CA: Corwin Press.
Tyack, D., & Tobin, W. (1994). The "grammar" of schooling: Why has it been so hard to change? *American Educational Research Journal*, *31*(3), 453-479.
Vatterott, C. (2015). *Rethinking grading: Meaningful assessment for standards-based learning*. Alexandria, VA: Association for Supervision and Curriculum Development.
Welsh, M. E., D'Agostino, J. V., & Kaniskan, B. (2013). Grading as a reform effort: Do standards-

based grades converge with test scores? *Educational Measurement: Issues & Practice*, *32*(2), 26-36.
Wiggins, G. (2012). Seven keys to effective feedback. *Educational Leadership, 70*(1), 10–16.
Wormeli, R. (2006a). Accountability: Teaching through assessment and feedback, not grading. *American Secondary Education*, *34*(3), 14-27.
Wormeli, R. (2006b). Fair isn't always equal. *Assessing & grading in the differentiated classroom.* Portland, ME: Stenhouse.
Wormeli, R. (2011). Redos and retakes done right. *Educational Leadership*, *69*(3), 22-26.
Wormeli, R. (2018). *Fair isn't always equal: Assessing & grading in the differentiated classroom*. Stenhouse Publishers.