



University of Nebraska at Omaha
DigitalCommons@UNO

Student Work

11-1-2017

Elementary Teachers' Grading Practices and Perceptions in Communicating Assessment of Students

Shannon A. Thoendel
University of Nebraska at Omaha

Follow this and additional works at: <https://digitalcommons.unomaha.edu/studentwork>

 Part of the [Education Commons](#)

Recommended Citation

Thoendel, Shannon A., "Elementary Teachers' Grading Practices and Perceptions in Communicating Assessment of Students" (2017). *Student Work*. 3646.
<https://digitalcommons.unomaha.edu/studentwork/3646>

This Dissertation is brought to you for free and open access by DigitalCommons@UNO. It has been accepted for inclusion in Student Work by an authorized administrator of DigitalCommons@UNO. For more information, please contact unodigitalcommons@unomaha.edu.



ELEMENTARY TEACHERS' GRADING PRACTICES AND PERCEPTIONS IN COMMUNICATING
ASSESSMENT OF STUDENTS

By

Shannon A. Thoendel

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Education

Major: Educational Administration

Under the Supervision of Dr. Kay A. Kaiser

Omaha, Nebraska

November, 2017

Supervisory Committee:

Kay A Keiser, Ed.D.

C. Elliot Ostler, Ed.D.

Tami D. Williams, Ed.D.

ProQuest Number:10680687

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10680687

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Abstract

ELEMENTARY TEACHERS' GRADING PRACTICES AND PERCEPTIONS IN COMMUNICATING ASSESSMENT OF STUDENTS

Shannon A. Thoendel, Ed.D.

University of Nebraska, 2017

Advisor: Kay A. Keiser, Ed.D.

Grading reform has been a major focus in school districts across the country (Guskey & Jung, 2012). Reporting student achievement through grades can have a lasting and profound impact on a student's academic career. Grades are often considered to have little relationship to student performance (Brookhart, 2004; O'Connor 2016). Grading authenticity is reached when the following standards are met: accuracy, consistency, meaningfulness, and supportiveness of the learning environment (O'Connor, 2016). A current movement to standards-based grading is believed to be a more accurate measure of student achievement than a traditional grading system (Guskey, 2001). Standards-based grading systems are believed to align with student performance evidence gained through standardized testing (Coladarci, 1986). The purpose of this exploratory mixed method research study was to identify how teacher perception of assessment practices through a traditional grading system impacted the correlation of student performance on a standardized assessment. The research was conducted in fourth grade classrooms districtwide in small suburban Nebraska school district. Quantitative achievement data was analyzed using a Pearson Product Momentum Correlation in order to identify the strength of the relationship between teacher

assigned third quarter report card grades and student achievement on the Nebraska State Assessment. The researcher then conducted a qualitative research approach to investigate how the fourth grade teacher grading practice perceptions impacted the assignment of grades. The analysis of compiled data lead to consequential implications for alignment in theory and grading practice.

Acknowledgements

It is with my sincerest appreciation that I recognize those that have supported and encouraged me in order to reach my goal of completing this dissertation. I first want to thank Dr. Kay Keiser for serving as my dissertation chair. Her ability to direct and challenge my thinking, expanded my knowledge and allowed my dissertation to form into a study that I am proud to present. Dr. Keiser was there to encourage, support, and build my confidence throughout the entire process.

I wish to thank the members of my committee: Dr. Elliot Ostler and Dr. Tami Williams. During my dissertation proposal their expertise assisted me in developing research methods that maximized the study results. Included in my journey are Dr. Dick Christie, who first supported my research in the area of grading during graduate studies coursework and Dr. Janice Garnett, who believed in me as a beginning administrator.

Thank you to Dr. Terry Haack for being a mentor, a strong influence on my professional development, and giving me feedback for growth. He is always there to listen and reassure. Thanks to Dr. Denise Hoge for being willing to share her resources and knowledge. She is never too busy or too tired to support growth of fellow educators. She is a role model and one of my biggest cheerleaders. I could not have made it through this advance degree without the encouragement and expertise of the Bennington administrative team, Dr. Matt Blomenkamp, Mr. Greg Lamberty, Mr. Shawn Hoppes, Mr. Michael Spoehr, Mr. Chad Boyes, and Mrs. Therese Nelson. The staff of Pine Creek Elementary is always looking for ways to continuously improve the learning

climate and meet the needs of diverse learners, which gave me the drive to grow and expand my skills as a school administrator.

My family has played an important role in my dissertation. The goal of dissertation attainment began when my dad shared that it was my grandfather Lloyd Dowden's dream for his grandchildren to reach the doctoral level. My parents, Larry and Cindy Dowden, have always emphasized the importance of education and a strong work ethic. Their love, belief in me, and instilment of strong foundations have made this achievement possible.

Finally, I want to express my love and gratitude to my husband, Martin Thoendel. He knew just what to say to anger me enough to put me back on track when I was feeling depleted. He never complained about the financial, time, and support commitment he had to give while I pursued this endeavor. My three children, Benjamin, Brooke, and Natalie, understood when I missed their special events or when dinner was leftovers again. I hope seeing mom go through this process will push them to follow their dreams and never give up when the task becomes difficult. I am blessed to have such a strong support system that was essential in helping me complete this dissertation and attain my doctoral degree.

Table of Contents

	Page
Abstract	ii
Acknowledgments	iv
Table of Contents	vi
List of Tables	ix
List of Appendices	x
Chapter One Introduction	1
Conceptual Framework	5
Teacher Cognition Theory	5
Thin-Slicing Theory	7
Statement of Problem	8
Purpose of Study	9
Research Questions	9
Definition of Terms	10
Assumptions	11
Limitations	12
Delimitations	12
Significance of the Study	13
Organization of Dissertation	14
Chapter Two Literature Review	16
History of Grading	16

No Child Left Behind Act and Every Student Succeeds Act	17
Nebraska State Accountability.....	19
Purpose of Grading	20
The Effects of Grading On Students	22
Common Grading Design	23
Achievement Factors versus Non-Achievement	23
Shift to Standards-Based Grading	24
Grading Reliability Factors	26
Standardized Testing	28
Correlation of Teacher Performance Judgements and High-Stakes Testing	30
Teacher Experience and Grading Practices	32
Practices to Improve Grading	33
Chapter Three Methodology	35
Research Design	37
Research Questions	38
Setting	39
Study Participants	40
Study Instruments	41
Nebraska State Accountability Assessment	41
Teacher Assigned Report Grades	45
Procedures	46
Data Analysis	47

Summary	48
Chapter Four Results for the Quantitative and Qualitative Findings	50
Overview of Results	50
Research Question 1	51
Research Question 2	52
Research Question 3	53
Research Question 4	60
Summary	75
Chapter Five Conclusion and Discussion	76
Purpose of the Study	76
Conclusion	77
Study Results	77
Research Question 1	77
Research Question 2	79
Research Question 3	80
Research Question 4	81
Discussion	85
Alignment with Conceptual Framework	87
Implications for Student Achievement Communication	91
References	95
Appendix A	103

List of Tables

Table 1.	Correlation of Student Performance on Student Report Card Percentage Grades and the Nebraska State Accountability Assessment Using a Pearson Product Moment	55
Table 2.	Scatterplot of Student Performance on Student Report Card Percentage Grades and the Nebraska State Accountability Assessment for Reading	56
Table 3.	Scatterplot of Student Performance on Student Report Card Percentage Grades and the Nebraska State Accountability Assessment for Mathematics	57
Table 4.	Scatterplot of Student Performance on Student Report Card Percentage Grades and the Nebraska State Accountability Assessment for Language Arts/Writing	58
Table 5.	Correlation by Teacher of Student Performance on Student Report Card Percentage Grades and the Nebraska State Accountability Assessment Using a Pearson Product Moment	59
Table 6.	Teacher Perception of the Assessment and Grading Practices	63
Table 7.	Evidence Most Often Used to Communicate Student Achievement	66
Table 8.	Difference in Assessment Practice by Content Areas	68
Table 9.	Received Feedback on Grading Practices and Communication of Student Achievement	70
Table 10.	Barriers for Consistent and Clear Communication of Student Achievement	72
Table 11.	Suggested Support for Consistent Grading Practices and Clear Communication of Student Achievement	74

List of Appendices

Appendix A	Teacher Participants Notification Letter and Teacher Questionnaire Elementary Grading and Assessment Practices	103
------------	---	-----

CHAPTER ONE

INTRODUCTION

Since the 1800's, schools across the United States have assigned grades to communicate a student's level of academic performance (Kirschenbaum, Simons, & Napier, 1971). School communities and postsecondary institutions have come to expect grade assignment as a strong academic performance indicator. Grades are used to communicate achievement to students and parents. They also inform postsecondary institutions and prospective employers of a student's potential, assist administration in making decisions such as class placement, promotion, and honors, and evaluate the effectiveness of instructional programs (Frisbie & Waltman, 1992; Guskey, 2002; Linn, 1983; O'Connor, 2016).

Ken O'Connor (2016) refers to grades as a summary statement of student performance, but he questions whether one number or letter, at the end of a set time period, gives a clear description of student performance levels. He further states grading should not just be a numbers game. Grades contain professional judgements, which are decisions made by educators that are derived from experiences, shared public standards, and established policies or guidelines. Fairness and uniformity in grading is obtained when all students are provided equitable learning opportunities (O'Connor, 2016).

A variety of purposes for grades are considered authentic, but educators rarely agree upon the main purpose (Guskey & Jung, 2012). According to O'Connor (2016) the main point in grading is to create a culture of continuous learning and not one of

grading or testing. This is achieved when a grading system meets the following standards: accuracy, consistency, meaningfulness, and supportiveness of learning environment. When these grading standards are achieved, the system communicates student performance accurately to all stakeholders (O'Connor, 2016).

Grades have a lasting and profound effect on the academic career of a student. They can influence a student's entire academic career. Grades can shape students' attitudes, behaviors, and motivations to learn. Even though grading systems play an important role in education, they are often considered to have little relationship to student performance (Brookhart, 2004; O'Connor, 2016). Grades are thought to measure more the willingness of a student to cooperate and work hard, rather than reporting learning. Therefore they may distort feedback regarding the student's ability. Finally, grades are influenced by nonacademic factors which distort the measurement of learning (Guskey & Jung, 2012; Winger, 2005). When nonacademic factors are included in grading, it blurs the meaning of grades by inflation or deflation of academic achievement (O'Connor, 2016).

According to Guskey (2001), a more accurate measure of student achievement is attained when measuring student achievement through learning standards, rather than comparing students to each other. Essentially, this means moving from a traditional norm-referenced grading system to a standards-based criterion-referenced grading system. The discourse into the effectiveness of traditional grading systems remains controversial. The movement to redefine grading seeks to improve teaching and learning. This paradigm shift has invigorated school systems across the nation to

reevaluation their student performance communication tools. School systems want to insure that current student achievement practices provide a true reflection of student learning and correlates with authentic student performance.

In addition to traditional grading systems, standardized assessments are commonly used in the measurement of student performance. Standardized assessments are given under uniform conditions, therefore they are considered, by some, to be a meaningful basis for evaluating performance (Coladarci, 1986). Not all researchers feel that standardized assessments are the solution to accurate student performance measurement. Brookhart (2004) agrees that grades rarely express student achievement levels, but feels the use of standardized assessments to communicate student performance levels is even more severe.

Popham (1999) describes three aspects of standardized achievement assessment that make them the wrong measurement tool to portray an accurate interpretation of student performance. The first problem is confounded causation. A student's score on a standardized achievement test is impacted by three factors: what's taught in school, a student's native intellectual ability, and a student's out-of-school learning. The quest for score variance creates the second problem with standardized achievement testing. Items on which students perform well are often excluded from the test. Those items often cover the content that is most stressed in schools because of its importance. Thirdly, there is almost certain to be a significant mismatch between what is taught and what is tested. Between school districts across the country there is considerable curricular diversity. Standardized achievement assessment developers create a one-

size-fits-all test in order to measure all content area knowledge and skills that are regarded as important. This is an impossible task. This one-size-fits-all assessment will always contain many items that are not aligned with what is emphasized instructionally in a particular educational setting (Popham, 1999).

The ability to create an accurate picture of a student's ability and communicate student performance for greater academic success has led to the rapid movement of grading reform (Guskey & Jung, 2012). Mixed messages are sent when student classroom performance does not match student performance on standardized tests. Educators need to gain a clearer understanding of performance evidence that is accurate, consistent, meaningful, and supportive of learning. Grading systems that align with these standards yield the greatest impact on teaching and learning (O'Connor, 2016).

Authors of assessment and grading practices question if either measure, report card grades or standardized tests, truly represent student achievement. It should not matter what teacher students are assigned to, what content area they are studying, what standardized assessment they take, or what score they are assigned. Students deserve to receive clear communication that accurately describes their level of performance in order to accelerate their learning. While the purpose for grading and assessment is valid, the application can be inconsistent. This inconsistency can cause discrepancy in communication of student performance. More exploration is needed to determine if student performance outcomes on a traditional grading scale and student

performance on a standardized assessment correlates, to offer a true reflection of student learning and provides authentic student performance statistics.

Teacher Cognition Theory

At the heart of the research done to investigate the reliability of student achievement on a standardized test, as comparable to the demonstration of classroom performance through grading, is the Theory of Teacher Cognition. Teacher cognition is defined as a teacher's mental structure. It is what teachers think, know, believe, or assume in relation to the aspects of the teaching profession. The framework for teacher cognition is constructed from three main themes: a teacher's accumulated experiences as a learner, the impact of teacher education through professional coursework, and the ongoing process of classroom practices and contextual factors (Borg, 2009).

The study of teacher cognition spans back to the 1960's where research on teaching focused on finding universal behaviors that could be applied by all teachers resulting in greater learning, which was typically measured through achievement tests. This kind of research was referred to as a process-product model of research. The teacher cognition model identifies there is a complex relationship between teacher knowledge, belief, and their instructional practice. Borg (2003) describes cognition as the unobservable mental dimension of teaching. He explains that teachers are active decisions-makers who make classroom decisions by relying on personalized and contextual pathways of knowledge, thoughts, and beliefs.

Borg (2003) found that teachers have cognitions about all aspects of their work.

He constructed a representation of the characteristics of teacher cognition (see Figure 1).

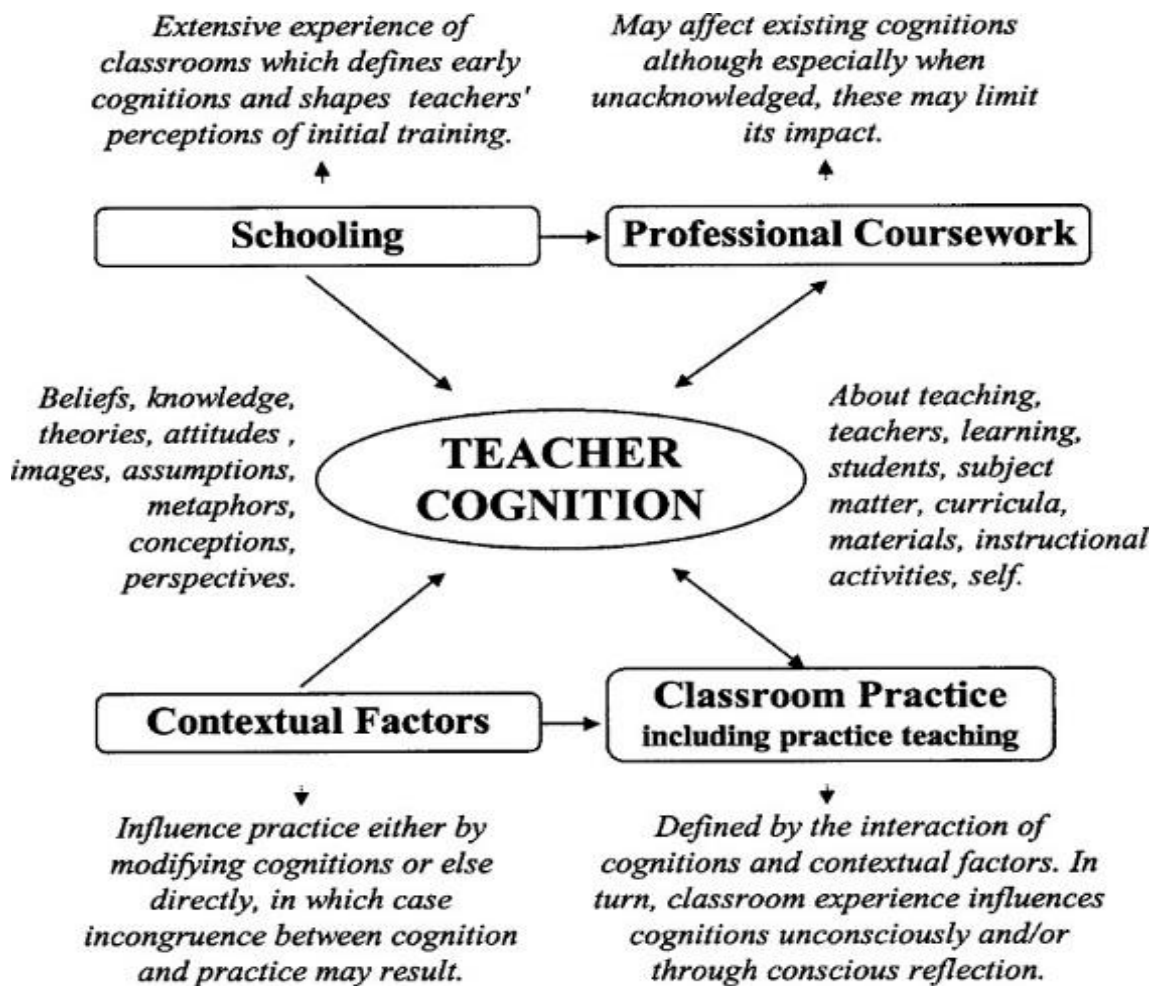


Figure 1. Teacher cognition, schooling, professional education, and classroom practice.

Teacher cognition is pivotal in determining the basis for teaching practices. It plays an important role in defining the extent to which teachers are able to implement compatible teaching practices. Therefore the Teacher Cognition Theory formulates the debate over what impact teacher cognition may have on classroom grading practices

and the consistency which grading practices predict student performance on standardized tests.

Congruency of student classroom performance judgements and standardized testing performance are often questioned. Researchers such as Borko, Cone, Russo, and Shavelson (1997) and Clark and Peterson (1986) suggest that teacher-based student reading comprehension level judgments are impacted by teacher cognition. These student achievement judgments are viewed as relevant assessments and are used for providing feedback to children, parents and school personnel (Hoge and Coladarci, 1989). These judgments are also used to make instructional decisions. Clark and Peterson (1986) further state that inaccurate student performance judgments, influenced by teacher cognition, lead to different decision making processes. Since grades, which are derived from classroom performance judgments, are perceived as a primary indication of student performance, it is important to understand the similarity of classroom performance judgments and standardized test performance.

Thin-Slicing Theory

Thin-slicing is a term that is used when an individual makes a quick decision on the basis of limited information (Gladwell, 2005). It takes the mind two seconds to form a judgment and arrive at a conclusion. Gladwell (2005) refers to this as a “blink.” Thin-slicing is usually applied when a person has to deal with an abundant amount of information that has to be understood in a complex situation in a short period of time. When thin-slicing, an individual will rely on their education, experience, and beliefs to gain quick understanding (Gladwell, 2005).

Thin-slicing occurs in education when student performance judgments are derived from a limited amount of evidence in a short period of time. Trying to combine copious amounts of performance standards into one standardized, criterion-referenced test may not yield the most accurate representation of student achievement. The outcomes of thin-slicing may have tremendous implications and significance for educators, especially when that thin-slice of information is used to determine and communicate student academic performance levels.

The Theory of Teacher Cognition has provided valuable insight into the mental structures of teachers, yet there does not seem to be a clear sense of unity in the area of predicting and communicating student achievement. Like teacher cognition, in the Theory of Thin-Slicing, a teacher will generally use their education, experience, and beliefs to gain a quick understanding from limited information. The perspective of teacher cognition and thin-slicing on the comparability of student achievement in the classroom and student performance on standardized tests deserves further exploration. This comparative predictive study will explore the Theory of Teacher Cognition and Theory of Thin-Slicing through fourth grade student achievement grades in the content areas of reading, math, and language arts with performance on the Nebraska State Accountability Assessment.

Statement of Problem

The expectation of success for every student has increased the necessity for true measures of academic achievement that are accurately communicated to parents, teachers, and administrators. The most relied upon source of student performance

communication is the assignment of report card grades by teachers and student performance on standardized assessments. This generates the question: Do traditional grading practices and standardized testing accurately measure and communicate student academic performance?

Purpose of the Study

This exploratory mixed method research study sought to identify how teacher perceptions of assessment practices through a traditional grading system compares to student performance on a criterion-referenced, standardized test. Further investigation into specific content areas was used to identify if there is a relationship in performance within this variable.

Research Questions

The study explored the relationship between teacher assigned report card percentage grades and student performance on the criterion-reference Nebraska State Accountability Assessment. The variance of relationship between the content areas of reading, math, and language arts for students enrolled in fourth grade was measured to determine if specific content areas influence the degree of association for the two performance measures. The study sought to answer the following research questions:

Research Question 1. How strong is the relationship between teacher assigned student report card grades in reading and scores on the Nebraska State Accountability Reading Assessment for students enrolled in fourth grade?

Research Question 2. How strong is the relationship between teacher assigned student report card grades in math and scores on the Nebraska State Accountability Math Assessment for students enrolled in fourth grade?

Research Question 3. How strong is the relationship between teacher assigned student report card grades in language arts and scores on the Nebraska State Accountability Writing Assessment for students enrolled in fourth grade?

Research Question 4. How do fourth grade teachers perceive grading and assessment practices to clearly and consistently communicate student achievement?

Definition of Terms

The researcher chose to define the following terms to clarify them for this study. To ensure clear understanding throughout the study, the following terms and acronyms have been defined:

Adequate Yearly Progress. Adequate Yearly Progress (AYP) is a term introduced with the *No Child Left Behind* legislation and indicates whether a school or system has made sufficient progress towards meeting school or system improvement goals.

Nebraska State Accountability (NeSA). Nebraska Legislative Bill 1157 passed in 2008 required a single statewide assessment of the Nebraska Academic Content Standards in the subject area of reading, mathematics, science, and writing for all K-12 Public schools. The assessment system was named Nebraska State Accountability (NeSA), with NeSA-R for reading, NeSA-M for mathematics, NeSA-S for science, and NeSA-W for writing. The assessments in reading and mathematics are administered in grades 3 through 8 and 11; science in grades 5, 8, and 11; writing in grades 4, 8, and 11

(For the purpose of this study, assessment results of fourth graders on the NeSA-R, NeSA-M, and NeSA-W were examined).

Content Standards. Content standards are broad statements that describe specific content areas that students should learn at each grade level. These standards define the knowledge within each discipline or subject area.

Standards-Based Grading System. Standards-based grading is a method of grading in which students are assessed based on their mastery of a specific skill or standard.

Traditional Grading System. Traditional grading is a system in which a letter grade is assigned based on a combination of either related or unrelated assessments of skills. For this study, a seven point grading scale will be used where every seven points below 100, a letter grade is assigned until a student reaches 70, at which point anything below 70 is a failing grade.

Criterion-referenced. Tests and assessments that are designed to measure student performance against a fixed set of predetermined criteria or learning standards.

Norm-referenced. Test, assessment, or evaluation which yields an estimate of the position of the tested individual in a predefined population, with respect to the trait being measured.

Assumptions

The research made the assumption that each teacher was highly qualified, had been trained in the grading process, and followed district policy for assigning student grades. The assumption were also made that teacher responses on a perception

questionnaire were honest and represented their true feelings. The adopted district curriculum was assumed to align with the Nebraska State Standards and were expected to be implemented system wide with consistency. The final assumption was that the NeSA Reading, Math, and Writing Assessments were an accurate measure of student content knowledge on the Nebraska State Standards.

Limitations

This research was a quantitative study dependent on public school teachers from the same school district teaching elementary school in the fourth grade. This study was specifically limited to a grading system that was adopted by one school district. In all grading, there is an element of human subjectivity. Teacher assigned report card grade data was collected. The range of teacher training in the area of assessment or grading and individual teacher beliefs on assessment or grading was not studied. Test preparation, exposure to test format, and the amount of practice on the standards measured by the NeSA assessments may vary between study participants and was not controlled. The only variables that was studied was the teachers' perception of student achievement communication, student report card achievement, and student achievement data on the NeSA-R, NeSA-M, and NeSA-W.

Delimitations

The primary goal of this mixed method exploratory study was to investigate teacher assigned report card grades using a traditional grading system and the degree to which it correlates to student performance on standardized reading, mathematics, and writing assessments. Caution was applied when making generalizations based on the

findings of this study as delimitations apply. The researcher imposed the following delimitations: The study involved only one Midwest suburban public school district that is limited to a total of eight fourth grade teachers. The research was conducted with one year's third quarter data examined. Given this, the study could have been strengthened if it included teachers from different school systems and multiple grade levels, and varying grading systems.

Significance of the Study

This study has significance as education institutions look to reform grading practices. A shift in grading from traditional practices to standards-based grading continues to be a focus nationally and locally. The results of this study can guide school districts in selecting best practices in reporting student achievement and assist in making informed decisions when recommending grading reform.

The purpose of grading is to collect evidence of student performance that creates equity in educational opportunity and leads to informed decisions in selection of curriculum and lesson planning for teachers. Gaining knowledge in the relationship of student report card grades and performance on a standardized test may enhance a school system's ability to accurately communicate student achievement, the primary purpose of assessment and grading. Furthermore, it will lead to important conversations, to determine if the right evidence is being gathered, if the right testing instruments are being used, and if communication tools accurately reflect student learning.

Professional knowledge of the relationship between grading and performance on standardized assessments will promote a clear, common description of student proficiency levels. No matter teacher cognition, years of teaching experience, or subject matter taught, grading practices should guarantee an accurate, consistent, meaningful statement that focuses on student learning. The above mentioned attributes are considered essential components of grading. Grading should be a summary statement of student performance (O'Connor, 2016). When a deeper understanding of student performance data is attained, the value of the data also increases. This value may lead to greater student motivation, better support for individualized learning, and clearer communication of student achievement between education institutions and families.

The intention of this research study was to provide school districts an examination of the accuracy of traditional grading systems to communicate and align with student performance on standardized assessments. The researcher strived to provide outcomes that assist in making informed decisions when implementing or reorganizing grading practices and philosophies. Furthermore, the researcher investigated if traditional report card grades have a stronger relationship with student performance on the Nebraska State Accountability Assessment based on specific content areas or a teachers' perceptions in communicating assessments of learning.

Organization of Dissertation

School districts across the country are investigating their current grading systems in a campaign to improve communication of student achievement and ensure mastery of content standards at every grade level. Many school districts have converted to a

standards-based system. The study of grade reporting is needed to determine whether a traditional grading system is a comparative measure of content standards achievement as reported by the Nebraska State Accountability Assessment.

Chapter one presented the background for this study, stated the problem, presented a brief overview of the research approach, and addressed the assumptions, limitations, and delimitations within the study. The first chapter concluded with a description of the significance this study may have for the research school district and for other districts looking at grading reform.

Related grading research and identification of grading practice characteristics are presented in Chapter Two. A description of factors that influence grading will assist in outlining the variation that may occur between student classroom performance and achievement on standardized assessments. Chapter Two will also relate aspects of grading that are considered practices that address the manner in which teachers assign grades.

Chapter Three presents an explanation of the research design, description of the school district, methodology for data collection, the means in which the data will be analyzed, and the instrumentation that will be used in this study. The results of the study will be presented in Chapter Four. A detailed statistical analysis of the data and an interpretation of the findings related to the research questions will be offered. In Chapter Five, a summary of the research and implications will be discussed.

CHAPTER TWO

LITERATURE REVIEW

Report card assignment of grades is one the most difficult tasks for teachers to complete, yet it is extremely important. Student grades are used for multiple purposes such as honors, retention, sorting of students, and evaluation (Mehring, Parks, Walter, & Banikowski, 1991). They are used for communication of strengths and weaknesses when indicating student progress to parents or as a motivation tool to increase student effort (Guskey, 1994). School districts set policies and procedures, but actual grading is left to the individual teacher's values and judgments on student achievement and behavior (Mehring, et al., 1991).

History of Grading

The history of grading suggests that before 1850, most schools grouped students of all ages and backgrounds into one cluster in a one-room schoolhouse. Grades were reported to parents orally during a home visit. Grading and reporting were unheard of in United States schools at this time (Guskey, 2013). In the late eighteen hundreds schools began to use formal progress evaluations. These were primarily narrative reports where teachers described the skills the student had mastered and where additional work was needed. The main objective of these reports was to communicate mastery of current level and readiness to move to the next (Edwards & Richey, 1947).

In the nineteenth and twentieth centuries, enrollment increased due to compulsory attendance laws. The number of United States schools increased dramatically. Subject specific content increased along with the number of high schools

(Gutek, 1986). Elementary schools continued to use narrative reporting for student performance, where high school teachers began to use percentages and other similar markings to communicate achievement (Kirschenbaum et al., 1971).

The shift to percentages was gradual and seemed natural due to the increased demands on high school education. In 1913, Daniel Starch and Edward Charles Elliott challenged the reliability and accuracy of percentages. The research found wide differences in assigned percentages to identical English papers. The teachers placed value on different elements of the paper they determined to be significant. Thirty different percentage grades were assigned to a single paper and scores had a range of more than 40 points (Guskey, 2013). The study was repeated with geometry papers and the researchers found even greater variation in grading. The math papers ranged in grades from 28% to 95% (Guskey, 2013). History has shown that the reliability in teacher grading practices has been questioned for over one hundred years.

No Child Left Behind Act and Every Student Succeeds Act

No Child Left Behind Act (NCLB) was signed into law by President Bush on January 8, 2002. With the implementation of this law came more stringent local, state, and federal accountability measures for school systems. Under NCLB, states were required to develop content and academic achievement standards. To measure how well all students in the state are acquiring the skills defined by the content standards, state accountability tests were also a requirement.

Annually, students in grades three through eight were tested in the areas of reading, mathematics, and science on the academic achievement standards, which was

a requirement of NCLB. Under NCLB student performance on the state mandated assessments were ranked into two high levels and a third lower level. All students were expected to reach a level of proficiency. NCLB required the total student population and specified subgroups to meet “adequate yearly progress” (AYP) by reaching the proficient level. This is defined by attainment of the two higher level rankings on the mandated state assessments. If schools and districts failed to meet AYP for two or more years, they were then classified as schools “in need of improvement” and faced such consequences as school transfer options, supplemental services, replacement of staff or administration, or a plan of restructuring (Great Schools, 2015).

In 2015 the reauthorization process of NCLB led to the Every Student Succeeds Act (ESSA), which focuses on three priorities: induction of student and school supports in state accountability, reduction in the amount of standardized testing in schools with decoupling high-stakes decision making and statewide standardized tests, and advocating for educators to be part of decision making at the federal, state, and local levels (Walker, 2015).

ESSA empowers educators as trusted professionals to make school and classroom decisions while keeping the focus on students most in need. Educators and schools whom were judged and punished by NCLB for poor achievement on high-stakes, mandated assessments, now can turn their focus to student learning and teaching. ESSA begins to close the opportunity gaps for students by providing a new accountability system that includes an “opportunity dashboard” with indicators of school success and student support. ESSA directives hold teachers and schools more

accountable for student academic results, by empowering them to use informed decision making on student performance to predict student and school success (National Education Association, 2015).

Nebraska State Accountability

The Nebraska State Accountability Assessment (NeSA) was designed in accordance with the federal enactment of the No Child Left Behind Act 2002 by the Nebraska Department of Education. Beginning in 2010, a centralized Nebraska State testing process, NeSA was developed. The first state reading results were available in 2010. Mathematics results were available in 2011. NeSA testing is required for all public schools in the State of Nebraska. The main goal of the assessment is to measure the attainment of state-wide academic standards. These academic standards provide the framework for teaching and learning in specific content areas. Test results are used to provide data to assist educators, policymakers, students, and families in understanding the achievement of every child in the State of Nebraska. Because the test is aligned to the standards of learning, a student's assessment results are perceived to communicate areas of particular proficiency and areas where additional learning might be necessary.

The NeSA Reading and Mathematics Tests asks questions in a multiple-choice format. The assessments are administered online unless a student receives an accommodation requiring a paper-and-pencil format. A student's Individualized Education Plan will describe needed accommodations which require the paper-and-pencil version.

The NeSA Reading and Mathematics Tests are given annually during a six-week testing window that begins in late March and runs through early May. The NeSA Writing Test is given between late January and early February. Prior to testing dates, teachers may access Check 4 Learning; a state developed database where teachers can obtain practice test resources that will familiarize students with NeSA questioning formats and content.

In the content area of reading, questions on the assessment cover vocabulary skills and reading comprehension ability. The mathematics test covers number sense, geometric concepts, measurement, algebraic concepts, data analysis, and probability. The NeSA Writing is an essay that requires students to demonstrate an ability to respond to a prompt in a specific writing genre in an organized format. Rubric scoring is applied to the writing response for ideas and content, organization, word choice and voice, sentence fluency and conventions.

Students receive a scaled score in individual content areas and a proficiency level indicator. The proficiency levels a student may reach on the NeSA Tests are Exceeds the Standards, Meets the Standards, or Below the Standards. Students who fall into the Below the Standards category are thought to be in need of remedial instruction for specific content areas that fall below the standard (Nebraska Department of Education, 2016).

Purpose of Grading

Grades are significant in American education systems. They are used to determine class placement, retention practices, college admissions, and scholarships.

Today's report cards are used to sustain state funding, generate positive feelings between school and community, assist teachers in increasing students' self-esteem, used to reward students' behavior, and create a chance to receive college funding (Stanley & Baines, 2004). Grades should simply reflect academic performance towards learning goals (Randall & Engelhard, 2010). The problem arises when grades are not just limited to communicating student achievement; they include self-esteem boosters, attitude, participation, and rewards (Stanley & Baines, 2004).

The pressure for successful performance on what would be considered high-stakes testing has increased the necessity for dependable measures of academic achievement that are accurately communicated to parents, teachers, and administrators. There is agreement in the education community that teacher assigned grades can be a reliable and effective tool for communicating a student's academic progress. Guskey (2007), found that diverse stakeholders perceive legitimacy of achievement indicators differently. Administrators view state, district, and national standardized assessments as a trustworthy source of academic achievement, whereas teachers perceived classroom observations and homework more trustworthy than administrators.

Frisbie and Waltman (1992), identified the purpose for grading by teachers, parents, and students. Their findings can be classified into six broad categories: (1) to communicate the achievement status of students to parents or others, (2) to provide information for student self-evaluation, (3) to identify certain pathways or instruction in education, (4) to provide learning motivation and incentives for students, (5) to evaluate

the effectiveness of instructional programs, and (6) to provide evidence of student effort or inappropriate accountability. In 2002, Guskey also sought to determine similarities and differences in all three stakeholders: parents, students, and teachers. Guskey wanted to determine their perceptions on the purpose of grading and reporting. Stakeholders ranked the six major purposes of grading mentioned above. Each group ranked “communication to parents” and “feedback to students” as the most important purpose for grading. The two least important purposes were “evaluation of instructional programs” and “lack of effort and accountability” (pp. 5-7).

The Effects of Grading On Students

Grades have been linked to having a strong and lasting impact on a student’s attitude, behavior, and motivation for learning (Brookhart, 1994). Researchers refer to three consistent effects that arise when an emphasis is placed on the importance of letter or number grades. Assigning an arbitrary letter or number grade tends to: (1) reduce the student’s interest in actual learning, (2) increase the frequency of students choosing the easiest task, and (3) lessen the quality of the student’s thinking (Brookhart, 1994; Kohn, 1999).

Research has shown that “focus on grades” and “focus on learning” are opposite of each other (Kohn, 1999). Kohn (1999), moreover, states that when students are told they need to know something, they tend to lose interest. If a student is focusing on the grade and there is pressure to receive the highest grade, the easiest intellectual path will produce this result. The end result being a reduction in thinking due to lack of interest and poor intellectual exploration.

Teachers do understand the importance of grades for students and parents. Students who receive good marks may get paid by parents and get to attend reward celebrations. Good grades build student self-worth. These are all factors that impact a teacher's ability to assess based on only achievement (Randall & Engelhard, 2010).

Common Grading Design

More detailed reporting methods such as checklists or narratives have their disadvantages too. These reporting methods offer more specific information on student achievement, but take greater amounts of time to prepare. More detailed reporting systems that increase the analytical process, are more likely to lead to subjectivity that influences grades (Ornstein, 1994). However, not all subjectivity in grading is bad. The teacher knows the students the best, understands the limitations of the work completed, and has knowledge of the progress made in class. This may produce a more accurate picture of the students' current academic performance (Brookhart, 1993).

Kohn (1999) would argue that several research studies found that students given numerical grades are far less creative than those students who received qualitative feedback. A combination of comments and numerical grading did not help attain high achievement. Students achieved the highest when comments were given instead of letter or number grades (Kohn, 1999).

Achievement Factors versus Non-achievement Factors

According to Reeves (2011) and Guskey (2015), traditional grading systems are inconsistent inaccurately conveying how students perform in relation to learning standards. Grades are influenced by a variety of factors that include effort and

behavior. In traditional systems that use numbers or letters to represent achievement, grades only slightly relate to performance on high-stakes external assessments. Today's education system demands large amounts of accountability. This was evident in NCLB and has continued importance in ESSA. Grading systems must reflect an accurate measure of students' learning towards the content standards. Many school districts have initiated standards-based grading practices in the hope of gaining a more reliable measure of student mastery of adopted learning standards.

Shift to Standards-Based Grading

Traditionally, grades have been constructed on assessment methods designed by educators and are based on comparing an individual student with a group. For grades to have meaning, there must be a clear understanding and a point of reference to compare student achievement. With the adoption of state and national student performance standards, grades that are based on standards will allow educators to use a criterion-referenced approach in assessing student achievement (Guskey, 2015; O'Connor, 2009).

To combat grading issues and misinterpretations of grades, Thomas Guskey (2015) and Ken O'Connor (2009), both outline a criterion-referenced grading system based on standards. Standards-based grading is where grades are strictly based on learning outcomes or performance standards that create a clear portrait on what students should know and be able to do. These standards create greater equity on learning outcomes for all students through consistent communication about student achievement among stakeholders (Schmoker, 2000).

Guskey (2001) compares norm-reference or traditional percentage grading to criterion-referenced or standards grading. He emphasizes the obligation a teacher has to communicate accurately a student's mastery towards specific learning goals. Traditional norm-referenced grading practices do little to communicate a clear picture of student performance and furthermore, create an atmosphere of "winners and losers."

A strong standards-based grading system includes many attributes. It views grading as a process. Quality criterion-referenced performance standards should be the reference point to determine student grades. Value or judgment attributes should be limited, and not all student work samples should be included in grades. Students should be allowed to work towards mastery, keeping grades written in pencil for the possibility of improvement. All grading procedures should align with learning goals. Standards-based grading is a teaching and learning process that involves properly recorded evidence of student achievement (Guskey, 2015; O'Connor, 2009).

One goal of standards-based grading systems, is to remove subjectivity in grading, thus providing reliable information about student learning (Hardegree, 2012). This was verified in a 2012 study by Hardegree. The study revealed that standards-based report card grades accurately correlate to the students' performance on a required high-stakes standards-based state assessment. In fact, the teacher assigned standards-based report card grades, exceeded the requirements of proficiency on the state assessment. By focusing teacher grade reporting and assessments on the standards, educators continue to increase the reliability of grading practices and provide

a clearer picture of student learning that is necessary to increase student achievement (Hardegree, 2012).

Students enter a classroom with varying backgrounds, prior knowledge, and rates of achievement. When averaging grades at the end of the quarter is the indication of performance, students who master material at the end of the quarter are penalized, even though their end performance may be greater than their counterparts. Giving students the opportunity to meet learning goals through continued practice allows for better obtainment of information to denote what a student has learned and what a student is able to do (Guskey, 2001; Guskey, 2015).

Grading Reliability Factors

Reeves (2008, 2011) expressed it is not new curriculum, replacement of a principal or teachers, or great technology that will improve schools; it is simply the need for a better grading system. He further stated that policies have to be set that require teachers to calculate grades solely on academic performance. Even teachers that agree with grading systems that only reflect academic performance, struggle to keep non-achievement factors out of their grading practices. In reinforcement of this statement, 81% of teachers polled and 70% of students agreed or tended to agree with the statement that achievement should be reported separately from other factors (Cross & Frary, 1999). Even though a high majority of the teachers expressed a belief in the statement, their actions do not indicate agreement. Seventy-two percent of the teachers in the same study indicated they raised the grade of low-ability students based on other factors than achievement (Cross & Frary, 1999). Similar results were found

when teachers professed to adhere to grading practices that were aligned with best practice research on grading; however, when they replied to a grading survey their responses indicated differently (Steidinger, 2011).

A student's characteristics have also been found to impact the consistency of grades. A teacher's perception of a student's behavior can significantly influence judgments of their academic performance. Four major factors are considered by teachers when assigning a final grade: Student academic achievement, student ability, student behavior, and student effort (Randall, & Engelhard, 2010; Südkamp, Kaiser, & Möller, 2012). A teacher's perception of a student's behavior can significantly influence the reporting of a student's academic performance (Hills, 1991). Even the neatness of a student's handwriting can influence a student's grade (Sweedler-Brown, 1985; Steidinger, 2011).

Brookhart (1993) demonstrates how value judgment and subjectivity can impact a student's grade. Teachers in a study were directed to assign a grade in two different situations. An average Algebra I student recorded grades on two tests for the grading period. On the first test he achieved an F and on the second test he achieved a low D. The teachers were asked to assign the student an overall grade for the period. The choices were an overall grade of an F based on the average of the two tests or an overall D because improvement of performance was demonstrated. Seventy-three percent of the teachers chose the D. The second situation was similar, except this Algebra I student achieved a B on his first test and a low A on the second test. The choices for a final grade were an overall grade of B which was the average of the two test grades or

an A with the consideration there was improvement. With an identical percentage, this time the teachers chose B as the grade (Brookhart, 1993).

Randall and Engelhard (2010) shared their examination of factors that impact borderline decisions in grade assignment. A student with a report grade of sixty-nine percent that demonstrates low achievement and low ability, but offers good behavior with high effort, on average receives a grade of 77%. This solidifies the thought that teachers reward lower achieving students at a higher grading rate due to good behavior and effort. This was not only true for low achieving students. Students that were on the borderline of achieving an A or a B, were consistently receiving the higher grade when their effort and behavior were excellent. Regardless of ability, students' grades improve with good behavior or effort (Randall & Engelhard, 2010).

Separating learning from attendance, behavior, effort, and other peripheral issues, educators can measure if students are actually learning, not just playing the game of grading. The goal of standards-based assessment is to provide a framework for all stakeholders to understand learning and maintain a focus on teaching and learning of communicated standards (Guskey, 2001; Scriffiny, 2008).

Standardized Testing

Standardized tests have been around for many generations. Never have these tests been given with such frequency and played such a large role in communicating student progress. The test outcomes are used to make important decisions that impact a student's educational career and why they are referred to as high-stakes. Common practice internationally is not to give standardized tests to students prior to their

sixteenth birthday. However, in the United States, tests are administered to students as young as six years of age. Experts of early childhood learning criticize the practice of testing at such an early age (Kohn, 2000).

Proponents of standardization argue that these tests can overcome subjectivity, which in return creates a more accurate evaluation of student knowledge and performance (Kohn, 2000). They also claim standardized tests supply evidence that interprets a student's knowledge and skill. The results of these tests are very useful in understanding student performance as compared to their counterparts (Popham, 1999).

Opponents feel the first fault of reliability for a standardized test is revealed in the creation of the test. Humans, with biased and varying opinions, formulate test questions and determine acceptable answers. Secondly, the validity of the scores can be impacted by the experience of the students. The prevalence of student test anxiety or a student who disregards the importance of the standardized test can significantly impact the outcome of a single evaluation of learning (Kohn, 2000). A final argument frequently discussed is the mismatch between what is being taught in the classroom and what is being measured on a standardized test. The localization of curriculum decisions makes it virtually impossible to create a standardized assessment that meets the needs of all learning institutions (Popham, 1999).

Standardized tests have revealed results that indicate a difference in performance for students based on their family socioeconomic status. Research repeatedly indicates that a school's poverty level, along with other variables, account for a high percentage of the difference in student test scores (Kohn, 2000). For some,

this demonstrates why standardized tests may be perceived as a poor measurement of learning and skill attainment.

Standardized test performance, like teacher assigned grades, are used as a primary source for judging student performance in the classroom. Education stakeholders challenge whether standardized tests provide an objective measure of learning or are a useful tool to improve teaching. The more knowledge that is gained on standardized tests, the more inclined the education community will be to make informed decision if these tests are an accurate measure of student achievement (Kohn, 2000).

Correlation of Teacher Performance Judgments and High-Stakes Testing

The idea of comparing a student to a set standard seems relatively simple. However, research shows variance in how standards are interpreted and presented to students can impact their achievement on high-stakes tests (Hill, 2001). D'Agostino, Welsh, and Corson (2007) found the degree to which a student experiences success towards a standard is impacted by a teacher's commitment to provide students the opportunity to learn the standard. Clearly, curriculum, instruction, and assessment should be well aligned and presented in the classroom in a similar way. Resnick (2006) warns there are dangers in standardized testing and the pressure high-stakes testing creates. Teachers find themselves teaching to the test and spending valuable instruction time in test-prep activities at the expense of valuable educational experiences to obtain goals set by mandated testing.

Not only is a teacher's ability to effectively communicate student performance questioned, Frederiksen and White (2004) discuss issues concerning the dependability of developing standards and assessments used to gauge student performance towards the standards. They question whether the high-stakes test can measure understanding and determine the necessary skills that may be important for success in future learning. Finally, they address the concern for the use of cut scores to identify students who have reached a level of proficiency and those who have not.

When teachers are asked to predict student performance on a standardized test, their judgement is not always found to be accurate. Teacher judgement accuracy has been found to differ between low and high achieving students. Teachers were able to accurately judge student responses for three quarters of the time on individual test items, but accuracy declined within subject subtests. Teachers were the least accurate when judging low achieving students, increasing accuracy for the high achievers. This was determined to be true because the high achieving student would answer a large amount of the questions correctly and the teacher assumed automatically the high achiever was efficient. This thinking would not apply to the lower achieving student since the student would answer many of the questions wrong. The implications of this are the students in greatest need of accurate performance appraisal and education assistance are the students who are at the greatest risk of being misjudged and in return, may lack the support they need to increase testing performance (Coladarci, 1986).

Teacher Experience and Grading Practices

There is limited research about the effect of teaching experience on grading reliability to communicate student performance. The majority of research in this area examines the impact assessment training has on a teacher's ability to accurately evaluate student learning; excluding the influence of teaching experience. Studies comparing a teacher's years of experience and grading accuracy have yielded conflicting results.

Receiving a fair amount of training on how to assign grades and score student performance can positively influence a teacher's understanding of assessing student performance. However, training on rating student performance does not necessarily assure consistency in grading (Schafer, Swanson, Bene, & Newberry, 2001; Stuhlmann, Daniel, Dellinger, Denny, & Powers, 1999). Factors such as teacher background and experience in evaluation have shown to affect grading practices; which can produce a lack of uniformity between student performance judgments (Schafer et al., 2001; Eckes, 2008).

A recent study conducted by Kan and Bulut (2014) examined the effects of teacher experience and the consistency of assigning scores. The study supported that the level of teacher experience has a considerable effect on the way teachers approach assessment. Teachers who have more experience tend to give grades more leniently than teachers with narrow experience. Contrary to this, Myford and Mislevy (1994) and Meyer (2000) found that teaching experience had an insignificant impact of rater imposed stringency on performance.

The inconsistency of findings in regard to teacher-related experience in grading, perpetuates the need for a greater understanding of how to improve current grading structures. The next section will suggest research-based practices that reduce the potential of negative impacts on grading, therefore increasing the consistency of student achievement communication.

Practices to Improve Grading

Regardless of teacher training on assessment, when assigning grades, most consider ability and effort (Brookhart, 1993). Forty-seven percent of elementary teachers reported using ability when assigning grades, especially when making borderline grading decisions (McMillan, Myran, & Workman, 2002). When a classroom teacher decides what factors will impact grading, one's personal training, official grading policies, and perceived consequence of assigned grades influence the outcome (Randall & Engelhard, 2010).

Few school leaders have extensive knowledge of the attributes of various grading methods and the impact different grading policies have on students (Brookhart, 2011, Brookhart & Nitko, 2008; Stiggins, 1993; Stiggins & Chappuis, 2011). An exercise that has produced higher levels of consistency in grading practices is the training of personnel, directly involved in the grading process, on the school district's philosophy and procedures for assigning grades (Mehring, et al., 1991). This conclusion was reached through a study that sought to define whether training on grading procedures would impact the degree of variability for grade assignment. Elementary principals and teachers in a large suburban school district were asked to complete a questionnaire

before and after they received training on their current report cards which contained two marking systems.

In the case of both marking systems, pre-training results indicated that the majority of the participants felt the marking systems were unclear and they were dissatisfied with the systems. Post-training, a difference in perception was collective. The majority of survey participants indicated the marking systems clearly communicated the students' present levels and they were in favor of retaining the current grading systems. Further noted, before training, vast differences in interpretation of both marking systems were evident. Greater consistency in assigning grades was achieved when teachers implemented a shared philosophy and defined criteria for assessing student achievement (Mehring et al., 1991).

Kan and Bulut (2014) support the necessity of establishing grading guidelines and communicating them system-wide. The study found that when teachers have a lack of guidance in grading, they will establish their own criteria to assess student performance which causes reliability and consistency to decrease among evaluators. The difference in teacher evaluation of student performance was negated when teachers used an established grading criteria in assessing student performance and received training on the criteria (Kan & Bulut, 2014; Schafer et al., 2001).

CHAPTER THREE

METHODS

This exploratory mixed method research study measured fourth grade student achievement through teacher assigned report card percentage grades and student performance on the Nebraska State Accountability Assessment, known as NeSA. This study compared scale scores on the NeSA in the content areas of math, reading, and writing to teacher assigned percentage grades, in the corresponding subject areas of math, reading, and language arts. Percentage grades which were converted to a letter grade are determined by an averaged percentage of graded work chosen and recorded by the classroom teacher. NeSA, a criterion-reference test developed by the State of Nebraska Department of Education, was given in the spring to all students in third through eighth and eleventh grades.

Collected within the study was the fourth grade teacher perceptions of grading and assessment practices that impact clear and consistent communication of student achievement. Teacher perception data was obtained in written format from the eight teacher study participants through the distribution of an electronic questionnaire. Key themes that emerged from the questionnaire were identified to understand how the beliefs, experiences, and practices of teachers impact clear and accurate student performance communication.

Under NCLB, which is now reorganized as ESSA, the Nebraska State Accountability Assessment is used to determine if schools are successfully educating their students. The law requires states to use a single accountability system for public

schools to determine whether all students, as well as individual subgroups of students, are making progress toward meeting state academic content standards. The goal of ESSA is to ensure equity and opportunity for all students in the United States.

NeSA is considered a measure of student learning towards the Nebraska State Standards. District and individual school performance on the NeSA assessments are reported and ranked within the State of Nebraska. The rankings identify high performing school districts and individual schools throughout the State. Cut scores assign students to one of three categories: Exceeds standards, meets standards, and below standards. The percentage of students exceeding or meeting the cut scores on the NeSA Assessment is then calculated. This percentage communicates a numerical amount of students who have met the proficient level of academic performance.

As stated by the research referenced in Chapter Two of this study, the accuracy of traditional norm-referenced grading systems to communicate student performance are under question. School systems frequently use teacher assigned grading systems that are assigned a letter that corresponds with a percentage or subjective opinion of a student's academic performance. The single letter grade or percentage is used to communicate student performance and develop an understanding of student academic achievement. Districts use report card grades as a summative measurement for student achievement of curricular learning, based on the teaching of the Nebraska State Standards. If the teacher-assigned letter grades or percentage grades do not accurately measure student performance and paint a false picture of proficiency attainment, the

outcomes could be detrimental which may lead to poor performance on high-stakes tests.

The purpose of this study was to investigate to what extent fourth grade teacher assigned report card percentage grades and student performance on the fourth grade Nebraska State Accountability Test co-vary for reading, math, and writing. In addition, further exploration was conducted to determine if teacher perceptions on communication of student achievement is significantly related to the congruency of assessed student performance in the classroom and student performance on standardized tests. The findings were used to explore if grades align with student performance on standardized tests and do they communicate a clear picture of student achievement to stakeholders.

This section of Chapter Three describes the sample population selected and instruments that were used for data collection. Included are the methods, materials, and procedures utilized to collect data, and the statistical procedures involved in the analysis of data collection.

Research Design

An exploratory mixed method research study was used to communicate to what extent two variables; teacher assigned report card percentage grades and student performance on the Nebraska State Accountability Assessment, co-vary. Correlation research was the appropriate design to explain the relationship between the two variables. This statistical test was used to describe and measure the relationship between the two sets of scores. The scores of this study were not controlled or

manipulated. This is a characteristic of correlation research. In this correlation design, the researcher used a correlation statistical test to investigate the pattern or tendency between report card grades that corresponds to a student's performance on the NeSA assessments (Creswell, 2012).

The exploratory research design was used to further investigate the patterns or tendencies of teacher belief, perception, and practice on student achievement communication. In this exploratory research the variable of different content areas was used to understand varying outcomes of student performance communication by content areas. All analysis was compared using a scatterplot to provide a visual picture of patterns and association (Creswell, 2012).

This study was a mixed method study that included the above mentioned quantitative analysis of student performance. To further understand the impact of teacher grading practices a qualitative questionnaire was administered to the eight teacher study participants.

Research Questions

This study explored the relationship between teacher assigned report card percentage grades and student performance on the Nebraska State Accountability Assessment. The essential question of this research was: Do traditional grading practices and standardized testing accurately measure and communicate student academic performance? The research questions that guided this study are:

Research Question 1. How strong is the relationship between teacher assigned student report card grades in reading and scores on the Nebraska State Accountability Reading Assessment for students enrolled in fourth grade?

Research Question 2. How strong is the relationship between teacher assigned student report card grades in math and scores on the Nebraska State Accountability Math Assessment for students enrolled in fourth grade?

Research Question 3. How strong is the relationship between teacher assigned student report card grades in language arts and scores on the Nebraska State Accountability Writing Assessment for students enrolled in fourth grade?

Research Question 4. How do fourth grade teachers perceive grading and assessment practices to clearly and consistently communicate student achievement?

Setting

This study examined students from a growing northwest Omaha Metropolitan area school district in full transition from a rural school district into a suburban school district with a 10.07% average increase in student population for the past five years. The District consisted of three kindergarten through sixth grade elementary schools, and one seventh through twelfth grade secondary school. The district-wide student population totaled 1,884 kindergarten through twelfth grade students. Characteristics of the system included limited diversity and high socio-economic status with a system-wide free or reduced lunch rate of 10.60%.

The district demographics were comparable to the 171 system-wide fourth grade students enrolled for the 2014-2015 school year. The student population enrolled in

fourth grade was predominately white at 87.1%. The remainder of the student population consisted of 5.8% Hispanic, 1.2 % Black, and 5.9% other or multi-racial. A greater percentage of the fourth grade population was identified as gifted learners at 17%. Fifteen and eight tenths percent of the students received special education services.

The school district recorded a less experienced teacher populace at an average of 11.41 years compared to the Nebraska state average of 14.34 years. Even though the system had less teacher experience, 70.73% of teacher's had a Master's Degree. That greatly exceeded the state average of 52.16%.

Study Participants

The target population for this study was fourth grade students enrolled in the school system during the 2014-2015 school year. The sample encompassed a total enrollment of 171 fourth grade students drawn from the three elementary schools in the system. The office of curriculum and assessment anonymously identified the students who enrolled in the school system the entire 2014-2015 school year and accumulated less than ten absences during the year. Absences that accrued more than ten days resulted in the student missing content instruction and test preparation which could impact student achievement data. The students were required to complete the entire battery of NeSA assessments. The population contained students who received special education as long as they were able to complete the entire battery of NeSA assessments. No restrictions of participation were enforced for those students that received accommodations, as stated in an Individual Education Plan, as long as the

student was not identified as taking an alternative assessment. The district office of curriculum and assessment provided third quarter report card grades and NeSA scores for the system wide fourth grade students who meet the above mentioned qualifications.

The teacher study group consisted of all teachers in the District assigned to a fourth grade classroom during the 2014-2015 school year. The eight teachers were all female, fulltime employees. The teachers' years of experience was: three teachers with three or less years of experience; three teachers with more than three, but ten or less years of experience; and two teachers with more than ten years of experience. The percentage of the fourth grade teachers in the system that had a Master's degree was 62.5%. This was above the Nebraska State average and below the District average.

Study Instruments

Nebraska State Accountability Assessment. This study used scores from the Nebraska State Accountability Assessment (NeSA), a standardized test required by the Nebraska Department of Education. Students in grades three through eight and eleven were required to take NeSA as a summative assessment in the areas of reading and math. Additionally, students in fourth, eighth, and eleventh grades were required to take NeSA-Writing. In the content area of science, grade levels fifth, eighth, and eleventh participated in NeSA testing. For this study, the researcher addressed the reading, math, and writing sections of the NeSA Assessment to determine the strength of the correlation between traditional teacher assigned report card grades in correspondence with content areas for all fourth grade students district wide.

The NeSA Assessment was a criterion-referenced test designed to examine how well students learn knowledge and skills within a specific content area, only testing achievement towards the Nebraska State Standards. Whereas, norm-reference tests highlight differences between students on an academic continuum. The required state assessment was designed to reveal how well all students in the State acquire the skills defined by the Nebraska Content Standards. The NeSA Assessment scores were used to

	Exceeds Standards (3)	Meets Standards (2)	Below Standards (1)
NeSA Reading	135 – 200	85 - 134	84 and below
NeSA Math	135 – 200	85 - 134	84 and below
NeSA Writing	57 – 70	40 - 56	39 and below

communicate academic achievement at student, class, school, district, and state level.

Student NeSA Assessment scores were ranked into two high levels and a third lower level. All students were expected to reach a level of proficiency, which was defined by attaining the two higher level rankings on the mandated state assessments. The NeSA scores were reported using the following measure based on cut scores for performance levels constructed on scaled scores: 1= below standard, 2= meets standard, or 3= exceeds standard (see Figure 2).

Figure 2. Spring 2015 grade four NeSA performance level ranges.

Reliability was one of the key aspects of testing quality that ensures that the same measurement or comparable result were received for the same student every time. The reliability of the NeSA Assessment was evaluated using statistical methods.

Reliability referred to the degree in which test scores, for the same group of testers, were consistent over multiple measures (DRC, 2015).

The reliability index used for the 2015 NeSA was Coefficient Alpha α . According to the Data Recognition Corporation, acceptable α levels usually range in the mid to high 0.80s to low 0.90s. All NeSA-Reading and NeSA-Mathematics Coefficient Alphas for the whole population ranged from the high 0.80s or low 90s. Specifically, the fourth grade NeSA-Reading α value was 0.89 and the fourth grade NeSA-Mathematics α value was 0.92. These α values provided evidence that the NeSA Assessment demonstrated good reliability.

The scoring of the NeSA-Writing involved two independent readers which may have caused random error due to measurement procedures and interpretation of measurement results by independent readers. To address this source of error, strict procedures in reader recruitment, reader training, and validity control were imposed. As a result, the degree of agreement among readers was acceptable at or above a 75% exact agreement rate. The acceptable consistency rate for individual readers, over time, was at or above 80% for exact agreement. The fourth grade reader decision accuracy on the 2015 NeSA- Writing Assessment was calculated at a rate of 90-92% accuracy. Individual reader consistency, over time, ranged from 86%-89% (DCR, 2015).

The other vital component of technical quality in testing was validity. Establishment of validity began with the purpose of the assessment and continued through item writing and review. With the NeSA-Reading and NeSA-Mathematics Assessments, which were criterion-reference standards-based assessments, strong

content validity evidence was a direct derivative of the test development process. The process ensured every test item aligned directly to the content standards. This alignment was the basis for all item writing on the assessment. Editors and review committees evaluated the alignment and made revisions when necessary. This resulted in consensus among developers and consumers that the test assessed what was intended (DRC, 2015).

The NeSA-Reading and NeSA-Mathematics Assessment was checked for validity based on internal structures. Correlations between test items, measured using the Pearson's Correlation Coefficients between strands within content areas, were reported. This resulted in positive intercorrelations between content area strands. The intercorrelations ranged from moderate to high in value. Disattenuated correlations were used to measure across content area strand correlation. Consistently, the correlations between the strands within each content area were higher than the correlations between strands across different content areas. Within content area strands, correlations were typically greater than 0.90, while across content area strand correlations ranged from 0.75 to 0.92. This pattern was expected since the content area tests were designed to measure different constructs. Overall, this data indicated good item discrimination and evidence of internal-structure reliability (DRC, 2015).

The NeSA-Writing Assessment was a criterion-reference test. The assessment was based on and directly aligned to the Nebraska Writing Content Standards. Similar as described above for the NeSA-Reading and NeSA-Math, item development and test creation process ensured that every test item aligned directly to the Nebraska State

Content Standards. A review committee checked alignment of the standards, which resulted in mutual agreement, among content specialists and teachers that the NeSA-Writing Assessment measured what was intended (DCR, 2015).

Teacher assigned report card grades. The school district used a percentage grading system communicated by a letter grade on a report card distributed four times annually. Report cards were constructed electronically by the classroom teacher. Student daily work and assessments were recorded into an online grading system called Student Information Management System (SIMS). Work samples were chosen by individual classroom teachers and may have varied throughout the school system. Report card grades were calculated by the total number of points accumulated by a student divided by the total number of points possible. This resulted in a percentage grade that was converted into a letter grade. Students' in third through twelfth grade achievement was calculated and communicated on a report card using the following grading scale (see Figure 3):

Mark	High	Low
A+	100	98.50
A	98.40	94.50
A-	94.40	92.50
B+	92.40	90.50
B	90.40	87.50
B-	87.40	85.50
C+	85.40	83.50
C	83.40	78.50
C-	78.40	76.50
D+	76.40	74.50
D	74.40	71.50
D-	71.40	69.50
F	69.40	0

Figure 3. Research District report card grade scale.

Procedures

With the assistance of the school district personnel, the researcher used the student information system, SIMS, to access district-wide fourth grade student report card information in the content areas of reading, math, and language arts. Through assistance from the district curriculum and assessment director, the Nebraska Department of Education NeSA Assessment portal was accessed to attain student performance data in the content areas of NeSA Reading, Math, and Writing Assessment for all fourth grade students system-wide. Teacher data information was obtained through the Office of Personnel. The researcher used no personal identifiers, so attainment of parental permission for the subjects was not necessary. All school information, teacher information, and personal information was removed from the data to comply with FERPA and protect the privacy of the study participants.

After obtaining the published 2014-2015 NeSA Assessment scores and third quarter teacher assigned report card grades, the researcher eliminated students who were not enrolled in the system the entire 2014-2015 school year, students who had accumulated 10 or more absences during the school year, students that had not completed the entire battery of tests, and students who had participated in an alternative assessment as stated by their Individualized Education Plan. The collected data for teacher assigned student third quarter report card grades and NeSA Assessment scores, for the students who met the study criteria, were entered into an excel spread sheet (No personal identification for students, teachers, or the schools were used). Teacher assigned third quarter report cards were examined to calculate the degree of correlation between teacher assigned report card percentage grades and student performance on the NeSA Assessment.

Data Analysis

A Pearson Product Moment Correlation Test was employed to analyze the linear relationship between report card grades and student performance on the NeSA Assessments. The statistical test was used to describe and measure the degree of the relationship between teacher assigned report card grades and student performance on the Nebraska State Accountability Assessment for research questions one through three. The strength of the correlation was calculated to determine if teacher assigned grades were congruent with student performance on the standardized Nebraska State Accountability Assessment.

The Pearson's Correlation was an appropriate test to define the linear relationship between the two variables, fourth grade student's report card percentage grades in reading and scale scores on the Nebraska State Accountability Reading Assessment. When significant, the relationship was found to be positive or negative. The linear measurement between the two variables was used to build understanding of the accuracy and consistency of the performance data the research school system was using to communicate student achievement.

Further data was collected on the belief, perception, and practice of the traditional grading system implemented in the research district. All teachers assigned to a fourth grade classroom during the 2014-2015 school year were emailed an electronic questionnaire. Teachers were asked to record their response to seven open-ended questions revealing information about the implemented traditional grading system. Frequencies and key themes that emerged from the eight teacher's written questionnaire responses were recorded and compared to the correlation data. This allowed the researcher to investigate the impact the patterns or tendencies of the teachers had on communication of student achievement (Appendix 1).

Summary

This chapter described the methods and procedures executed to examine comparability of report card assigned grades in relationship to student performance on (NeSA), in specific content areas. Secondly, it indicated how this mixed method study collected the beliefs, perceptions, and practices of the teacher study participants in regard to their traditional grading system. The problem, research design, research

questions, setting, study participants, test instruments, data collection, and data analysis were presented. The main purpose of the study, was to determine if a traditional grading system congruently aligned with student achievement on a criterion-reference, standardized assessment. In Chapter Four, the presentation of data addresses the research questions and reveals an answer to the purpose of this study. A summary, discussion, and conclusion of the findings is included in Chapter Five. Additionally, implications for practice and recommendations for improved student performance communication will be offered.

CHAPTER FOUR

RESULTS FOR THE QUANTITATIVE AND QUALITATIVE FINDINGS

The purpose of Chapter Four is to report the data collected and communicate the findings of the relationship between teacher assigned report card grades and student performance on the criterion-reference Nebraska State Accountability Assessment. This chapter first reports the results of the statistical analysis of variance in the content areas of reading, math, and language arts on the two performance measures for students enrolled in fourth grade. The second part of this chapter reports fourth grade teacher perception of grading and assessment practices that impact clear and consistent communication of student achievement. Teacher perception data was collected in written format from the eight teacher study participants through the distribution of an electronic questionnaire. Key themes that emerged from the questionnaire were identified and are communicated in this chapter. The themes are indicated in a table along with teacher direct responses that support the identification of key themes. The final response on the questionnaire asked for teachers to provide any further information they felt beneficial for informing the research study on grading and communication of student achievement.

Overview of Results

The participants for the statistical analysis of this mixed method study were 171 fourth grade students drawn from the three elementary schools in the system. Students from the total population were eliminated from the study if they: were not enrolled in the school system the entire 2014-2015 school year, accumulated ten or more days of

absence resulting in the student missing content instruction, did not complete the entire battery of NeSA Assessments, and were identified as taking an alternative assessment. The end sample encompassed a total of 137 fourth grade students enrolled in the school system during the 2014-2015 school year that met the above mentioned criteria. A Pearson Product Moment Correlation Test was employed to analyze the linear relationship between report card grades and student performance on the NeSA Assessments. An alpha level of .05 was imposed to help elevate Type I error.

The Pearson Product Moment Correlation was measured on a standard scale representing the effect size. It communicated the strength of the relationship between the two variables, student report card grades and student performance on the NeSA Assessments. Cohen and Manion's guide that describes the strength of the relationship between two variables was used to interpret the size of the coefficients in this study. In Cohen and Manion's correlation a coefficient of .20-.35 was thought to represent a slight statistical significance; a correlation coefficient of .35-.65 was considered limited or some correlation; a correlation coefficient of .66-.85 was considered a very good or strong correlation; and .86 and above the correlation coefficient was very strong and seldom achieved (Creswell, 2012).

Research Question 1. How strong is the relationship between teacher assigned student report card grades in reading and scores on the Nebraska State Accountability Reading Assessment for students enrolled in fourth grade? The statistical analysis of the correlation rate for fourth grade students' report card percentage grades in reading and scale scores on the Nebraska State Accountability Reading Assessment revealed a strong

correlation coefficient of .747. This correlation rate was the highest for all content areas. This information can be found in Table 1 and Table 2.

The reading correlation was then broken down and ranked by individual teacher. The highest correlation for an individual teacher was .877 and the lowest was .656. According to Cohen and Manion's description of effect size, all teachers recorded a strong correlation (Creswell, 2012). The correlation difference among the teachers was .221. The least amount of variance of correlation amongst the content areas.

Research Question 2. How strong is the relationship between teacher assigned student report card grades in math and scores on the Nebraska State Accountability Math Assessment for students enrolled in fourth grade? The statistical analysis of the correlation rate for fourth grade students' report card percentage grades in math and scale scores on the Nebraska State Accountability Math Assessment discovered a strong correlation coefficient of .664.

The math correlation level was minimally below the reading correlation at a difference of .083. Even though the math relationship of student achievement data fell below reading, a stronger correlation for math over language arts was revealed. This information can be found in Table 1 and Table 3.

The math correlation was then broken down and ranked by individual teacher. According to Cohen and Manion's description of effect size, five teachers recorded a strong correlation and three of the teachers documented limited or some correlation (Creswell, 2012). The highest correlation for an individual teacher was .870 and the lowest was .569. The strongest individual teacher math correlation was only .007 lower

than the strongest individual teacher reading correlation. The correlation difference among the teachers was .301. Whereas, the correlation difference among the teachers for reading was less at .221.

Research Question 3. How strong is the relationship between teacher assigned report card grades in language arts and scores on the Nebraska State Accountability Writing Assessment for students enrolled in fourth grade? The statistical analysis of the correlation rate for fourth grade students' report card percentage grades in language arts and scale scores on the Nebraska State Accountability Writing Assessment discovered a moderate correlation coefficient of .384.

The language arts coefficient level had the weakest correlation of all content areas in this study. Even though the relationship in language arts was of moderate strength, there was still a positive, linear relationship between teacher assigned report card grades in language arts and scores on the Nebraska State Accountability Writing Assessment. This information can be found in Table 1 and Table 4.

The language arts correlation was then broken down and ranked by individual teacher. The highest correlation for an individual teacher was .787 and the lowest was .134. The discrepancy of scores amongst the teachers was the highest for all content areas in this study. The spread of teacher language arts' correlations reached a difference of .653. Another inconsistency established from the analysis of the language arts' correlation was that the range of scores showed that two of the teachers achieved a strong relationship between the two sources of student achievement data; five of the teachers obtained a moderate or limited relationship between the two scores; and one

teacher did not demonstrate a relationship between the two scores. Once again, Cohen and Manion's description of effect size was used to communicate the relationship of individual teacher correlations (Creswell, 2012).

Table 1. Correlation of student performance on student report card percentage grades and the Nebraska State Accountability Assessment using a Pearson Product Moment.

	Reading Percentage Grade	Math Percentage Grade	Language Arts Percentage Grade
NeSA Reading Scale Score	.747		
NeSA Math Scale Score		.664	
NeSA Writing Scale Score			.384

Table 2. Scatterplot of student performance on student report card percentage grades and the Nebraska State Accountability Assessment for reading.

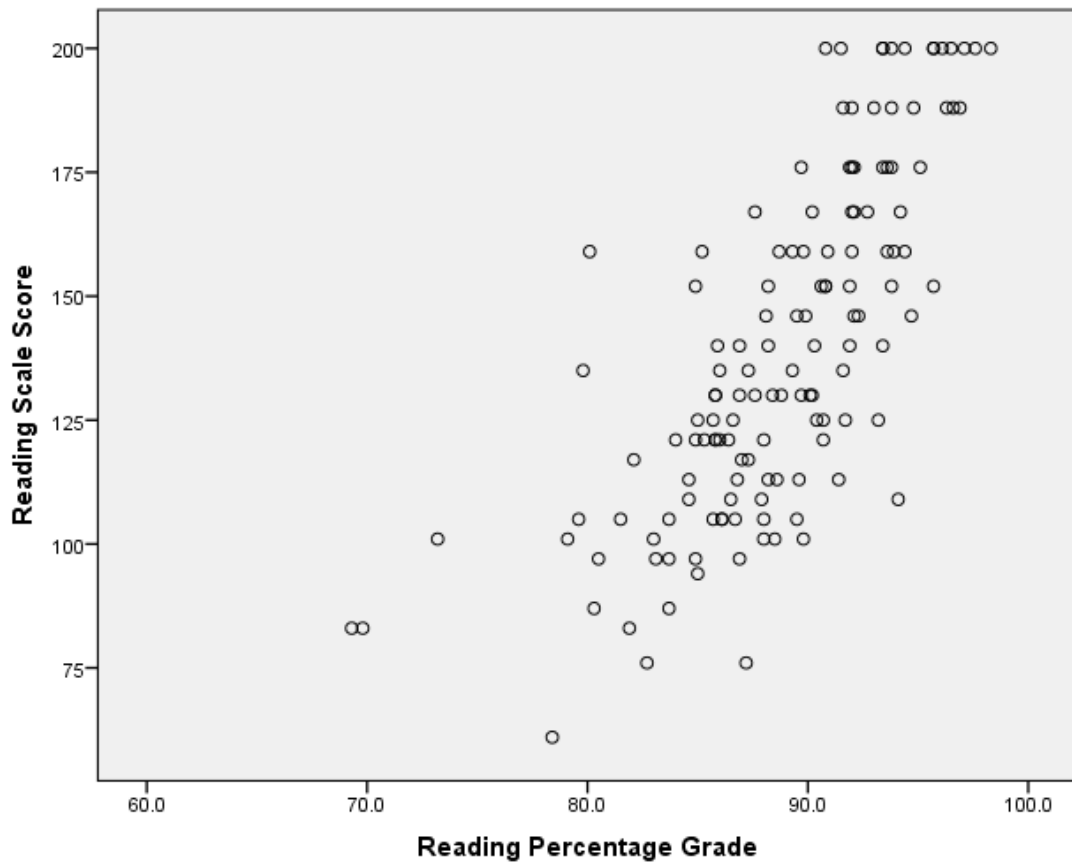


Table 3. Scatterplot of student performance on student report card percentage grades and the Nebraska State Accountability Assessment for mathematics.

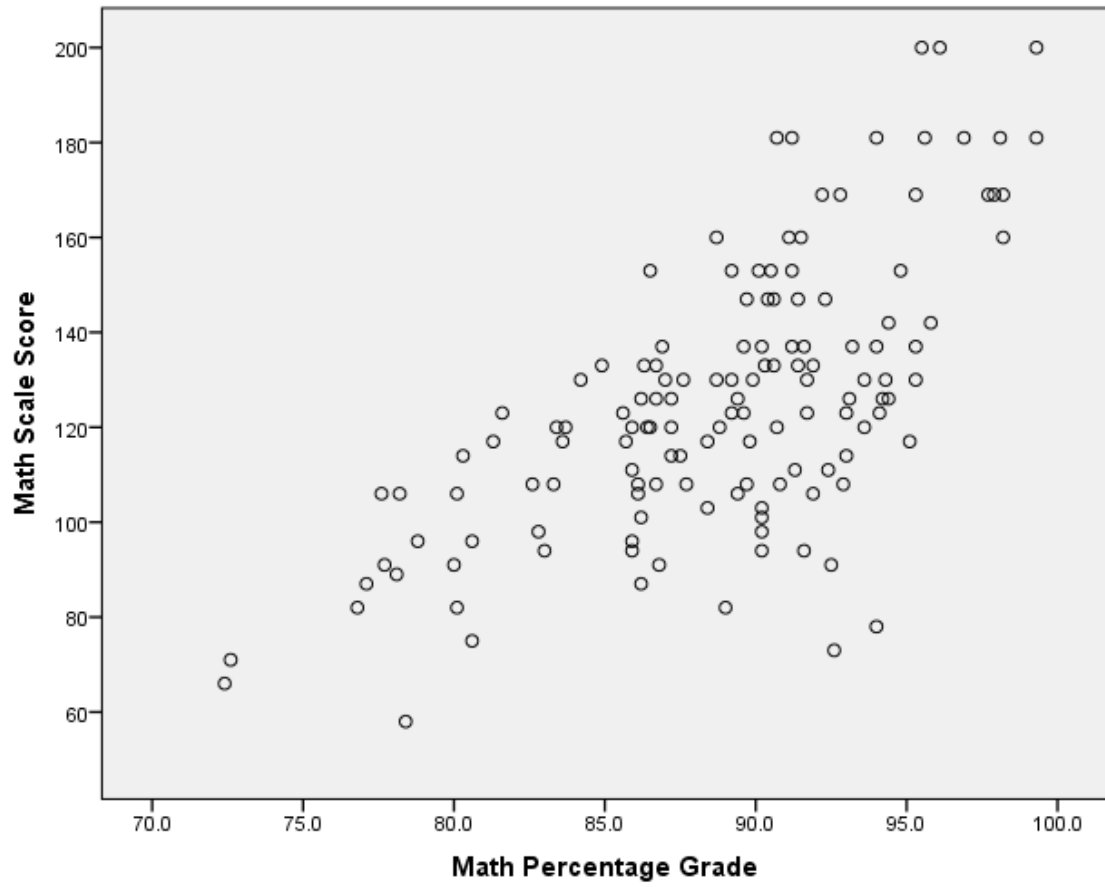


Table 4. Scatterplot of student performance on student report card percentage grades and the Nebraska State Accountability Assessment for language arts/writing.

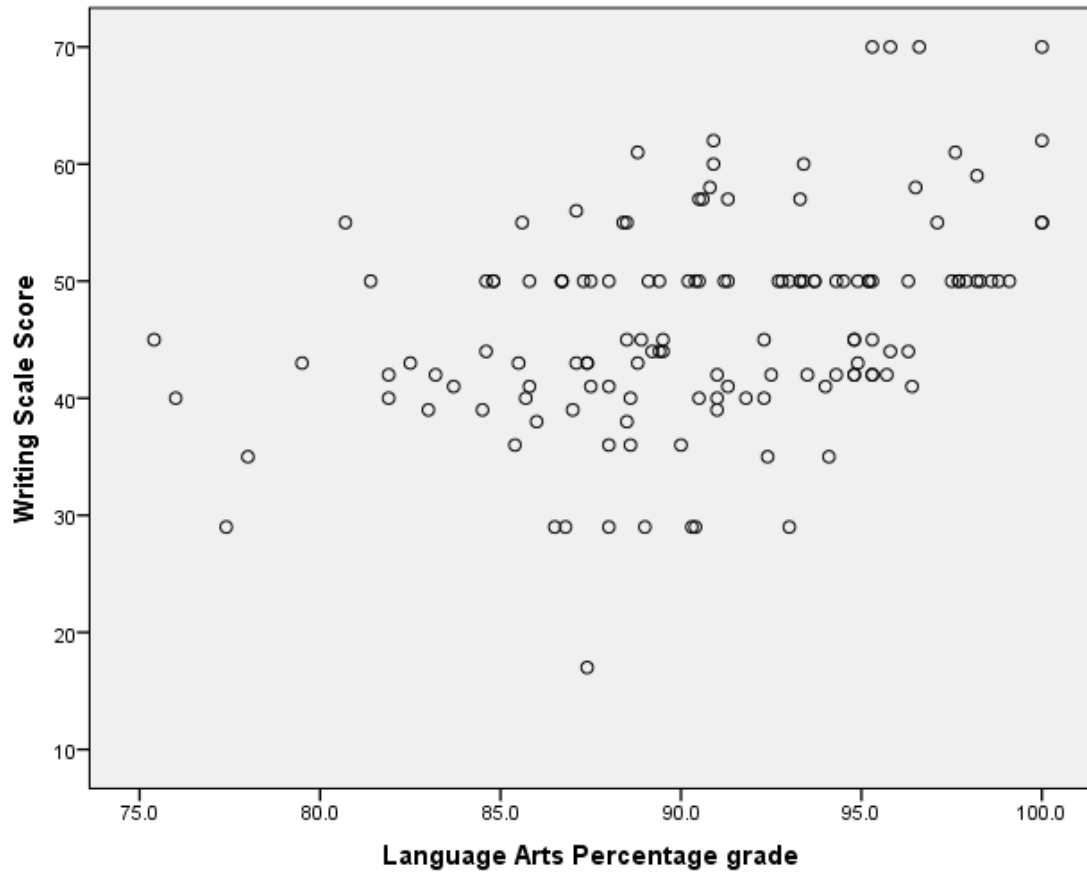


Table 5. Correlation by teacher of student performance on student report card percentage grades and the Nebraska State Accountability Assessment using a Pearson Product Moment.

Rank Order by Teacher for Strength of Correlation	Reading Percentage Grade/NeSA Reading Scale Score	Math Percentage Grade/NeSA Math Scale Score	Language Arts Percentage Grade/NeSA Writing Scale Score
Rank 1	.877	.870	.787
Rank 2	.823	.779	.705
Rank 3	.807	.772	.638
Rank 4	.789	.766	.520
Rank 5	.786	.737	.432
Rank 6	.771	.638	.387
Rank 7	.690	.601	.357
Rank 8	.656	.569	.134

This section of the mixed method study conveys the qualitative findings and records the reactions of eight teacher respondents. All District teachers assigned to a fourth grade classroom during the 2014-2015 school year participated in the qualitative portion of the study. The eight fourth grade teachers were female, fulltime employees. The teachers' years of experience consisted of: three teachers with three or less years of experience; three teachers with more than three, but ten or less years of experience; and two teachers with more than ten years of experience. The range of experience allowed for perspectives from teachers at different stages in their career. The fourth grade teachers in the system also possessed a high level of advanced degree opportunities with 62.5% of them attaining a Master's degree.

Research Question 4. How do fourth grade teachers perceive grading and assessment practices to clearly and consistently communicate student achievement? Frequencies of responses for the key themes that emerged and direct quotations from the eight teachers' written questionnaire responses were recorded and summarized. The frequency of responses was recorded in Tables 6-11.

Questionnaire Question 1. What is your perception of the assessment and grading practices currently implemented in your school district? The most common perception of grading from the eight questionnaire respondents was inconsistency. Six of the eight teacher study participants described assessment and grading practices as inconsistent. One teacher explained her negative perception of assessment and grading by stating,

My perception of assessment and grading practices in our District is that it is completely inconsistent...they differ grade level to grade level, school to school, and even teacher to teacher. There is nowhere in a handbook or anywhere that tells us what needs to be included...I had no idea what I was doing and there is so many things to include. I feel there isn't consistency for the special education students as well.

In addition to inconsistency in practice, another study participant talked about the lack of unification in forms of assessment. This response also revealed support for subjectivity in grading. The second highest reaction, where five out of the eight teachers wrote, subjectivity was a characteristic of grading practice.

Teachers differ in the way they implement forms of assessment and grading practices. Both formative and summative are used in assessment and grading practices. This can be beneficial in that it allows that teacher to grade and assess in such a way that the teacher sees as best in meeting the needs of their classroom...these differences also make it more difficult to understand what a student would need to do to earn a grade.

Two teachers expressed sufficiency in current assessment and grading practices. One teacher commented, "District grading is necessary. I don't think there is a perfect system, but I feel our District does a sufficient job." Three of the teachers expressed concern for the amount of assessments conducted, as three others stated that through

multiple pieces of evidence they have gained an accurate picture of student achievement.

In the School District my perception of assessment is we really see our student in full circle. As a teacher we assess students daily, hourly, and anytime a task is given...it may be a simple teacher observation. Overall, I think our District does a good job assessing our students so we can get a full picture of their learning.

Table 6. Teacher perception of the assessment and grading practices.

Teacher Perceptions of Assessment and Grading Practices	Number of Teacher Responses
Grading is inconsistent	6
Grading is subjective	5
Have too many assessments but they most accurately communicate student achievement	4
Grading is not reflective of individual student achievement	3
Use many different pieces of evidence to determine a grade	3
Too much emphasis on just the grade	3
No communicated practices or policies on grading	3
No consideration of students with special needs	2
Grading practices are sufficient	2
The differences in grading practices can be good to allow for individual and classroom needs	1

Questionnaire Question 2. What pieces of evidence (homework, tests, district assessments, NeSA, Terra Nova, work habits/study skills) do you most often use to communicate student achievement? The strongest rated piece of student achievement evidence indicated by the teacher participants was district created assessments. These assessments were identified as both formative and summative assessments. They were explained as a true reflection of student achievement.

Homework was also identified as evidence frequently used to communicate student performance data. It was mentioned by six of the teachers, making it the second highest piece of evidence used to communicate student achievement.

Homework was not perceived by all to be a good reflection of student performance, which was summarized in the below statement.

The evidence I use most to communicate true student achievement is work done in class. Homework is too often done too much by parents...I have really liked the results and information that District adopted assessments have given me...easy to communicate to parents and shows differences and growth.

The majority of the teachers specified they did not categorize standardized assessments as a strong piece of evidence to communicate student achievement. Only one teacher felt it was a good representation of student ability. The respondents described standardized assessments as evidence of their teaching and instruction.

Unit tests were the third most indicated evidence of student performance, with five teachers in favor of using unit tests for assessment of learning. One teacher summarized her thoughts on achievement evidence by stating,

I use homework as a small piece of evidence...tests at the end of a unit really assess student learning and application of skills. Standardized test scores show

where students are at that day, but more importantly help guide my instruction so I know where to go next with each student.

Study skills and other factors, according to the research, should be separated from student performance evidence. It is thought when student behavior and effort are included in grading, communication of student achievement is distorted (Reeves, 2011; Guskey, 2015). Three of the eight teachers listed study skills as communicated evidence of student achievement. One response clarified the separation of study skills from academic achievement. "Work habits and study skills get their own grade, but it is solely based on teacher observation." Where another expressed the importance of study skills in their communication of student achievement, "I focus most on work habits/study skills."

Table 7. Evidence most often used to communicate student achievement.

Pieces of Evidence	Quality of Evidence	Number of Teacher Responses
District assessments	Strong piece of evidence	7
	Weak Piece of evidence	0
Homework	Strong piece of evidence	6
	Weak Piece of evidence	1
Unit tests	Strong piece of evidence	5
	Weak Piece of evidence	0
Work completed in class	Strong piece of evidence	3
	Weak Piece of evidence	0
Study skills	Strong piece of evidence	3
	Weak Piece of evidence	1
Standardized tests	Strong piece of evidence	1
	Weak Piece of evidence	6

Questionnaire Question 3. What are the differences in practice between the content areas of reading, mathematics, and writing when assessing and communicating student achievement? The differences in content area assessment and grading was expressed in three forms: the easiest to assess, the most difficult to assess, and the area in which was most accurately assessed. The research participants were then asked to elaborate on their thoughts.

Overpoweringly, mathematics was recorded as the easiest and most accurately assessed content area. Seven of the eight teachers believe that mathematics is the easiest content area to assess and six of those teachers further stated that math is the most accurate content area assessed. A teacher supported this by stating, ““Math! Student responses are purely objective.” Another teacher wrote, “Math seems to be the easiest content area to assess because it is pretty black and white. Either they were able to grasp the concept or they weren’t.”

The most difficult content area to assess was identified as writing; five teachers expressed difficulty. Reading was a close second with three teachers reporting difficulty. A teacher comment that concurred this was,

Writing is the most difficult for me to assess because it is more subjective than reading or math. It is also more difficult to grade as frequently as math or reading because producing an end product in writing takes longer.

Another teacher explained their position on why reading was difficult to grade.

I struggle giving reading grades because the skills change each week, so a student’s grades might fluctuate weekly. The skills area applied to many different situations and contexts, so I feel an overall grade is more difficult to give.

Table 8. Difference in assessment practice in content areas.

Differences in Practice	Content Areas	Number of Teacher Responses
Easiest content area to assess student achievement		
	Math	7
	Reading	1
	Writing	0
Most difficult content area to assess student achievement		
	Writing	5
	Reading	3
	Math	0
Most accurate content area to assess student achievement		
	Math	6
	Reading	2
	Writing	0

Questionnaire Question 4. What feedback do you receive about your grading practices and communication of student achievement? Overall, when the teacher study participants were asked about the feedback they receive in regards to grading practices and communication of student achievement, they felt the feedback was very limited. The feedback was focused on a perceived unhappiness with a grade a student received. Only one teacher recorded a response that they had received feedback for growth through a formal appraisal system.

Described feedback from parents was grade focused. One questionnaire respondent said, "I don't receive a lot of feedback from parents, but when I do, it's because they are all too concerned about the grade!" Half of the teacher survey participants associated the lack of feedback from administration and parents due to the understanding of their grading procedures. This conclusion was explained,

I do not usually receive any feedback. I begin conferences explaining to parents what the grade consists of. I feel like if you educate the parents and communicate with them, they trust your judgement.

Even though half of the teachers said a lack of feedback was due to statements such as, "I can explain my grading practices to that parent." An expressed desire for feedback on grading and communication of student progress was recorded in the teacher questionnaire. "I don't receive much feedback about my grading practices, but I think it would be very beneficial to me and my students."

Table 9. Received feedback on grading practice and communication of student achievement.

Feedback on Grading and Communication of Student Achievement	Number of Teacher Responses
I do not receive feedback on grading	4
I receive grading feedback from parents when students receive low report card marks	4
Administration, parents, and teachers understand grading procedures so feedback is not necessary	4
I receive grading feedback during informal conversations with colleagues	1
I receive grading feedback during my end of the year summative review from administration	1
I receive grading feedback throughout the school year from administration	0
I receive grading feedback from parents when students receive high report card marks	0

Questionnaire Question 5. What barriers (time, school policy, understanding, personal beliefs, grading system, pieces of evidence, etc.) do you experience in

providing consistent and clear communication of student achievement? When questioned in regards to the barriers that impact communication of academic performance, seven answers generated support from at least 50% of the teacher study participants. The teacher sentiments on communication consistency is revealed in this statement,

Time is a huge barrier, there is never enough of it. Also school policy dictates when grade communication goes home and what that communication is. There isn't even always a policy in place for certain assessments as to what goes home and when.

A second respondent shared feelings of barriers in this recorded sentiment,

Because grades are used as evaluation of student work, it's important that grades accurately reflect the quality of student work and that student work is graded fairly. Grading with accuracy and fairness can take a lot of time, which is often short.

Two out of the eight noted barriers were internal factors controlled by teachers.

Those being individual beliefs and fairness in grading. The remainder of the barriers were external factors impacted by the School District practices and procedures.

Table 10. Barriers for consistent and clear communication of student achievement.

Barriers to Consistent and Clear Communication of Student Behavior	Number of Teacher Responses
The time it takes to grade and understand student achievement	5
No district policies or common practices in grading	5
No training or professional development provided on grading	5
Focus is on a grade not learning	4
Grading does not reflect student growth	4
Fairness in grading	4
Personal beliefs	4
A grade is only one piece of student achievement	3

Questionnaire Question 6. What added support would assist you in achieving consistent grading practices and clear communication of student achievement? There was a premier questionnaire response that was believed to support clear communication in grading practices. Seventy-five percent of the replies mentioned a need for a clearly defined philosophy on grading. These replies included, “A clear outlined plan of what to include for final grades, when and what to send home. So the entire district would follow the same philosophy/policies.” and “Consistency across grade levels and the District...there is a need to use in-service days to discuss and implement same practices.”

A contradiction to the top requested grading and assessment supports: defined philosophy, process, and practice, was the request for greater autonomy to make grading decisions that support specific content areas, distinct student needs, and different classrooms. Support for autonomy was expressed, “I appreciate autonomy to grade and assess in a way that reflects my teaching style and student needs.”

Overall, the questionnaire revealed a necessity for guidance and understanding of grading practices. The teacher reactions exposed a belief that there was a need for a high level of support and guidance system wide. Over half of the suggested support concentrated on actions that would originate from school leadership.

Table 11. Suggested support for consistent grading practices and clear communication of student achievement.

Support for Consistent Grading and Clear Communication of Student Achievement	Number of Teacher Responses
A clearly defined philosophy on grading	6
Professional development on best practices in grading	4
More time to assess students individually	4
A clearly defined grading process	3
Greater autonomy to make grading decisions that support specific content areas, distinct student needs, and different classrooms	2
Ask stakeholders for feedback on current grading practice	2
Communicate common grading practices to parents	1

Summary

Chapter Four presented the statistical results of the data collected based on student performance on their third quarter report card and the Nebraska State Accountability Assessment. The chapter also presented grading perceptions of the fourth grade teachers' that were assigned the student subjects during the 2014-2015 school year. In Chapter Five, the researcher will complete an analysis of the data and offer a conclusion. The chapter contains a discussion on how the quantitative data aligned with the qualitative collection of teacher perception on grading practices. Finally, the conclusion will compare Teacher Cognition Theory, Thin-Slicing Theory, and related research to the study results.

CHAPTER FIVE

COUNCLUSION AND DISCUSSION

Purpose of the Study

The purpose of this study was to explore evidence in order to gain knowledge in regards to how well traditional grading practices and the use of standardized assessments communicate a clear and consistent picture of student achievement. The two variables, student report card grades and student performance on Nebraska State Accountability Standardized Assessments, were compared and statistically correlated to identify the strength of the relationship. The correlations were compared to teacher grading perception and beliefs to determine the impact they may have on grading practices.

The student study participants were sorted by the individual teacher they were assigned to during their fourth grade academic year. The report card data in reading, mathematics, and language arts, along with NeSA reading, mathematics, and writing data was available for 137 students assigned to eight classroom teachers. All student and teacher data was anonymously coded in the study documents to protect the identity of the participants. All eight teacher participants completed a questionnaire in regards to their perception and beliefs on current grading practices. Teacher survey responses were linked to performance correlations for students who were assigned to the corresponding teacher. Common themes were collected from the teacher questionnaire responses and matched to student performance data.

By recognizing the connection between student performance and teacher grading perception this study attempted to respond to the broad question: Do traditional grading practices and standardized assessments accurately measure and communicate student academic performance?

Conclusion

The following conclusions were drawn from this mixed method exploratory study.

Study Results

This study examined the relationship between fourth grade students' third quarter percentage grades and the students' scale scores on NeSA for specific content areas. A Pearson Product Momentum Correlation, with a two-tailed .05 Alpha level, was employed.

Research Question 1. Research Question 1 was used to examine the relationship between teacher assigned student report card grades in reading and scores on the Nebraska State Accountability Reading Assessment for students enrolled in fourth grade. The results showed a significant linear relationship between student assigned percentage grades in reading and the students' achieved scale score on the NeSA-Reading. The correlation coefficient reached .747, which was the highest content area correlation of the study.

This strong correlation was reached even though the majority of the teacher study participants voiced that reading was not the easiest or the most accurate content area to assess. In fact, only one teacher thought that reading was the easiest content

area to assess and only two of the teachers thought it was the most accurate. This respondent voiced differently than the other professionals, "Reading is the easiest to assess...the curriculum provides great resources to use to assess knowledge."

Colleagues in this study overwhelmingly articulated the subjectivity they employ when grading reading. This response demonstrated their thoughts, "Reading and writing are the most difficult. Most responses in reading are open ended and grading is based on teacher discretion, which is not consistent."

The expressed subjectivity when grading reading, created doubt in the teachers' ability to accurately assess student reading achievement. Educators shared a comfort level when they assumed there to be a right or wrong answer. As one respondent stated, "There are multiple resources that are used to accurately measure student knowledge in reading and many require the uses of teacher judgment."

The elementary teacher has been perceived to be an expert in teaching young students to read. Elementary teachers have been provided in-depth professional development in the use of reading resources and strategies. The emphasis of importance in reading, within the elementary classroom, was demonstrated by the State of Nebraska's recommended time allotment for weekly language arts instruction. The State recommendation for language arts instruction was identified as 35% of the weekly classroom time allotment, where compared to math instruction at 15% of the weekly classroom allotment (Nebraska Department of Education, 2011). Elementary classrooms frequently dedicate two blocks of time daily to reading instruction. This leads to the conclusion that the emphasis placed on developing literacy skills in the

elementary classroom and the extensive exposure to reading resources, strengthened the correlation between the two reading achievement evidence pieces.

Research Question 2. Research Question 2 was used to examine the relationship between teacher assigned student report card grades in math and scores on the Nebraska State Accountability Math Assessment for students enrolled in fourth grade. The comparison of student assigned percentage grades in mathematics and the students' achieved scale score on the NeSA-Math showed a significant relationship of .664.

The linear association was lower in math by .083 when compared to the relationship of performance data for students in reading. These results contradicted the majority of the teacher responses in regards to the subject area they most accurately assessed. Overwhelmingly, six of the eight questionnaire responses indicated math as the most accurately assessed content area.

Elementary teachers are responsible to teach all subject areas. They are not expected to specialize in specific content areas and rarely are required to take higher level math courses during teacher preparation programming. Elementary teachers lack the necessary mathematical understanding to help students comprehend concepts well. Their own understanding of math facilitates a need to assess on right or wrong answers.

Students who are taught mathematics and assessed merely on the memorization of rules, are less likely to apply what they have learned and are less likely to develop deep understanding of mathematical concepts (Lubienski, 2007). This thought can also be applied to a teacher's ability to assess mathematics. When a teacher has merely

memorized math rules, their lack of deep understanding leads to a need for a clear answer. Questionnaire statements such as, "Math is my strongest subject to assess...it is pretty cut and dry to see if students are able to understand the content." and "I am most accurate in math due to clear cut answers.", support the idea that the strength of the correlation and participant perception on grading was impacted by a shallow understanding of mathematical concepts by the fourth grade teachers.

Research Question 3. Research Question 3 was used to examine the relationship between teacher assigned student report card grades in language arts and scores on the Nebraska State Accountability Writing Assessment for students enrolled in fourth grade. The correlation in this subject area was the lowest association of all content areas investigated. The relationship was still significant at a level of .384.

It is important to note that all statistical correlation analysis revealed a linear relationship between the traditional grading system and performance on the standardized assessment. In the questionnaire, responses aligned with the statistical results in the content area of writing. Five of the teachers specified writing as the most difficult content area to assess and none recorded assessing writing accurately. One teacher conveyed, "Writing! It is hard to teach. We can use many different tools to help, but ultimately it is up to the student to use them. By fourth grade their writing style is set."

It was evident in the questionnaire that the teachers' self-efficacy, belief in their own ability and capacity to successfully assess student learning in writing, was low. Time, complexity, and difficulty in writing assessment was expressed in the following

teacher statement, "Writing is the most difficult for me to assess. There is no clear cut way to know how students have grown. Grading is subjective and very time consuming." A lack of belief in teacher ability to assess writing impacted the strength of the correlation comparative to the investigated content areas.

Research Question 4. Research Question 4 was used to record the perception of the teacher participants on the clarity and consistency of communicated student achievement. Since none of the eight fourth grade teachers achieved the highest correlation in all subject areas, it was important to investigate how their grading perceptions impacted their ability to clearly and consistently communicate student achievement. It is essential when reviewing the study results to understand that all eight teachers recorded a significant correlation in all subject areas, except one teacher in language arts, when comparing communicated report card performance and NeSA Assessments. Where the study discovered variance, was in the strength of correlation among the fourth grade teachers and between the content areas.

The questionnaire responses from the teacher study participants revealed a strong opinion that current grading practices contained inconsistency. As articulated in this questionnaire response, "Our grades are not consistently done across grade levels or buildings so it makes it hard to compare. Grades do not reflect the individual."

Seventy-five percent of the teacher participants felt grading and assessment practices within the study district were inconsistent. The inconsistencies documented by the questionnaire participants included: "What a student needs to do to earn a grade, what is included in a grade, implementation of assessments, and difference in

teaching to the assessments.” In addition, more than half admitted that grading is subjective, “Grading is subjective, some content areas are more subjective than others. The more subjectivity the more difficult it is to grade.” Only one teacher felt that individuality in grading was beneficial. Two of the teachers felt current grading practices were sufficient and half of them felt fairness in grading was a barrier to accuracy.

Further inconsistency was discovered in perceived quality evidence used to communicate student performance. The difference in evidence quality was demonstrated in the following responses, “I rely heavily on homework and study skills such as teamwork, collaboration, character, and moral. These traits are not measured in a typical standardized test or district assessment.” and “I have really liked the results and information that district adopted assessments have given me...easy to communicate to parents and shows differences and growth.”

All but one teacher felt the greatest piece of student performance evidence was assessments created at the local level. Student homework ranked second and curriculum unit tests ranked third as quality performance evidence. Standardized assessments were identified by only one teacher as a strong piece of evidence in communicating student achievement.

When referring back to the correlation statistics, reading scores reached the highest linear relationship among the three content areas in the study. Whereas, the teachers questionnaire identified math as the easiest content area to assess and the most accurate depiction of student performance. The content area of writing, statistically and perceptually, was identified as the least accurate portrayal of student

performance. Teachers' perception of their ability to easily and consistently assess student learning did not match the correlation data collected on student performance.

The barriers to consistent student performance communication identified by the teacher participants included a lack of feedback from school administration. One teacher reported receiving feedback for growth. Whereas, other participants only received parental feedback when students demonstrated low performance. Half of the teachers recorded never receiving any feedback on their grading and assessment practices. The lack of feedback was contributed to stakeholders possessing an understanding of grading procedures. Further barriers that were strongly communicated through the teacher questionnaire were: personal beliefs impacted grading, grading procedures were not fair, no district policies or training allowed for individualized grading decisions, and not enough time to properly assess student learning.

A deeper analysis of the individual teachers' grading perceptions aligned with the correlational relationships of student performance data revealed the following results. Two of the teachers did not express a concern for inconsistent grading practices. In fact, they conveyed grading practices were sufficient. The expression of grading sufficiency lead to a bottom ranking in strength of correlation in all content areas for one teacher. The second response of sufficiency linked to inconsistent rates of correlation, ranking at the top among the teachers for writing and below the median in the area of math.

Strong opinions that district grading practices were inconsistent, insufficient, subjective, and did not reflect individual student achievement were transcribed by the majority of the respondents. Even though the majority communicated inadequate feelings in regards to district grading practices, their grading practices correlated the highest with student NeSA performance, specifically in the subject areas of math and reading.

Within the questionnaire teachers discriminated which subject areas were easiest and most accurately assessed in the classroom. All eight teachers except one, recorded math as the easiest content area to assess. Six out of the eight teachers also indicated that math is the most accurately assessed. More than half of the teachers expressed concern for the difficulty in accessing student writing achievement.

Teacher perception of grading by content areas exposed a lack of understanding by the teacher study participants in their own abilities to assess student performance by content area. Every teacher scored the highest correlation rate in student performance on NeSA Reading when compared to reading report card percentage grades. This suggests the perception that math is the easiest and most accurate content area of assessed student achievement is faulty. One teacher did indicate that reading was the easiest and most accurate graded content area. It should be noted that this teacher achieved the highest correlation rate in the content areas of math and reading. The reading correlation rate was .877. Consequently, the perceptions that the teachers have of their current grading practices did not match the quantitative data that showed

teacher grading judgements correlate with student performance on the NeSA Standardized Assessment.

Teacher perception of accuracy was addressed in the above discussion of the study results. The association of perceived accuracy in communication of student performance was attributed to three factors. The first factor was the emphasis of reading assessment and instruction in the elementary classroom. This is documented in the State of Nebraska's suggested time allotment for language arts in an elementary classroom (Nebraska Department of Education, 2011). Secondly, the elementary teacher is not required to specialize in specific content areas. Teacher preparatory courses focus on pedagogy and not higher understanding of skill development. The lack of understanding concepts well, simplifies assessing student learning by looking for right and wrong answers. Lastly, teacher self-efficacy impacted teacher confidence in their ability to assess student learning. This was distinguished particularly in the content area of writing.

Discussion

Randall & Engelhard (2010) expressed a need for grades to simply reflect academic performance towards learning targets. Even though grading systems play an important role in education, they are often considered to have little relationship to student performance (Brookhart, 2004; O'Connor, 2016). This study supported that two highly used performance data sources, student report card percentage grades and student achievement on the NeSA Standardized Assessment, communicated congruent student performance results. The research presented here revealed a significant

relationship between a traditional grading system and student achievement on the NeSA Assessment. The study supports that the communication of performance evidence shared with stakeholders in the research District is consistent.

A concern recognized in this study was that teachers did not always perceive their grading and assessment practices to be accurate and reliable. The number one purpose of grading and assessment identified by Frisbie and Waltman (1992) is to communicate the achievement status of students accurately to parents and other stakeholders. The study supports the research of Guskey (2007) when he found that diverse stakeholders perceived legitimacy of achievement indicators differently. According to Guskey, administrators viewed state, district, and national standardized assessments as a trustworthy source of academic achievement, whereas teachers perceived classroom observations and homework more trustworthy than administrators. Through the teacher questionnaire, study participants designated local created assessments and homework as the best pieces of student performance evidence, which are a component in the reporting of student achievement on the traditional grading scale. A questionnaire response summarized the evidence selection based on personal preference, "Personally I use homework, tests, and district assessments. Work habits and study skills get their own grade, but it is solely based on teacher observation." Also, communicated on the questionnaire was limited support for the use of standardized assessment as student achievement evidence. This is depicted in the statement, "Standardized assessments I prefer to use only to look at student growth. It helps me as a teacher know what I need to improve on or spend more time

on in the classroom.” Both of these perceptions support the research findings of Guskey (2007).

According to Reeves (2011) and Guskey (2015) traditional grading systems are inconsistent and inaccurate in conveying how students perform in relation to learning standards. The NeSA Assessment measures student performance on the adopted Nebraska State Standards. The traditional grading system implemented in the research district significantly correlated to the student performance measurement of achieved State Standards in the content areas of reading, mathematics, and language arts. It is believed a more accurate measure of student performance is attained when measuring student achievement through learning standards Guskey (2001). Guskey, (2015) and O’Connor (2009) emphasized in their research a need for standardized grading due to the impact of external factors on grading outcomes such as: behavior, effort, and student background. According to the presented research in this study, a traditional grading scale aligns with the measurement of achievement through standards as measured through the NeSA Assessment. A majority of the teacher study participants believed autonomy in grading decision making created variance in grading practice. There was no evidence that the variance in practice impacted the relationship between the two communicated pieces of student performance data.

Alignment with Conceptual Frameworks

Factors such as teacher background and experience in evaluation have shown to affect grading practices, which can produce a lack of uniformity between student performance judgments (Schafer et al., 2001; Eckes, 2008). This conceptual theory is

referred to as Teacher Cognition Theory. The teacher questionnaire revealed half of the teachers confirmed that personal beliefs impact their grading practices, with such expressions as,

I am appreciative of the amount of freedom teachers are given in the grading and assessing practice they use in their classroom. It allows teachers to assess and grade in a way that better reflects their teaching style and the needs of their students.

The development of teacher beliefs are explained to be formulated from contextual factors, early experiences as a student, professional coursework, and classroom practice (Borg, 2003). The teacher questionnaire did not disclose a strong indication of which cognition component greatest impacted teacher grading and assessments practices. It did expose that individual teacher cognition impacted grading practice as expressed in the following response,

I was not given a clear direction and purpose and as a new teacher I used my own opinion to choose what goes into a grade. When I asked another teacher about grading, I was told we are like God when it comes to grading...I was terrified if what I was doing was the right way...I began to develop my own grading system I felt comfortable with.

Experience in the field changed teacher cognition. Respondents shared how grading and assessment practices have altered with experience in the classroom. A response expressed how pieces of evidence changed in different academic years, "Currently my grades consist of homework, all district assessments, tests, quizzes, and projects...other years I did not include district assessments."

The below response linked teacher early schooling to their grading practices today. What they encountered to be good components of writing was viewed to limit their ability to grade student writing today.

I would say writing is most difficult to assess...my bias of how I learned to write comes into play or we are distracted by other issues such as spelling, grammar and conventions. I don't really see the students' overall writing.

Another response revealed how a teacher modified District practices in order to align with their individual beliefs on student understanding which were derived from cognition factors.

What I have learned is that a good rule of thumb is that when students are performing at 80% or higher, they are showing good mastery of the skill and are ready to learn the next skill. However, our District would label that a very low grade."

Teacher Cognition Theory was reinforced by the teacher study questionnaire responses. Within the responses, teachers revealed their experience, knowledge, influence, and belief impacted grading and assessment practices. Even though the recorded responses supported the Teacher Cognition Theory, this study revealed that individual Teacher Cognition Theory did not significantly impact the communication of student performance through a traditional grading scale when correlated to student performance on a standardized assessment.

The lowest level of correlation was in the area of writing, where the research teachers expressed difficulty in assessed accuracy of student performance, "Writing is the most difficult content area to assess student achievement because writing is subjective...I have a hard time grading student on their own experiences and writing skills." The amount of subjectivity in grading writing conveyed by the teachers alludes to the uses of cognition to formulate decisions in grading. When there is greater subjectivity in practice, it allows for Teacher Cognition Theory to influence enactment of grading practice.

Popham (1999) stated standardized achievement assessments are the wrong measurement tool to portray an accurate interpretation of student performance. This study demonstrated that the communication of student achievement through a standardized assessment aligned with student measures of daily performance in the classroom. The research presented here did not determine whether standardized assessment are the right piece of performance data, rather it confirmed it aligned with teacher communicated student performance on a traditional grading scale.

The Theory of Thin Slicing (Gladwell, 2005) states that conclusions can be formulated quickly with limited information. A standardized assessment can be classified as a thin slice, a quick assessment that yields a student performance outcome based on limited information. The teachers explained, "Standardized assessments seem to be a snap shot." and "Standardized test scores only show where students are at that day." Both comments denounce standardized assessments accurately communicate student achievement and a thin slice can communicate accurate student performance measures.

The thought that greater amounts of evidence creates stronger accuracy in communication of student performance was supported by the teachers. Within the questionnaire they stated accuracy was promoted when multiple pieces and a variety of evidence was available when assessing student achievement. This thought was shared, "I am most accurate in assessing student achievement in math because it is more objective and I assess students most often in math." A greater amount of opportunity to assess and collect evidence did not equate to greater accuracy.

According to the results of this study, the thin slice of student performance data, NeSA Assessment, significantly correlated with the ongoing performance calculation of student achievement when using the traditional grading scale. This negates the teacher perception that the Theory of Thin Slice or the NeSA Assessment is a poor piece of evidence to communicate student achievement.

Implications for Student Achievement Communication

The study showed variance in the correlation level amongst the content areas and between the teachers of the same grade level. Even though all individual teachers achieved significant correlation in communication of student performance, the degrees of significance varied. Improved and consistent student performance communication could be reached through established philosophies, policies, and practices for the research district. The teacher participants repeatedly responded similarly, requesting the development of shared grading philosophy, policy, and practice.

I would benefit from having research-based clear expectations from the district level as to what should be included to maintain a level of consistency across the District...My building fourth grade team and I all have different grading policies about what we include in grades. We know that more consistency and guidance would be best.

According to supporting research, the difference in teacher evaluation of student performance was eliminated when teachers used an established grading criteria in assessing student performance and received training on the criteria (Kan & Bulut, 2014; Schafer et al., 2001).

The indication that the establishment of common grading practices were nonexistent is further evidence of a need. This was apparent in a recorded teacher account,

I am completely unaware of how other 4th grade teachers give grades to students. There has not been any dialogue to share ideas and practices in grading, so perhaps this would be a good place to begin working towards common grading practices.

Teachers believed that the formation of grading criteria would improve current communication of student achievement. The teachers felt that further training in this area is a high need for improved student performance communication.

The perceived presence of grading inaccuracies was higher than the actual occurrence of incongruent student achievement communication. Teachers received limited or no feedback on their grading practices. This led to uninformed perceptions of their grading and a lack of confidence or self-efficiency in their ability to assess student performance by a majority of the study teachers.

This may be due to a lack of understanding of best practice in grading by school leaders. Research shows few school leaders have extensive knowledge of the attributes of various grading methods and the impact different grading policies have on students (Brookhart, 2011, Brookhart & Nitko, 2008; Stiggins, 1993; Stiggins & Chappuis, 2011).

School leaders should be trained to observe, identify, and communicate research based grading practices. The establishment of a shared grading philosophy and practice would assist the classroom teachers in congruence of grade assignment. One teacher quantified the establish of a grading philosophy like this,

The support that would be helpful in assessing students is having a clearly defined process on how we come up with grades. Do we use summative or formative assessments? Or both at different percentages? A policy that would be consistent across the District. A policy needs to be in place. Principals and teachers receive professional development on what this policy looks like and how to implement it in the classroom. It would also be beneficial to communicate with parents on the policy.

In addition, it would benefit administration in providing feedback for improved communication of student performance. A formal system of grading practice feedback may be established through the teacher appraisal process.

There is a current trend in education to transform grading systems to a standards based measure. Use of a standards based reporting system is thought to increase the reliability of grading practices and provide a clearer picture of student learning (Hardegree, 2012; Guskey, 2015; & O'Connor, 2009).

The outcome of this study indicated a strong understanding of student achievement can be communicated with the use of a traditional grading system. It also suggested that a traditional grading system can align with an assessment that measures student achievement towards adopted learning standards; even when student performance is measured by a standardized assessment defined as a thin slice of evidence. The need to adopt a standards based grading system not supported and the emphasis of grading communication improvement focused on defining grading philosophy, practice, and procedure through ongoing professional development and appraisal.

There was evidence that the classroom teachers were able to clearly and accurately communicate student performance through a traditional grading system,

even though their perception of grading practices was inconsistent and the influence of cognition was recorded. The assignment of traditional grades correlated with student performance on standardized assessments. Even though the correlation rate of these two pieces of student performance evidence was significant, there was room for consistency improvement amid the teachers and content areas. Fairness and uniformity in grading is obtained when all students are provided equitable learning opportunities (O'Connor, 2016). School districts should consider committing to the development of defined beliefs and practices to improve differences and provide clear, consistent communication of student performance.

References

- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know believe, and do. *Language Teaching*, 36(2), 81-109.
- Borg, S. (2009). Introducing language teacher cognition. Retrieved from <http://www.education.leeds.ac.uk/research/files/145.pdf>
- Borko, H., Cone, R., Russo, N. A., & Shavelson, R. J. (1997). Teachers' decisions making. *Research on Teaching: Concepts, Findings, and Implications*. Berkeley, CA: McCutchan.
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, 30, 123-142.
- Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education*, 7(4), 279.
- Brookhart, S. M. (2004). *Grading*. Upper Saddle River, N.J.: Merrill Prentice Hall.
- Brookhart, S.M. (2011). Starting the conversation about grading. *Educational Leadership*, 69(3), 10-14.
- Brookhart, S. M., & Nitko. A. J. (2008). *Assessment and Grading in Classrooms*. Upper Saddle River, NJ: Pearson Education.
- Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. *M. C. Wittrock Handbook of research on teaching* (3rd ed.). New York, NE: Macmillan.
- Coladarci, T. (1986). Accuracy of teacher judgments of student responses to standardized test items. *Journal of Educational Psychology*, 78(2), 141-146.
- Creswell, J. C. (2012). *Educational research: Planning, conducting, and evaluating*

quantitative and qualitative research (4th ed.). Boston, MA: Pearson.

Cross, L. H. & Frary, R. B. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied Measurement in Education*, 12(1), 53-72.

D'Agostino, J.V., Welsh, M.E., & Corson, N.M. (2007). Instructional sensitivity of a state's standards-based assessment. *Educational Assessment*, 12(1), 1-22.

Data Recognition Corporation. (2015). *Technical Report Nebraska State Accountability Reading, Math, and Science*. Retrieved from http://www.education.ne.gov/Assessment/pdfs/Final_2015_NeSA_Technical_Report.pdf

Data Recognition Corporation. (2015). *Technical Report Nebraska State Accountability Writing Test*. Retrieved from http://www.education.ne.gov/assessment/pdfs/Final_2015_NeSA_W_Technical_Report.pdf

Eckes, T. (2008). Rater types in writing performance assessments: A classification approach to rater variability. *Language Testing*, 25, 155-185.

Edwards, N. & Richey, H. G. (1947). *The school in the American social order; the dynamics of American education*. Boston: Houghton Mifflin Co.

Every Student Succeeds Act, Pub. L. No. 1147-95, § 129, Stat. 2192 (2015).

Frederiksen, J. R. & White, B. Y. (2004). Designing assessments for instruction and accountability: An application of validity theory to assessing scientific inquiry. *Yearbook of the National Society for the Study of Education*, 103, 74–104.

Frisbie, D. & Waltman, K. (1992). Developing a personal grading plan. *Educational Measurement: Issues and Practice*, 11(3), 35-42.

- Gladwell, M. (2005). *Blink: The power of thinking without thinking*. New York: Little, Brown and Co.
- Great Schools. (2015). *No child left behind requirements for schools*. Retrieved from <http://www.greatschools.org/definitions/nclb/nclb.html>
- Guskey, T. R. (1994). Making the grade: What benefits students? *Educational Leadership*, 52(2), 14.
- Guskey, T. R. (2001). Helping students make the grade. *Educational Leadership*, 53(1), 20-27.
- Guskey, T. R. (2002). Perspectives on grading and reporting: Differences among teachers, students, and parents. Paper presented at the annual meeting of the American Education Research Association. (Eric Document Reproduction Service No. ED 490411)
- Guskey, T. R. (2007). Multiple sources of evidence: An analysis of stakeholders' perceptions of various indicators of student learning. *Educational Measurement: Issues & Practice*, 26(1), 19-27.
- Guskey, T. R. (2013). The case against percentage grades. *Educational Leadership*, 71(1), 68-72.
- Guskey, T. R. (2015). Your grading check-up: Practices to adopt, refine or drop now [Webinar]. *Educational Research Newsletter & Webinars*. Retrieved from <http://www.ernweb.com/shop/standards-based-grading-struggling-exceptional-learners/>
- Guskey, T., & Jung, L. (2012). *Grading and Reporting for All Students*. New York, NY:

Knowledge Delivery Systems.

Gutek, G. L. (1986). *Education in the United States: An historical perspective*. Englewood Cliffs, N.J.: Prentice-Hall.

Hardegree, A. C. (2012). *Standards-based assessment and high stakes testing: Accuracy of standards based grading* (Doctoral dissertation, School of Education Liberty University).

Hill, H.C. (2001). Policy is not enough: Language and the interpretation of state standards. *American Educational Research Journal*, 38, 298-318.

Hills, J. R. (1991). Apathy concerning grading and testing. *Phi Delta Kappan*, 72(7), 540-545.

Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgements of academic achievement: A Review of literature. *Review of Educational Research*. 59(3), 297-313.

Kan, A., & Bulut, O. (2014). Crossed random-effect modeling: Examining the effects of teacher experience and rubric use in performance assessments. *Eurasian Journal of Educational Research*, 57, 1-28.

Kirschenbaum, H., Simon, S. B., & Napier, R. (1971). *Wad-ja-get?: The grading game in american education*. New York: Hart Pub. Co.

Kohn, A. (1999). Grading is degrading. *Education Digest*, 65(1), 59.

Kohn, A. (2000). *The Case Against Standardized Testing: Raising the scores, ruining the schools*. Portsmouth, NH: Heinemann.

Linn, R. L. (1983). Testing and instruction: Links and distinction. *Journal of Educational*

Measurement, 20(2), 179-189.

Lubienski, S. T. (2007). What we can do about achievement disparities. *Educational Leadership*, 65(3), 54-59.

McMillan, J.H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research*, 95(4), 203.

Mehring, T., Parks, C., Walker, K., & Banikowski, A. (1991). Report cards: What do they mean during the elementary school years. *Reading Improvement*, 28(3), 162-68.

Meyer, L. (2000). Lingering Doubt Examiners: Results of Pilot Modeling Analyses. *AEB Research Report*.

Myford, C. M., & Mislevy, R. J. (1994). *Monitoring and Improving a Portfolio Assessment System*. Princeton, NJ: Educational Testing Service.

Nebraska Department of Education. (2011). *Planning the Weekly Schedule in the Elementary Classroom*. Retrieved from <https://www.education.ne.gov/wp-content/uploads/2017/07/Planning-Elementary-Schedule-August-2011-1pg-per-sheet.pdf>

Nebraska Department of Education. (2016). *Data Reporting System*. Retrieved from <http://drs.education.ne.gov/Pages/StateAssessment.aspx>

No Child Left Behind Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).

O'Connor, K. (2009). *How to Grade for Learning*. California: Corwin.

O'Connor, K. (2016). How to grade for learning by using 15 fixes for broken grades [PDF

document]. Retrieved from <http://www.gfps.k12.mt.us/sites/default/files/Ken%20Connor%20GFS%20Aug%202013,%202012.pdf>

- Ornstein, A. C. (1994). Grading practices and policies: An overview and some suggestions. *NASSP Bulletin*, 78(561), 55-64.
- Popham, W. J. (1999). Why standardized tests don't measure educational quality. *Educational Leadership*, 56(6), 8-15.
- Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26(7), 1372-1380.
- Reeves, D. B. (2008). Leading to change/Effective grading. *Educational Leadership*, 65(5), 85-87.
- Reeves, D. B. (2011). *Elements of grading: A guide to effective practice*. Bloomington, IN: Solution Tree Press.
- Resnick, L. B. (2006). Making accountability really count. *Educational Measurement: Issues and Practice*, 25, 33-37.
- Schafer, W., Swanson, G., Bene, N., & Newberry, G. (2001). Effect of teacher knowledge of rubrics on student achievement in four content areas. *Applied Measurement in Education*, 14(2), 151-170.
- Schmoker, M. (2000). Standards versus sentimentality: Reckoning-successfully-with the most promising movement in modern education. *NASSP Bulletin*, 84(620), 49-60.
- Scriffiny, P. (2008). Seven reasons for standards based grading. *Educational Leadership*, 66(2), 70-74.

- Stanley, G., & Baines, L. (2004). No more shopping for grades at B-mart: Re-establishing grades as indicators of academic performance. *The Clearing House*, 77, 101-104.
- Steidinger, R. (2011). *A study to determine the concurrent validity between course grades and standardized assessment results* (Doctoral dissertation, Department of Educational Leadership East Carolina University).
- Stiggins, R. J. (1993). Teacher training in assessment: Overcoming the neglect. *Teacher Training in Measurement and Assessment Skills*. Lincoln, NE: Buros Institute of Mental Measurements.
- Stiggins, R. J. & Chappuis, J. (2011). *An Introduction to Student-Involved Assessment for Learning*. Boston, MA: Pearson Assessment Training Institute.
- Stuhlmann, J., Daniel, C., Dellinger, A., Denny, R. K., & Powers, T. (1999). A generalizability study of the effects of training on teachers' abilities to rate children's writing using a rubric. *Journal of Reading Philosophy*, 20, 107-127.
- Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104(3), 743-762.
- Sweedler-Brown, C. O. (1985). The influence of training and experience on holistic essay evaluations. *The English Journal*, 74(5), 49-55.
- U.S. Department of Education. (2015). *Adequate Yearly Progress*. Retrieved from <http://www2.ed.gov/nclb/accountability/ayp/edpicks.jhtml>

Walker, T. (2015). With passage of every student succeeds act, life after NCLB begins.

NEA Today. Retrieved from <http://neatoday.org/2015/12/09/every-student-succeeds-act/>

Winger, T. (2005). Grading to communicate. *Educational Leadership*, 63(3), 61-65.

Appendix A

IRB # 805-16-EX

Hello!

As part of my dissertation research of teachers' perceptions and practices in communicating student achievement, I would appreciate about 20 minutes of your time to complete the following questions. All responses will be collected anonymously, with pseudonyms used if necessary to protect your confidentiality. Your participation is voluntary, and you can opt out of responding at any time. If you have any questions, you may contact me at 402-238-2372 or at sthoendel@bennps.org.

Thank you for your input. The results of this study may help our district and the broader field of education to clearly and consistently communicate student achievement, and also to support teachers in their efforts.

1. What is your perception of the assessment and grading practices currently implemented in your school district?
2. What pieces of evidence (homework, tests, district assessments, NeSA, Terra Nova, work habits/study skills) do you most often use to communicate student achievement?
3. What are the differences in practice between the content areas of reading, mathematics, and writing when assessing and communicating student achievement?
4. What feedback do you receive about your communication of student achievement?
5. What barriers do you experience in providing consistent and clear communication of student achievement?
6. What added support would assist you in achieving consistent and clear communication of student achievement?
7. What other information would help inform this research study on communication of student achievement?