

A Tale of Two Grades:

An Evaluation of *Grading for Learning*, A Middle School Grading Reform

A Dissertation

SUBMITTED TO THE FACULTY OF
THE UNIVERSITY OF MINNESOTA BY

Katie Pekel

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF EDUCATION

Advisor: Jean A. King

Date: November, 2013

© Katie Pekel, 2013

Acknowledgments

I would like to thank my advisor, Dr. Jean King, and my committee members, Dr. Lynn Searcy, Dr. Peter Demerath, and Dr. Sandra Christenson for their encouragements and insights not only on the writing of this dissertation, but on my pursuit as a scholarly educational leader.

I would also like to acknowledge the incredible students and staff of Ellis Middle School where I served as principal for the duration of the five-years that Grading for Learning was developed and implemented. Being a part of change is not always easy, but the staff at Ellis worked to insure that the components in this reform supported students and their learning.

Without the initial teacher inquiry, passion and leadership of Eric Harder, Curtis Bartlett, and Lynn Hemann, Grading for Learning would not have taken place at Ellis Middle School. Teacher and counselor leaders, Lisa Glynn, Jill Rollie, and Angie Taylor, along with the entire 8th grade team of teachers, greatly assisted in the development of Grading for Learning. The expanded leadership of The Grading for Learning Working Group, school counselors Lea Oelfke and Kirsten Norby, Grade Level Team Leaders and Department Chairs helped embed Grading for Learning into the culture of Ellis Middle School.

To my administrative colleagues, Ron Prybelski, Jessica Cabeen, and Brad Bergstrom a debt of gratitude for providing sounding boards for the ideas and feedback on the realities. My administrative assistant Kelly Pryor insured the products of the reform to parents, students and the broader K-12 community were professional, easy to follow, and always accurate and on-time. Our Director of Information Technology, Corey Haugen, went above and beyond in the technical support of implementing the grading reforms and statistically analyzing their impacts; it is because of Corey that we have sound correlation data. Without Corey's belief in what we wanted to do at Ellis and his support, we would not have been able make such significant changes. I also want to acknowledge the Austin Public School's district leadership and school board as they saw the value and fully supported this teacher initiated, site-led initiative for change.

I want to acknowledge faculty within the Department of Organizational Leadership and Policy Development who coordinated the program. Their vision for the development of scholarly leaders was one that allowed for my work as principal to be honored while also encouraging growth.

Finally, I wish to thank my dad, who helped to frame my understanding of the importance of education by asking, "Where are you going to college?" not "Are you going to college?" And for following that up twenty years later with encouraging calls asking, "How's your dissertation coming?" not "Are you going to get it done?"

Dedication

This paper is dedicated to my husband, Kent Pekel, not only for his support and guidance in the writing of it, but because he taught me the incredibly important lesson that *smart is not what you are; it is what you work to become*, a lesson that will remain with me far beyond the conclusion of our doctoral program.

I also dedicate this paper to the four other most important people in my life, our children. Adam, Thomas, Victoria, and Molly, thank you for your witty humor, calm patience, heartfelt support, and amazing questions as Kent and I were sometimes distracted from the things that interested you at 12, 10 and 9 years old as we both completed our dissertations. I hope by going through this with us, you will always be gently reminded that learning is a journey, not a destination.

Abstract

This developmental evaluation used multiple methods to evaluate the implementation of a grading reform initiated by two math teachers in a large traditional middle school in greater Minnesota. The reform, titled “Grading for Learning,” was developed based on a review of scholarly research and a collaboration of an interdisciplinary group of teachers and the principal. The approach required assigning two grades: a knowledge grade, which was based primarily on student assessments, and a life skills grade, which was based on a rubric that assigned scores for effort, behavior and timeliness. The evaluation studied the perceptions of students and staff regarding the separation of the two grades, if they found the changes in practice useful and if the reform in grading promoted greater alignment between:

- a. Subjective teacher evaluation of student knowledge as measured by the knowledge grade and an objective measure of student knowledge as indicated by scores on a standardized achievement test
- b. Teacher evaluation of student knowledge as measured by the knowledge grade and teacher evaluation of student effort as measured by the life skills effort grade
- c. Teacher perception of student effort as measured by the life skills effort grade and student perception of effort as measured by a survey

The evaluation concluded that staff and students generally perceived the separation of a single grade into the knowledge grade and life skills grade positively. Students and staff reported that the changes in grading were useful for many reasons, although most notably

because they had a clear understanding of what students knew about a particular subject as reported by the knowledge grade. Finally, changes in practices of calculating the knowledge grade increased the correlation of the knowledge grade and state standardized test scores. There was also moderate correlation between knowledge grades and life skills grades.

Table of Contents

	<u>Page</u>
Abstract	iv
List of Tables	vii
List of Figures	viii
Chapter 1: Purpose and Background	1
Chapter 2: Literature Review	15
Chapter 3: Research Methods	27
Chapter 4: Findings	44
Chapter 5: Discussion	77
Bibliography.....	89
Appendix A: Grading for Learning Overview for Students and Parents.....	93
Appendix B: Analysis of Teacher and Student Perceptions and Effort.....	96
Appendix C: Student Survey	99
Appendix D: Protocol for Student Group Interviews.....	104
Appendix E: Staff Survey.....	105
Appendix F: Protocol for Staff Interviews and Questionnaires	110
Appendix G: Results for Student and Staff Surveys.....	112
Appendix H: Unused Template for Parent/Guardian Survey... ..	124
Appendix I: Unused Template for Parent/Guardian Interviews	129

List of Tables

Table 1: Comparison of Traditional Grading to Grading for Learning	8
Table 2: Measures and Indicators of Evaluation Study	28
Table 3: Measure 1 Sample Size.....	29
Table 4: Measure 2 Sample Size	31
Table 5: Measure 3 Survey Questions	32
Table 6: Measure 4 Survey Questions and Group Interview Questions	34
Table 7: Measure 5 Survey Questions and Group Interview Questions.....	37
Table 8: Measure 6 Survey Questions and Interview and Questionnaire Questions.....	39
Table 9: Measure 7 Survey Questions and Interview and Questionnaire Questions	42
Table 10: Measure and Methods Used to Evaluate Grading Practices Reform Initiative.	44
Table 11: Summary of Indicators of Statistical Correlations	72
Table 12: Summary of Survey of Student Perceptions	73
Table 13: Summary of Survey of Teacher Perceptions,.....	74

List of Figures

Figure 1: 2007-2008 Grade 8 Pre-Algebra Correlation Analysis Scatterplot	46
Figure 2: 2008-2009 Grade 8 Pre-Algebra Correlation Analysis Scatterplot	48
Figure 3: 2009-2010 Grade 8 Math Correlation Analysis Scatterplot	49
Figure 4: 2010-2011 Grade 8 Math Correlation Analysis Scatterplot	50
Figure 5: 2011-2012 Grade 8 Math Correlation Analysis Scatterplot	51
Figure 6: 2011-2012 Grade 7 Math Correlation Analysis Scatterplot	52
Figure 7: 2011-2012 Grade 6 Math Correlation Analysis Scatterplot	53
Figure 8: 2011-2012 Grade 8 Math Grade and Effort Grade Correlation.....	55
Figure 9: 2011-2012 Grade 7 Math Grade and Effort Grade Correlation	56
Figure 10: 2011-2012 Grade 6 Math Grade and Effort Grade Correlation Scatterplot....	57
Figure 11: Relationship Between Teacher and Student Perception of Effort.....	59

Chapter 1

Background and Purpose

Purpose

Much has changed in education, but grading largely remains the same. While major school reforms in recent decades have brought about changes in curriculum, staff development, supplemental services, and restructuring schools, the basic structure of reporting student progress has changed little in over 100 years. Grades as we know them originated in the educational system in 1780, when Yale University began using a four-point scale to provide feedback to students. Today, more than 230 years later, over 80% of schools use letter grades to evaluate student performance during the upper elementary and secondary years (Marzano, 2000). The practice of reporting student progress via grades is now deeply rooted within and beyond schools.

Grading matters because it is the primary way teachers evaluate students and communicate their evaluation of student performance to the students themselves, to parents and to other educational institutions, such as schools and colleges that the student will later attend. Following a study of district and school grading policies and practices, Austin and McCann (1992) concluded that the five primary purposes for grading include administrative purposes (such as course placement), feedback about achievement, guidance, instructional planning, and motivation. Austin and McCann note that while many educators assume that providing feedback about achievement is the most important of these objectives, reviews of school documents and surveys of school staff suggest that,

in fact, managing administrative processes is the primary function that grades perform in many schools.

Multiple scholars and practitioners have concluded that there are many problems with the way grades are given and interpreted. It is extremely difficult, to cite one example, for a single grade to accurately summarize both a student's knowledge of academic content and the academic behaviors he or she demonstrates inside and outside of class. Studies have shown that grading practices often vary widely between teachers, including those who teach in the same school and even the same grade level and subject area. For example, a series of studies have found that the degree to which teachers include non-academic factors such as attendance, effort, participation and behavior in a grade is highly variable (Austin & McCann, 1992; O'Connor, 2007; Brookhart, 1994; Guskey, 2009).

Studies have also shown that even when teachers have participated in training or professional development activities that emphasize particular research-based approaches to grading, in practice they often do not use those approaches in their classrooms. This often occurs because those teachers do not view the new approaches as fair to certain students and because they conflict with established classroom management practices (Guskey, 2009; Brookhart, 1994; Friedman & Frisbie, 1995). Many leading educational measurement experts have, for instance, urged educators to separate academic and non-academic factors in assigning grades (O'Connor, 2007; Allen & Lambating, 2001; Guskey, 2000), but the practice of combining those factors into a single grade remains widespread.

Origin and Focus of the Evaluation

The purpose of this study was to evaluate how an effort to reform grading practices in a middle school in Minnesota affected student and staff perceptions of grading and the degree of alignment between grades and standardized test scores and other variables. The grading reforms that were implemented generally reflect the recommendations of researchers who advocate separating the evaluation of students mastery of academic content knowledge from their mastery of desired behaviors and other non-academic outcomes.

The middle school where the study took place serves a large and rapidly changing population of diverse students in a small Minnesota city. At the time of the study, the school served a population of 973 students. Of those students, 49.7% were female, 50.2% were male, 51.7% received free or reduced lunch, 31.2% were students of color, 11.2% received special education services, and 8.5% were English language learners.

The reform of grading practices in the school took place over a five-year period, as teachers and administrators in the school gradually studied, developed and implemented the reforms. This evaluation focused on the fifth year of the grading reforms, which occurred during the 2011-2012 school year. Permission to conduct the evaluation was obtained from the Institutional Review Board of the University of Minnesota, the superintendent of the school district, and the participants.

The primary reform in grading practice that took place at the school and that was the focus of this evaluation was splitting the traditional single letter grade into two separate grades: a knowledge grade and a life skills grade. Through this fundamental

change in grading practice, students received a knowledge grade using the scale of A, B, C, D, and F. A life skills grade that used the scale of 4, 3, 2, 1 was given for effort, behavior, and timeliness.

In addition to separating grades into the two components of knowledge grades and life skills grades, a series of other supporting changes were made in grading practices at the school to improve the accuracy of grading and to increase student motivation and effort. Those changes required that all teachers in the school do the following:

- Assign a 10% weight to homework and 90% weight to assessments in determining a student's final grade in a course
- Discontinue awarding extra credit
- Allow students to re-take tests if all of their homework was complete prior to the test first being given
- Accept work that is turned in late without lowering the grade
- Give a student 50% of the credit for an assignment or an assessment for which he or she had earned fewer than 50% of the total possible points.

While teachers were still expected to record the number of points the student actually earned, students who scored below 50% were still given credit for 50% of the total points to ensure that it remained mathematically possible for them to pass the course even after failing a test or an assignment.

Background on Implementation of the Reforms

As noted above, implementation of reforms in grading practices happened organically over a period of five years at the school study site, as teachers, administrators and counselors dialoged with each other, examined data, read scholarly works, and responded to student and parent concerns. There was no multi-year strategic plan that guided the reforms from the outset. Rather, school leaders made decisions and responded to events as they occurred. A summary of the evolution of the reforms follows by year:

Year 1: 2007-2008 School Year

The grading reforms originated with the participation of three of the school's mathematics teachers in a math and science improvement initiative known as the University of Minnesota/Hormel Foundation Teacher Fellowship. During the fellowship, those three teachers decided to study grading practices to fulfill one of the requirements of their graduate program. They chose to focus on grading because they felt that grades they gave their students often were not accurate representations of the students' mastery of the academic content they had taught. The teachers also noted that the grades they assigned also often reflected their assessment of students' life skills such as organization and behavior. In other words, the math teachers felt that a student's grade was often not a true representation of what he or she had learned or was able to do. The teachers concluded that it was possible for a student to earn a letter grade of B in mathematics, and yet not score at the proficient level on the state standardized math assessment. To inform their project on grading reform, the teachers read O'Connor's *A Repair Kit for Grading 15 Fixes for Broken Grades* (2007).

Year 2: 2008-2009 School Year

After the three pioneering math teachers completed their project for the Teacher Fellowship program, two of them decided to implement a new approach to grading in the 8th grade math courses that they taught. The teachers decided that they would assign and evaluate homework, but would not include points earned on homework assignments in students' course grades. The teachers also decided that if students received a grade on the cumulative final exam that was higher than their course grade, the higher grade would be awarded. In addition, the teachers decided not offer extra credit in their classes.

While the teachers encountered some concerns about these changes among the parents of students in their classes immediately after they were announced, after talking with the teachers most parents felt that the changes enhanced their understanding of how well their children were learning mathematics.

When these reforms were implemented in the two math teachers' classrooms, other teachers in the school grew interested in the reasons for and impact of the changes. Two 8th grade interdisciplinary teams were particularly intrigued. At the beginning of the year, they did not necessarily agree with what the math teachers were doing, but as the year progressed and many of the teachers became frustrated with students' work ethic, lack of responsibility, and occasional academic dishonesty, they began talking as a grade level about the possibility of adopting the practices that the two math teachers had pioneered in the school.

With the principal's support and in preparation for a pilot project the following year, the 8th grade teams met throughout the summer and developed their approach to

grading. This approach later came to be known as ‘Grading for Learning,’ and was designed for use in math, science, social studies, language arts, physical education, and a speech and writing course during the following school year.

Year 3: 2009-2010 School Year

The Grading for Learning approach developed and implemented by the interdisciplinary teacher teams required assigning two grades: a knowledge grade, which was to be based only on student assessments, and a life skills grade, which was to be based on a rubric that assigned scores for homework, preparation for class, behavior, and class participation.

The major differences in Grading for Learning as the teams initially designed the program and traditional grading are shown on Table 1 on the following page.

Table 1: *Comparison of Traditional Grading to Grading for Learning*

Traditional Grading	Grading for Learning	
Single Grade Consists of:	Knowledge Grade Consists of:	Life Skills Grade consists of:
Total points or weighted grading may be used	Total points	Rubric score
Assessment scores	Assessments (tests, quizzes, projects, essays, speeches and labs)	Homework
Homework calculated at varying percentages		Preparation for class
Participation, attendance, and behavior may or may not be included in grades		Behavior
Extra-credit may be factored in		Participation
Non-academic assignments (permission slips, progress reports, Kleenex) may be included		
Late-work may or may not be penalized		

The students and parents who would participate in the new grading system were told about the change via a mailing in August and at back to school conferences. As had been the case when the two math teachers changed grading practices in their classrooms during Year 2 of the reforms, parents initially did not express significant opposition to or concern about the change. As the school’s fall conferences in November approached, some teachers began to anticipate a parent and student backlash because the new practices had seemed to have lowered the first quarter grades they would be sharing at the conferences. All of the 8th grade teachers spent staff development time going through

their life skills grades and formulating talking points from which all felt comfortable speaking.

As some of the teachers had predicted, after the fall conferences some parental push back did occur. Parents reported that they were concerned that because homework did not factor into the knowledge grade, it didn't count. In an effort to address these concerns, the school held an information session during which parents were presented with evidence (which is described below) that significant misalignment existed between students' grades in mathematics and their test scores in mathematics on state standardized assessments. This data persuaded many parents that there was a valid reason for the reforms and that they should be given time to be implemented and evaluated.

During Year 3 of the reforms, the 8th grade team leaders and the principal also began briefing the rest of the school's staff and the 9th grade team at the high school on the rationale behind and the substance of the changes. After that information was shared, the school began a formal decision making process to consider further expansion of Grading for Learning. Feedback was gathered from staff following a number of informational sessions and book study of O'Connor (2007). The elements of Grading for Learning were then agreed upon among the building leadership team in April of 2010 and shared with the full staff in May of that year. In preparation for implementing the approach throughout the school during the following school year, a Grading for Learning working group was formed to achieve three objectives:

1. Develop a clear and concise written guide for teachers, students, and parents

2. Develop a two-hour staff development session for the staff to be delivered in August
3. Develop a 15-minute presentation for parents that explained the overarching reasons for adopting for Grading for Learning

Year 4: 2010-2011 School Year

An optional staff development session that focused on Grading for Learning was held in August, 2010, and was attended by over 77% of the school's staff. Additional sessions were offered during back-to-school workshop week to prepare the entire school staff to implement Grading for Learning during the 2010-2011 school year. During that year, the skills and behaviors that were being evaluated through the life skills grade were also incorporated into a new college readiness course that all 6th grade students took.

As teachers implemented Grading for Learning during the first year of school-wide implementation, there were times of frustration as some teachers lamented the absence of grading practices that they previously used to encourage and enforce compliance with their expectations. School administrators and teacher leaders attempted to help teachers address those concerns through weekly staff development sessions and collaborative work sessions held during early release time. At those sessions, teachers spent a significant amount of time examining assignments and assessments with their colleagues and talking about the grades that they thought those examples of student work should be given using the criteria of Grading for Learning.

Students were also provided with extensive information on the grading program. A primary vehicle for providing students with this information was the daily advisory

sessions in which all students in the school were enrolled. In those sessions, students were taught about the criteria for assigning knowledge grades and were also introduced to the rubric that would be used to assign life skills grades. The handout used for this instruction can be found in Appendix A. The students were also taught how to use their online student portal accounts to monitor their grades once every two weeks during the advisory period.

Providing parents with information on the reforms was also a priority. Parents were given information on Grading for Learning at back-to-school conferences, which were held in August. They were also provided with hands-on instructions for and demonstrations of logging into the online Parent Portal to monitor their child's grades. There was a 93% parent attendance rate at these conferences.

Year 5: 2011-2012 School Year

The second year of the building-wide implementation of Grading for Learning was supported by a college readiness course then required at all grade levels. The instruction on the grading parameters that previously took place in the daily advisory sessions now occurred in this course. Progress monitoring of grades continued to occur in the advisory sessions where each student was required to check their online portal account a minimum of every two weeks.

During the second year of implementation, teachers were more comfortable with their own knowledge of the parameters and began to explore differentiated ways to support students. After school study sessions to prepare students for tests, a teacher

staffed after school homework room, and additional after school bussing were a few changes that took place to support students.

This evaluation tells the story of what happened during Year 5 of the grading reforms, which occurred during the 2011-2012 school year. Using multiple methods, it also evaluates the impact that Grading for Learning had on teacher and student perceptions of grades and on the degree of alignment between grades, standardized test scores and other measures.

Research Questions

This evaluation sought to ask and answer the following questions about the Grading for Learning reforms that were conducted at the school study site:

1. How did staff and students perceive separation of the traditional single grade into two grades?
2. Did staff and students find the changes in grading practices useful? If so, how? If not, why not?
3. Did the reform in grading and related practices promote closer alignment between:
 - a. Subjective teacher evaluation of student knowledge as measured by the knowledge grade and an objective measure of student knowledge as indicated by scores on a standardized achievement test?

- b. Teacher evaluation of student knowledge as measured by the knowledge grade and teacher evaluation of student effort as measured by the life skills effort grade?
- c. Teacher perception of student effort as measured by the life skills effort grade and student perception of effort as measured by a survey?

Significance

Reform of grading practices is a growing priority for schools across the nation. The lessons learned from this evaluation can inform those efforts. This evaluation will inform the efforts of other educators to improve grading practices by helping them understand how the views that teachers and students had about grading system changed through grading reforms in a large Minnesota middle school. The understanding of these perceptions can enhance the continuing implementation of Grading for Learning at the study site and beyond.

Limitations

The evaluation described in this paper has a number of limitations of which readers should be aware as they review its findings. Those limitations are:

The study sample consisted of a single school community. While the school in question was a relatively large middle school with an enrollment of approximately 1000 students, the process and outcomes of reforming grading at other schools might logically be very different from what occurred at the study site. It is therefore important to be very cautious about generalizing the findings from this study to other middle schools.

The researcher for this study was also the principal of the school at which the study was conducted. I was actively involved in the development and implementation of the Grading for Learning program. The fact that I also held positional authority in the school where the reform was evaluated may have influenced the results of the evaluation. That said, my position as principal also ensured that I had access to all of the data and personnel that could benefit conducting the evaluation.

The comparatively short timeline for this evaluation, which focused on a single year of the implementation of Grading for Learning, does not allow for the extensive longitudinal analysis that would help to generate more robust findings. For example, the students were not followed through their high school years to see if the new approach to grading has an impact on the decisions they made about post-secondary enrollment.

Finally, this evaluation did not include the perspective of parents. I initially sought to include that input, but response rates to a parent survey and invitations to participate in parent interviews were so low that the perception of parents was eliminated as a subject of the evaluation. Because parents are a major recipient of grades and because their reactions to grades influence both their children and their teachers, the parent component is a major missing piece of the puzzle this evaluation seeks to put together.

Chapter 2

Literature Review

A number of important studies have noted that the effectiveness of the grading enterprise depends upon mutual understanding of the signal being sent about performance. For example, Austin and McCann found that many educators, students and parents see “grading systems [as] ‘shorthand’ languages for communicating evaluative information about students” (1992, p. 4).

Similarly, in an article about symbolic validation and state-mandated high school testing, Airasian (1998) cites Lindblom and Cohen’s observation that, “the commonality sought in public symbols rests largely on the cultural experiences and shared understandings of the social group” (1997, p. 303). Applying that concept to grades, most Americans have the shared experience of receiving grades; therefore, they may believe they have a shared understanding of their meaning. However, as Airasian (1998) explains, “the reaction to [or understanding of] the symbol is dependent on this commonality, because the symbol abstracts only the important aspects of the referent, leaving the individual’s memory to fill in the details” (p. 303).

Perhaps most important to this evaluation, both scholars and practitioners have found that people don’t agree on the meanings of the symbols of grades. Educational experts themselves highlight this disagreement. O’Connor (2007) suggests that grades have two purposes: to communicate student achievement and to inform teachers for instructional planning. In contrast, Erickson (2010) argues that grades should simply and accurately represent what a student knows and is able to do. And in a study conducted by

Guskey (2009), 74% of the over 500 teacher respondents indicated that the purpose of grades for report cards is to communicate achievement to parents.

Teachers, students and parents do not all agree on the meaning, or purpose, of the symbols. A study by Baron (2000) found that teachers, students and parents generally agree that tests and papers are the most important out of the following eight identified functions of assigning grades: class participation, attendance, homework, improvement (from the previous year or semester's performance), tests (and quizzes), papers, effort, and growth. However, Baron also found significant difference among the groups when they identified the impact they believed attendance, effort, and class participation have in determining a grade (Baron, 2000).

Researchers have also shown that elementary and secondary teachers have significantly different perceptions of grading. For example, 90% of elementary teachers report that the primary purpose of grading is parent communication, while 27% of secondary teachers report the primary purpose is evaluation of programs (Guskey, 2009). Austin and McCann (1992) analyzed documents of school boards, districts, and schools and found that the primary purpose of grades was to provide information about student progress.

Internal and external evaluators and audiences, such as K-12 and higher education also do not agree. In a study of the Oregon Proficiency Based Admission Standards System, a project to establish admission standards for the Oregon University system, data regarding grades and proficiency scores was gathered from 78 high school teachers. Assignments scored by the high school classroom teacher were compared to numerical

proficiency ratings on the same assignments scored by independent, trained scorers. A score of 3 was likened to the standards needed to do successful college entry work. In the results, many students who scored below a 3 were getting A's. The lack of relationship between the two – grades and the proficiency-based score system – may be attributed to the fact that the proficiency system captures student performance only while grades “capture a more varied aspect of the classroom experience” (Conley, 2000, p. 20). Further reason for this found discrepancy is that no two teachers used the same grading system (Conley, 2000).

This disagreement is not just about the understanding of the symbol. There is disagreement over what should be measured and how. Sometimes grades are based on criterion references and sometimes on normative comparisons. These are fundamentally different approaches and research suggests they will have different impacts. There is reason to believe that there are observed differences between grades and standardized test scores because they are not designed to measure and communicate the same things (Willingham, Pollack, & Lewis, 2002). Although given the era of accountability and the need for seemingly objective evidence to make decisions, standardized tests are viewed as appealing because they symbolize order and control, appear to measure important educational outcomes or content, and measure what the public values – high achievement and rigor (Airasian, 1988). The problem is that the state mandated tests “supply largely redundant information, but in a manner that is perceived to be more objective and fair than school-based judgments” (Airasian, 1988, p. 308). The value of classroom grades can be viewed as diminishing as a result.

Conley (2000) and Willingham, Pollack, and Lewis (2002) note that proficiency assessments and classroom grades are measuring different experiences. This has not gone unnoticed. Teachers anecdotally have argued that the plethora of standardized tests students take do not accurately measure what they teach and students learn. Brookhart (2003) suggests we may need to look at psychometric theory which has long informed testing experts and consider the role of classroom assessments a bit more carefully because “most of the information that students have about their learning and what it means about the subject, about themselves, and about their futures – comes from classroom assessments” which in turn inform grades (Brookhart, 2003, p. 5). Teachers are not trained, nor when they write classroom assessments do they “calculate reliability estimates, standard error of measurement, validity coefficients, item discriminations, or standard scores, nor do they construct detailed test blueprints. These techniques are based on principles for developing large scale objective tests, with limited relevance to the assessment context of classroom teachers” (McMillan, 2003, p. 34). Because of this, Brookhart (2003) urges a closer study of the relationship between psychometric theory and what she has coined “classroometric theory”.

In an era of assessments based on standards, criterion referenced assessment and grading based on pre-set standards, is the recommended practice over assessment against the norm, or students’ peers (Guskey, 2000). Students respond to criterion referenced assessments over normative assessments. In a study by Wilburn and Felps (1983) where a quasi-experimental non-equivalent control group design was used to study two groups

of middle school math students, students achieved at higher levels under the criterion referenced method of grading versus the students whose scores were determined in comparison to the group. In addition to the performance outcomes, students were asked to respond to how they felt about the grading system they were being assessed under on a Likert scale. They found middle school students evaluated by criterion referenced methods demonstrated significantly more positive attitudes toward the subject than the control group (Wilburn & Felps, 1983). While this thirty year old study is confirmation of what many teachers attempt to utilize as a philosophical foundation for their grading system today, it is the practices that they employ within that grading system that negate the criterion referenced nature of the grade.

Experts say that a major cause of the disagreement over the symbols of grades and teachers inadvertently negating a criterion referenced system of assessment is confounding academic achievement with student behaviors. Among the educational measurement experts there is a belief that grades “should be based exclusively on measures of current achievement and growth, [and that] ability, effort, conduct and other non-academic factors should not be considered” (Cross & Frary, 1999, p. 53). However, many studies report that senders (teachers) and receivers (students and parents) do not necessarily agree or even see that happening in practice (Cross & Frary, 1999; Baron, 2000; Allen & Lambating, 2001). There is evidence that teachers confound academic and non-academic factors into a “hodgepodge” approach to grading where attitude, effort, and achievement are all included in one grade (Cross & Frary, 1999; Brookhart, 1993). They conclude that this general approach is utilized to protect both students and teachers

from negative social or professional impacts of accurately representing grades because grades would no doubt fall if they strictly reported achievement.

In a survey of 307 middle and high school teachers and 8,664 students, respondents confirmed use of this ‘hodgepodge’ approach to grading (Cross & Frary, 1999). In the same survey, it was found that teachers’ view of ideal practice was inconsistent with their responses regarding their actual practice. This was especially true when they were asked if they considered aptitude and ability in grading (Cross & Frary, 1999). This study reported that 25% of teachers raised grades for effort fairly often. In addition, 39% said they considered behavior in assigning grades (Cross & Frary, 1999). In stark contrast and further proving that teachers’ ideal practice is inconsistent with their actual practice, 81% of teachers and 70% of students in the Cross and Frary (1999) study felt that “achievement, effort, and conduct should be reported separately” (p. 61).

The inclusion of non-academic factors like behavior can also mask accurate achievement levels resulting in inflation or deflation of grades. If teachers choose to bump up a grade because a student is well behaved and shows effort, though does not have a strong grasp of the content, there is a misrepresentation of their knowledge. This inflation can be detrimental to a student who believes they are prepared and then does not score well on college entrance examinations or even in college classrooms (Howley, Kusimo, & Parrott, 1999). Conversely, this issue can be problematic if a student’s grade is deflated due to non-academic factors like lack of participation in class. Opportunities may be missed because the student and parent may not understand or believe the student is capable of challenging content or coursework.

There is also evidence that grades are awarded inconsistently within and beyond schools. In a 2010 study of factors affecting grading, Randall and Englherd found “behavior even more so than effort is an important factor to teachers when dealing with borderline students” (p. 1376). In their study of 516 teachers in a stratified random sample in a major metropolitan school district, Randall and Englherd (2010) looked at borderline grading cases, when a student was between a B and a C for example, and found that “students with both high effort and excellent behavior all receive a boost regardless of ability or achievement level” (p. 1376). For grades to be considered valid and reliable, they must only communicate one message: achievement (Marzano, 2000; O'Connor, 2007; Guskey, 2000). The inclusion of non-achievement factors in an academic grade compromises the validity of the grade (Allen & Lambating, 2001; Baron, 2000; ACT, 2005).

Grade inflation may in part be caused by the failure to differentiate academic knowledge from behavior. Evidence of invalid and unreliable grading in the form of grade inflation, especially in low-income schools, is found in ACT’s multi-year study looking at student test scores and their self-reported grades.

An important difference between grade inflation and other factors influencing the reliability of grades is that grade inflation is an increase in grades *over time* for students who are at the same level of achievement, while other factors affecting variability in grades result in assignment of different grades to students at the same level of achievement during the *same* period of time. Because of this, grade inflation is not easy to detect. It requires an examination of grades across time and a stable measure which to compare them (ACT, 2005, p. 1).

In an analysis done by ACT in their thirteen-year study of student self-reported grades in 23 courses, evidence of inflation was found. When comparing GPA against composite

score from 1991 through 2003, there was an average grade inflation of 0.25, or one quarter of one grade point on a four point scale every year. ACT reports that over the 13 years “GPAs in ACT tested public high school graduates increased by about 6.25% without an accompanying increase in ACT composite score” (ACT, 2005, p. 3).

Similarly, the College Board reports that “since 1987, the population of students with A plus, A, and A minus grade-point averages has gone from 28% to a record 37%, while their SAT scores have fallen an average of 13 points on verbal and 1 point on math” (Marzano, 2000, p. 10).

With evidence of grade inflation given the comparison of grades to two nationally normed tests, it is plausible that the messages being sent to students and parents are invalid or unreliable. These messages can be detrimental to students from schools of poverty. In a 1994 study conducted by the Office of Educational Research and Improvement, researchers discovered in a national sample of 8th grade students from high poverty schools that grades were inflated (Marzano, 2000). Students who earned A’s in high poverty schools scored at roughly the same proficiency ratings on reading and math tests given as part of the U.S. Department of Education’s National Education Longitudinal Study as students in affluent schools who earned C’s and D’s. This finding was reinforced in a 1997 study done by Christopher Cross where he found that students in high poverty schools earned A’s and again scored at the same levels on the same exam as students from low poverty schools earning C- to D+ grades (Marzano, 2000). Experts agree that academic knowledge and behavior should be reported separately because of the inaccurate messages that are sent when the two are reported together.

There are technical problems with the ways that grades are often calculated that distort achievement. Awarding a zero for missing or failing work in percentage based grading is an example of this. The use of a zero in grade calculation when a student has a missing assignment or as a consequence for cheating does not allow for evidence of learning. When using percentage based grading, all the other grades are representative of 10% of the grading scale where an F is representative of 59% of the scale. Because of this, ideally, percentages should not be used as they are not equal intervals representing each letter grade. If percentage based grading is used, all scores below 50% should be adjusted up to 50% when there is evidence to score (Guskey, 2004; Guskey, 2000; Erickson, 2010; Reeves, 2004; O'Connor, 2007).

Another technical inaccuracy that allows for the inclusion of student behavior is the use of extra credit. Including extra credit in a grade can mask a student's failure to complete and/or master the content. Often times, the extra credit being offered is not related to the curriculum and is offered with the purpose of raising a grade. This is sometimes referred to as "the tissue box effect" describing a practice where students bring boxes of tissue for the classroom use and receive extra points to raise their course grade. This disadvantages students of poverty as their families may not be able to purchase tissue to inflate their grade as their affluent peers can (Erickson, 2010; O'Connor, 2007).

Reducing scores for late work also distorts the degree to which the grade reflects mastery of the course content. Turning in work late is a behavior and should be reported

separately. All assessments should be assessed for achievement, not compliance (O'Connor, 2007; Wormeli, 2006).

Even when teachers are exposed to grading practices that are considered sound, research shows that it does not impact their practice. In part this is perhaps because many get no measurement training in their teacher preparation programs. In a random sample of 47 such programs, for example, Allen and Lambating (2001) discovered that less than half required a measurement course. Measurement training for in-service teachers does not have a significant impact on changing the practices of teacher grading to align with the recommendations of measurement experts to provide valid and reliable grades (Brookhart, 1993; Guskey, 2009).

It is puzzling that teachers would not adjust their practice when provided with information about sound measurement and grading practices. One answer could be that teachers are concerned with fairness and how grades will be used, so they take it upon themselves to justify the inclusion of factors like effort and behavior because they cannot guarantee that grades will be used only as measures of achievement (Brookhart, 1993; Guskey, 2009).

Educational measurement experts argue that a grade should be a reflection of a student's achievement and should not include non-achievement factors (Austin & McCann, 1992; Cross & Frary, 1999; Guskey, 2000; Marzano, 2000; O'Connor, 2007). Yet, the inclusion of student behavior in grades is a practice often utilized as a form of classroom management and control or extrinsic motivation (Erickson, 2010; Guskey, 2009; Randall & Engelhard, 2010). Evidence of this practice is found in Cross and

Frary's (1999) survey of over 307 middle school and high school teachers and 8,664 students from the same schools. From this sample 25% of the teachers raise grades for high effort 'fairly often' and 39% take conduct and attitude into consideration when assigning final report card grades. Again, contrary to the practice cited, 81% of teachers and 70% of students agreed or tended to agree that the non-academic factors like attitude and effort should be reported separately.

In a state-wide study of policy and practice documents regarding grading, Austin and McCann (1992) found participation to be the second most cited criteria after student performance for grading. Even though this practice is against the recommendations of measurement experts and what teachers and students believe should be happening, teachers want students to succeed and think that including these factors may provide engagement and motivation (McMillan, 2001; Brookhart, 1993; Guskey, 2009). What the motivational literature suggests, however, is that extrinsic motivation, like grades, is not effective in the long term and in the case of learning, it is problematic. In a study in a large southwestern university of how grading systems motivate students, Docan (2006) divided students into two groups: earners and maintainers. The earners started with zero points and added throughout the semester, yielding positive reinforcement while the 'maintainers' began the semester with the maximum number of points available and lost points as work was assessed yielding negative incentives. While both systems motivated students, the overall extrinsic motivation of grades decreased the intrinsic motivation to learn (Docan, 2006). The practice of including non-academic factors like effort and

attitude or overall behavior into a grade to motivate students will likely motivate compliance instead of engagement in learning.

Summary

The reforms instituted at the evaluation site attempted to address both of the broad issues discussed in this literature review: the fact that there are not common understandings of the meaning of grades and the technical problems that have caused some scholars and practitioners to conclude that grading is broken. Thus, the evaluation of those reforms should shed light on these issues identified in the literature.

Chapter 3

Research Methods

Methods Overview

This evaluation utilizes a mixed methods approach to evaluate the effectiveness of the school's initiative to reform grading practices. The goal of the initiative was to provide students, parents and other teachers with a clearer and more accurate evaluation of the two distinct aspects of students' performance in school: mastery of academic content and mastery of life skills such as effort, behavior and timeliness. To achieve this goal the initiative focused on extracting the life skills into a separate grade and reforming the practices that contributed to the accuracy of calculation of the knowledge grade. This is a developmental evaluation intended to analyze impact of the reforms to provide lessons learned to other schools and districts that are seeking to reform grading practices.

Measures and Indicators

The evaluation uses data from surveys, interviews, questionnaires, grades and standardized tests to evaluate the outcomes and implications of the grading reform. The evaluation focuses on seven measures using seven different indicators. These can be seen in Table 2 on the next page with explanation to follow.

Table 2: *Measures and indicators of evaluation study*

Measure	Indicator
1. Correlation between knowledge grades and state standardized test scores	Regression analysis of math knowledge grade and MCA math test score
2. Correlation between knowledge grades to teacher perception of effort	Regression analysis of math knowledge grade and teacher assigned effort grade
3. Relationship between teacher and student perceptions of effort	Student responses to college readiness survey questions measuring effort and self regulatory strategies compared to teacher assigned effort grade
4. Student perceptions of separating the knowledge and life skills grades	Responses to student survey and student focus groups
5. Student perceptions of related changes in grading practices for calculating the knowledge grade	Responses to student survey and student focus groups
6. Teacher perceptions of separating the knowledge and life skills grades	Responses to teacher survey, interviews and questionnaires
7. Teacher perceptions of related changes in grading practices for calculating the knowledge grade	Responses to teacher survey, interviews and questionnaires

Measure 1: Correlation between knowledge grades and standardized test scores

The evaluation analyzed how well knowledge grades assigned by teachers correlated with student scores on standardized tests. The correlation between knowledge grades and state standardized test scores was an important measure because if grades contained both content knowledge and life skills, it was expected that there would be misalignment between the grades and an objective measure of the content knowledge such as a standardized achievement test. In addition, if the grade was calculated using unsound grading practices like using a zero for missing work, allowing for extra credit, or deducting points for work turned in late, the knowledge grade would not accurately reflect the content mastered by the student. If the grading reform separated life skills

and academic content knowledge and reformed the methods of grade calculation took place, it was expected that the knowledge grade and the test scores would be more closely correlated after the reforms were implemented.

To assess this, the evaluation used 8th grade math grades from 2007-2008 and 8th grade math knowledge grades from 2008-2009 through 2011-2012 by student ID number obtained from the school district's student information system and the state of Minnesota's standardized test, the MCA (Minnesota Comprehensive Assessment) 8th grade math standardized test scores by student ID number obtained from the school district's data warehouse. These data were compared using a linear regression analysis to determine the correlation coefficient between the grade and standardized test score. For each year of the correlation study, the linear regression analysis sample size was as follows:

Table 3: *Measure 1 sample sizes*

Year	Number of Students
2007-2008	141
2008-2009	242
2009-2010	231
2010-2011	261
2011-2012	279

The correlation of course grades and standardized test scores was determined by calculating the linear association between grades (x) and the state standardized test (y) to

determine the correlation coefficient. The correlation between 8th grade math grades and students' performance on the state standardized test was the only data that extended outside of the period of the study, although it is what originally prompted the statistical evaluation of the grading system. Student grades and standardized test scores were data available for analysis over time, so a consistent approach to evaluating the correlation in 8th grade was established at the outset of the change.

Using grades and standardized test scores had limitations. One limitation of the correlation study was that it examined grades and standardized test scores for each year only at the 8th grade and was not broken out by student groups. Consequently, the correlations that either got closer or had little change could have been affected by the change in student cohorts. Another limitation was that the changes that were made in the approaches to grading occurred with a particular group of students, i.e., only those enrolled in 8th grade Pre-Algebra in the pair pilot in 2008-2009. As a result, the data from prior to the changes as represented in the correlation study of 2007-2008 only included the students in Pre-Algebra in 8th grade as that was the only course that experienced changes to grading practices in 2008-2009. Changes were made and used in all 8th grade math classes 2009-2012 and in all math classes in the study site 2010-2012.

Measure 2: Correlation between knowledge grades to teacher perception of effort

The study analyzed how well teacher knowledge grades correlated with teacher perceptions of student effort. The correlation between knowledge grades and effort grades was an important factor to measure because if grades were high and effort was reported as low, one would assume that there was misalignment between either the effort

needed to achieve a high grade or the teacher’s perception of the effort the student was putting forth to earn the grade. The reverse of this was also important to understand because if a student was earning a low knowledge grade and the teacher reported the student had put forth a great deal of effort, one would assume the student either struggled greatly with the content and an intervention was necessary or the teacher had a misperception of the student’s effort.

To assess this, using student ID numbers the correlation study used 2011-2012 final knowledge grades in math and the 4th quarter life skills effort grades in math obtained from the student information system for grades 6, 7, and 8 respectively. These data was compared using a linear regression analysis to determine the correlation coefficient between the knowledge grade and the effort grade. For each grade level in the study, the linear regression analysis sample sizes were as follows:

Table 4: *Measure 2 sample sizes*

Grade	Number of Students
6	318
7	274
8	279

Measure 3: Relationship between teacher and student perceptions of effort

The evaluation analyzed how well aligned teachers and students were in their perceptions of student effort. The relationship between teacher and student perceptions of effort was an important factor to measure because one would hope that teachers and

students would have similar perceptions of effort so that the signal being sent by the teacher in the form of an effort grade was accurately understood by the student for action.

To assess this an effort grade analysis was conducted by CAREI, Center for Applied Research and Educational Improvement in the College of Education and Human Development at the University of Minnesota, using eight questions from a larger survey given to the entire student body in the winter of 2012 as part of an evaluation of the school's participation in the University of Minnesota's college and career readiness program, Ramp-Up to Readiness™. The survey questions used to square the correlations of teacher perception, the effort grade, to student perceptions are listed in Table 5 and can be found in Appendix B.

Table 5: *Measure 3 survey questions*

Instrument (<i>n size</i>)	Question
CAREI Survey (<i>n=247</i>)	Q22 – If, on a scale of 106, a one is the laziest student you know, and a six is the hardest working, what number are you?
CAREI Survey (<i>n=247</i>)	Q25 – How often do you work hard in class and on assignments even when you don't like the class?
CAREI Survey (<i>n=247</i>)	Q27 – How often do you give up TV, video games, Facebook, texting, or time with friends to study for a test or do an assignment for school?

Using student ID numbers, the 2012 third quarter math knowledge grades of the 8th graders and the third quarter math life skills effort grades of the 8th graders were obtained from the student information system and provided to CAREI for matching with the survey data.

Measure 4: Student perceptions of separating the knowledge and life skills grades

The evaluation analyzed views on separating the knowledge and life skills grade. Student perception of separating and communicating mastery of academic content divorced from mastery of life skills was an important factor to measure because if students did not see the two as distinct and valuable in their own right, the goal of the reform initiative would not be met in part.

To assess this, eight questions from the student survey and two questions from student focus groups were analyzed to determine if students saw value in the separation of the two grades. The questions from these two instruments are listed in Table 6 on the following page.

Table 6: *Measure 4 survey questions and group interview questions*

Instrument	Question
Student Survey (n=719)	Q5 – The way that grades are given at this school is fair
Student Survey (n=712)	Q7 – A students’ grade should be based only on how well he or she has learned what was taught in the class
Student Survey (n=708)	Q8 – A students’ grade should include things like effort and behavior along with how well the student has learned what was taught in the class
Student Survey (n=707)	Q13 – I like the approach of giving students separate Knowledge and Life Skills grades that is being used at this school
Student Survey (n=712)	Q14 – The Knowledge grades that I have received at this school are usually a good measure of how well I learned what was taught in class
Student Survey (n=715)	Q15 – The Knowledge grade at this school is made up of enough information (tests, quizzes, projects, assignments) to accurately represent what I know
Student Survey (n=718)	Q16 – The Life Skills gradethat I have received in my classes at this school are a good measure of my behavior in those classes
Student Survey (n=717)	Q17 – The Life Skills grade that I have received in my classes at this school are a good measure of the amount of effort I put into those classes
Student Group Interview (n=36)	Q4 – When you are given a Knowledge grade in a course at this school, what do you think that grade means?
Student Group Interview (n=36)	Q6 – When you are given a Life Skills grade in a course at this school, what do you think that grade means?
Student Group Interview (n=36)	Q8 – Do you think it is a good idea to give students both Knowledge grades and Life Skills grades? Why?

The student survey administered in the spring of 2012 was optional and anonymous. Passive consent was obtained from parents for student participation in the survey and possible participation in a group interview. Parents of students were informed

through a letter mailed to the home of every student in both English and Spanish that the online survey would be administered through the students' online portal account in the student information system and that a small sample of students would be asked if they wanted to participate in a group interview. Parents were offered the opportunity to decline their child's participation in either the survey or the group interviews two weeks before the survey was launched. Remaining students were then provided an opportunity to decline the survey themselves at the time it was administered either prior to the survey starting, or by exiting at any point while taking the survey. Students responded to the survey during their homeroom time using the school's computer labs. Students could only take the survey once, though could do so at any time from any computer with internet access during the survey window of approximately two weeks. Both a summary and detail report were obtained by the evaluator from the student information system at the conclusion of the survey period. A copy of the student survey can be found in Appendix C. The response rate for the student survey was quite high at 74.6%, 726 out of 973 eligible students took the survey.

Student group interviews were also part of the evaluation design. There were six student group interviews – two at each grade 6, 7, and 8 with approximately 6-8 students in each group. In the spring of 2012, students were identified by school counselors to represent a cross section of the student body in the group interview process. Counselors were asked to identify students representing members from different racial, gender and socioeconomic backgrounds. In addition, counselors were asked to factor in academic achievement, and behavior of the students in order to have a population similar to that of

the school represented. Counselors then contacted the students to see if they had an interest in participating. Parents had already provided passive permission to allow their student to participate if they did not originally respond to the notice mailed to the home outlining the survey and group interview process. Group interview times were scheduled and students missed the end of whichever class they were in at that time. Students who came to the group interviews were provided with an informed consent form that outlined the process and explained that participation was voluntary. Students signed the informed consent form prior to the focus group beginning. The protocol used for the informed consent process and student group interviews can be found in Appendix D. The group interviews were recorded and coded to identify similar response subjects per question. Like responses were then analyzed for frequency by grade-level and as a whole.

Measure 5: Student perceptions of grade calculation changes

The evaluation study analyzed how students viewed technical changes in the way grades were calculated. Student perception of related changes in grading practices for calculating the knowledge grade was an important factor to measure because if students did not see the changes in practices that impacted the calculation of the knowledge grade as being advantageous to accurately communicating mastery of academic content, then the goal of a clearer and more accurate way of communicating mastery of academic content would not be met. The specific practices that were examined included the elimination of the use of zero, the elimination of the use of extra credit, allowing students to re-take tests, and not deducting points for work turned in late.

To assess this, seven questions from the student survey and one question from the student group interviews were analyzed to determine if students perceived the changes in knowledge grade grading practices as clearly and accurately communicating mastery of academic content. Those questions are listed in Table 7.

Table 7: *Measure 5 survey questions and group interview questions*

Instrument	Question
Student Survey (n=719)	Q5 – The way that grades are given at this school is fair
Student Survey (n=714)	Q9 – A student who completes a test or assignment should not receive a score lower than 50% on that test or assignment in order to insure that the student always has a chance of passing the course
Student Survey (n=712)	Q10 – Students should always have the opportunity to retake a test in order to raise their score
Student Survey (n=720)	Q11 – Students should not be allowed to earn extra credit in order to raise their grade in a course
Student Survey (n=715)	Q12 – A student’s grade should not be lowered for turning in work late
Student Survey (n=717)	Q18 – Which of the following factors should be the most important in determining a student’s final grade in a course in your opinion? Attendance, Classroom behavior, Completion of homework assignments, Effort, Extra credit, Grades on homework assignments, Grades on essays and papers, Participation in class, Scores on quizzes, Scores on tests
Student Survey (n=709)	Q19 - Which of the following factors should be the least important in determining a student’s final grade in a course in your opinion? Attendance, Classroom behavior, Completion of homework assignments, Effort, Extra credit, Grades on homework assignments, Grades on essays and papers, Participation in class, Scores on quizzes, Scores on tests
Student Group Interview (n=36)	Q2 – How accurately do you think the course grades you receive at this school capture what you know and can do?

Measure 6: Teacher perceptions of separating the knowledge and life skills grades

The evaluation analyzed teacher views on separating the knowledge and life skills grades. Teacher perceptions of separating the knowledge and life skills grades was an important factor to measure because it had been the focus of staff development and the major change initiative in the building the year prior to the study. If teachers' reported perceptions did not align with the goal of more accurately communicating the two distinct aspects of student performance, the staff development and change initiative would not have been effective. Additionally, the change in practice to separate reporting academic content mastery and life skills required not only new learning about grade calculation, it ultimately required more work of the teacher. It was important to know that given the additional work, teachers saw the separation of grades as clearly and accurately communicating student performance.

To assess this, eight questions from the teacher survey and two questions from the staff interviews and questionnaires were analyzed to determine if teachers' perceptions of the separation of mastery of content from mastery of life skills more clearly and accurately communicated student performance. These questions are listed in Table 8 on the following page.

Table 8: *Measure 6 survey questions and interview and questionnaire questions*

Instrument	Question
Teacher Survey (n=42)	Q5 – The way that grades are given at this school is fair
Teacher Survey (n=42)	Q7 – A students’ grade should be based only on how well he or she has learned what was taught in the class
Teacher Survey (n=42)	Q8 – A students’ grade should include things like effort and behavior along with how well the student has learned what was taught in the class
Teacher Survey (n=42)	Q13 – I like the approach of giving students separate Knowledge and Life Skills grades that is being used at this school
Teacher Survey (n=42)	Q14 – The Knowledge grades that I assign at this school are usually a good measure of how well a student has learned what was taught in class
Teacher Survey (n=42)	Q15 – The Knowledge grade at this school is made up of enough information (tests, quizzes, projects, assignments) to accurately represent what students know and are able to do
Teacher Survey (n=42)	Q18 – Which of the following factors should be the most important in determining a student’s final grade in a course in your opinion? (Please choose 3): Attendance, Classroom behavior, Completion of homework assignments, Effort, Extra credit, Grades on homework assignments, Grades on essays and papers, Participation in class, Scores on quizzes, Scores on tests
Teacher Survey (n=42)	Q19 - Which of the following factors should be the least important in determining a student’s final grade in a course in your opinion? (Please choose 3): Attendance, Classroom behavior, Completion of homework assignments, Effort, Extra credit, Grades on homework assignments, Grades on essays and papers, Participation in class, Scores on quizzes, Scores on tests
Teacher Interview/Questionnaire (n=20)	Q8 – Do you think it is a good idea to give students both a Knowledge grade and a Life skills grade? Why?
Teacher Interview/Questionnaire (n=20)	Q9 – Have you observed any changes in students’ behavior and/or attitudes that you believe were significantly influenced by the adoption of Grading for Learning at this school? If so, what are they?

The teacher survey was administered in the spring of 2012 through the staff's access to the online student information system. The complete survey can be found in Appendix E. All staff were told about the survey and its purpose for the study via an email from the principal a few days before it was launched into their task inbox within the student information system. It was explained to staff that the responses would be anonymously recorded and that it was a voluntary survey. Once the survey was launched, staff chose whether or not they wanted to open it and participate. Their responses were recorded anonymously within the student information system. Both a summary and detail report were then obtained by the evaluator from the student information system at the conclusion of the survey period. Staff could complete the survey only once from any computer with an internet connection during the two week window. The response rate was fairly high at 68.8%; 44 of 64 eligible staff members participated.

The teacher interview process took place using both face to face interviews and by emailing the interview protocol to respondents to complete. Teachers were initially invited to participate in the one on one interviews with the principal in the spring of 2012. All teachers we invited to participate in a voluntary interview using the protocol found in Appendix F. Teachers who participated were given the informed consent form to sign. The initial response rate in spring of 2012 for interviews was low with only 6 of a possible 64 teachers volunteering to participate in the interview. At that time the principal decided to wait until the fall of 2012 to conduct further interviews. Access to the teachers for face to face interviews became challenging to complete in the fall of

2012 as the evaluator was no longer the building principal. After obtaining permission from the evaluator's advisor, the evaluator emailed the interview protocol including the informed consent form as seen in Appendix F to 21 teachers in the school. The response rate of the emailed interview questionnaires was 14 with some being electronically returned and some mailed back to the investigator. The combined interview and questionnaire response rate was 31% with 20 out of a total 64 teachers participating. The interviews and questionnaires were all individually coded using like subject codes for responses to each question. The comments were then analyzed for the frequency with which subjects were cited.

Measure 7: Teacher perceptions of grade calculation changes

The evaluation analyzed teacher views of technical changes in the way grades are calculated. Teacher perceptions of related changes in grading practices for calculating the knowledge grade was an important factor to measure because the key practices changed previously distorted the accuracy of the knowledge grade. Practices like deducting points for late work, allowing for extra credit, using percentages below 50%, and not allowing for test retakes, distorted communication about a student's mastery of the academic content. If teachers' perceptions regarding these changes in practice did not support the changes, then the staff development leading to the grading reform would be seen as not effective.

To assess this, five questions from the teacher survey and five questions from the teacher interviews and questionnaires were used to analyze teacher perceptions of the

changes in practices used for calculating the knowledge grade. These questions are listed in Table 9.

Table 9: *Measure 7 survey questions and interview and questionnaire questions*

Instrument	Question
Teacher Survey (n=42)	Q5 – The way that grades are given at this school is fair
Teacher Survey (n=42)	Q9 – A student who completes a test or assignment should not receive a score lower than 50% on that test or assignment in order to insure that the student always has a chance of passing the course
Teacher Survey (n=42)	Q10 – Students should always have the opportunity to retake a test in order to raise their score
Teacher Survey (n=42)	Q11 – Students should not be allowed to earn extra credit in order to raise their grade in a course
Teacher Survey (n=42)	Q12 –A student’s grade should not be lowered for turning in work late
Teacher Interview/Questionnaire (n=20)	Q2 – What factors do you consider when you assign a Knowledge grade for a course at this school?
Teacher Interview/Questionnaire (n=20)	Q9 - Have you observed any changes in students’ behavior and/or attitudes that you believe were significantly influenced by the adoption of Grading for Learning at this school? If so, what are they?
Teacher Interview/Questionnaire (n20)	Q10 – What do you think students at this school think is the reason they are allowed to retake tests?
Teacher Interview/Questionnaire (n=20)	Q11 – What do you think students at this school think is the reason they are not allowed to earn extra credit?
Teacher Interview/Questionnaire (n=20)	Q12 – What do you think students at this school think is the reason their grades are not lowered for turning work in late?

Summary

The mixed methods design of this evaluation held the potential to yield valuable information on the effectiveness of the grading reforms at the study site and provide insights on the broader discussions of grading that are taking place among practitioners and scholars across the nation.

Chapter 4

Findings

This evaluation examined an initiative to reform grading practices with the goal of providing students, parents, and other teachers a clearer and more accurate evaluation of the two distinct aspects of students' performance in school: mastery of academic content, and mastery of life skills such as effort, behavior, and timeliness. Seven measures were used to evaluate the effectiveness of the initiative. The measures and indicators used are listed in Table 10, and each will be discussed.

Table 10: *Measures and methods used to evaluate the grading practices reform initiative*

Measures	Indicator
1. Correlation between knowledge grades and state standardized test scores	Regression analysis of math knowledge grades and state standardized test scores
2. Correlation between knowledge grades and teacher effort grade	Regression analysis of math knowledge grades and teacher assigned effort grades
3. Correlation between student perception of effort and teacher perceptions of effort	Regression analysis of student responses to college readiness survey questions measuring effort and self regularorty strategies compared to teacher assigned effort grade
4. Student perceptions of separating the knowledge and life skills grades	Responses to student survey and student group interviews
5. Student perceptions of related changes in grading practices for calculating the knowledge grade	Responses to student survey and student group interviews
6. Teacher perceptions of separating the knowledge and life skills grades	Responses to teacher survey, interviews and questionnaires
7. Teacher perceptions of related changes in grading practices for calculating the knowledge grade	Responses to teacher survey, interviews and questionnaires

Measure 1: Correlation between knowledge grades and standardized test scores

The findings on the correlation of course grades to the standardized test scores are reported in two parts. First, the correlation of grades to standardized test scores are reported over time for grade 8 where a change in the degree of correlation was found. Prior to the changes in grading, the correlation coefficient was considered negligible to low, falling in the range of .00-.20. Following the intervention the correlation coefficient was considered moderate, falling in the range of .40-.60 (Ravid, 2005). Second, the degree of correlation of grades to the standardized test score for the study year for grades 6, 7 and 8 are reported. There was a moderately correlated finding for all grades.

Grade 8 Findings Over Time

Correlation of course grades to standardized test scores increased in grade 8. In 2007-2008, one year prior to any changes in grading in the 8th grade Pre-Algebra math classes, the R^2 of the 141 students enrolled in the Pre-Algebra courses was .197. This low correlation coefficient reinforced what the teachers feared, which was that the grade students received was not accurately representing what they knew about math. Although some teachers thought primarily that grade inflation was occurring prior to the correlation study, it can be seen in the visual representation of the original correlation in the form of a scatterplot in Chart 1 that in fact both inflation and deflation were occurring. Students who were not passing the state standardized test were receiving As and Bs, evidence of inflation; at the same time, students passing the state standardized test were receiving Ds and Fs, evidence of deflation.

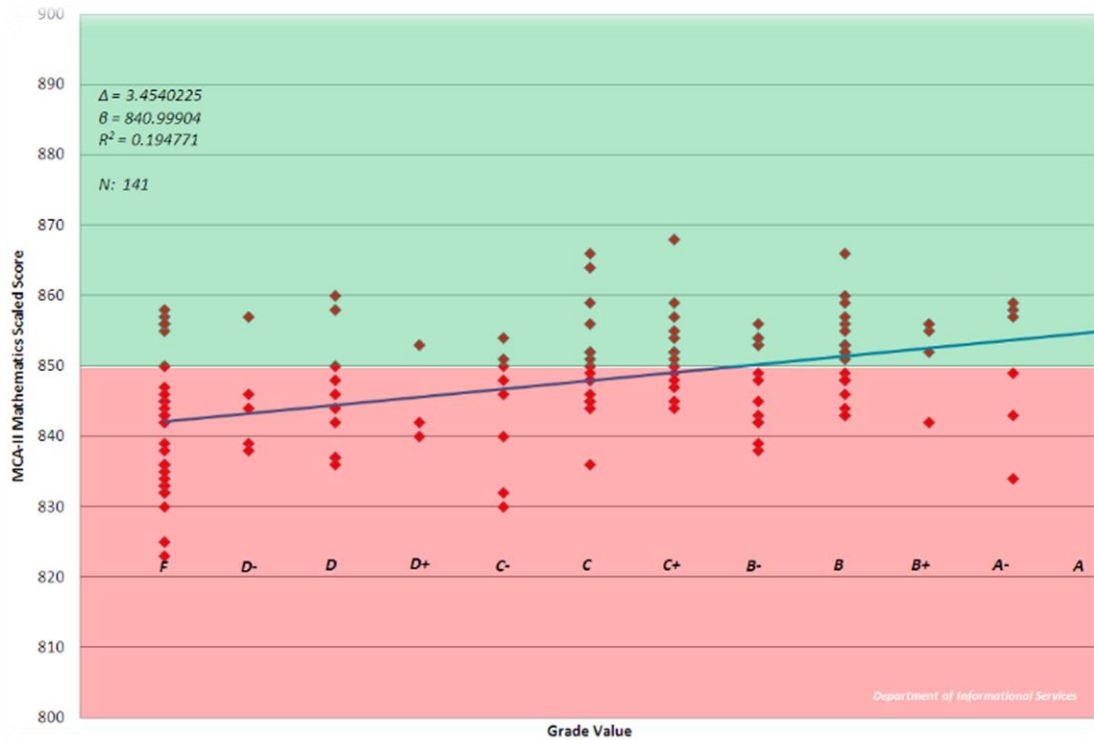


Figure 1. 2007-2008 grade 8 Pre-Algebra correlation analysis scatterplot

In 2008-2009, the two 8th grade math teachers changed their grading practices with the intent of eliminating factors that artificially inflated or deflated math grades. Their goal was to insure that their math grades represented what students had learned and knew about the standards taught in the 8th grade Pre-Algebra class and tested on the state standardized assessment, the MCA.

Changes that the teachers made included calculating homework at a 0% weight in the overall grade. Homework was recorded in the gradebook so that students and parents could see how students were doing, but the points earned were not calculated in the course grade. The teachers scored what the students demonstrated they knew on the assesment, and not the process of their practicing (Guskey & Bailey, 2010; Marzano,

2000; O'Connor, 2007). In addition, the teachers did not offer extra credit, nor did they include any in grades. This did not allow grades to be artificially increased (O'Connor, 2007).

End of quarter final exams were given. If a student's final exam grade for the quarter was higher than the average of quizzes and tests that made up their quarter grade, the final exam grade became the final quarter grade. The rationale for this was the teachers felt if the student demonstrated mastery of the content by the end of the quarter, he/she should not be penalized for not learning it sooner (Guskey, 2000; O'Connor, 2007).

With the new practices in place for 123 students in the Pre-Algebra courses, the R^2 was .427. This was considered a moderate correlation coefficient in the correlation between course grades and the state standardized test (Ravid, 2005). Figure 2 illustrates that there was less inflation and deflation by the observed decrease in students who received As not passing the state standardized test and no students who received Fs passing the state standardized test.

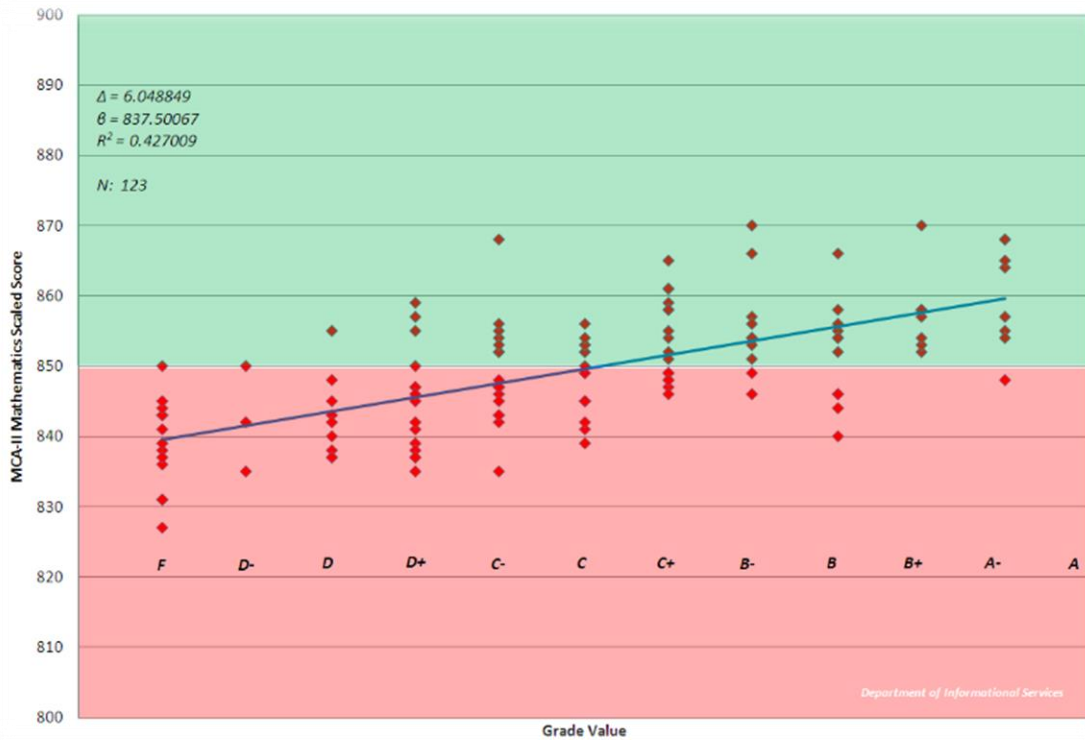


Figure 2: 2008-2009 grade 8 Pre-Algebra correlation analysis scatterplot

The correlation study changed to include all 8th grade math students in 2009-2010. That was also the year that all aspects of Grading for Learning were included in the change. In 2009-2010 the students were given two grades:

- The life skills grades allowed teachers to report to parents about the factors that no longer influenced the knowledge grade. The life skills grade was based on a rubric and reported on practice (timeliness of homework completion), preparation for class, behavior, and teamwork/participation.
- The knowledge grade eliminated the factors that artificially inflated and deflated grades like extra credit and points for participation and continued to calculate homework, now called practice work, at 0%. Students were allowed

to retake tests for a better score if they reviewed their first test with the teacher.

The R^2 for math grades to the state standardized test in 2009-2010 was .430, a decrease over the previous year. A visual representation of this correlation analysis can be found in the form of a scatterplot in Figure 3.

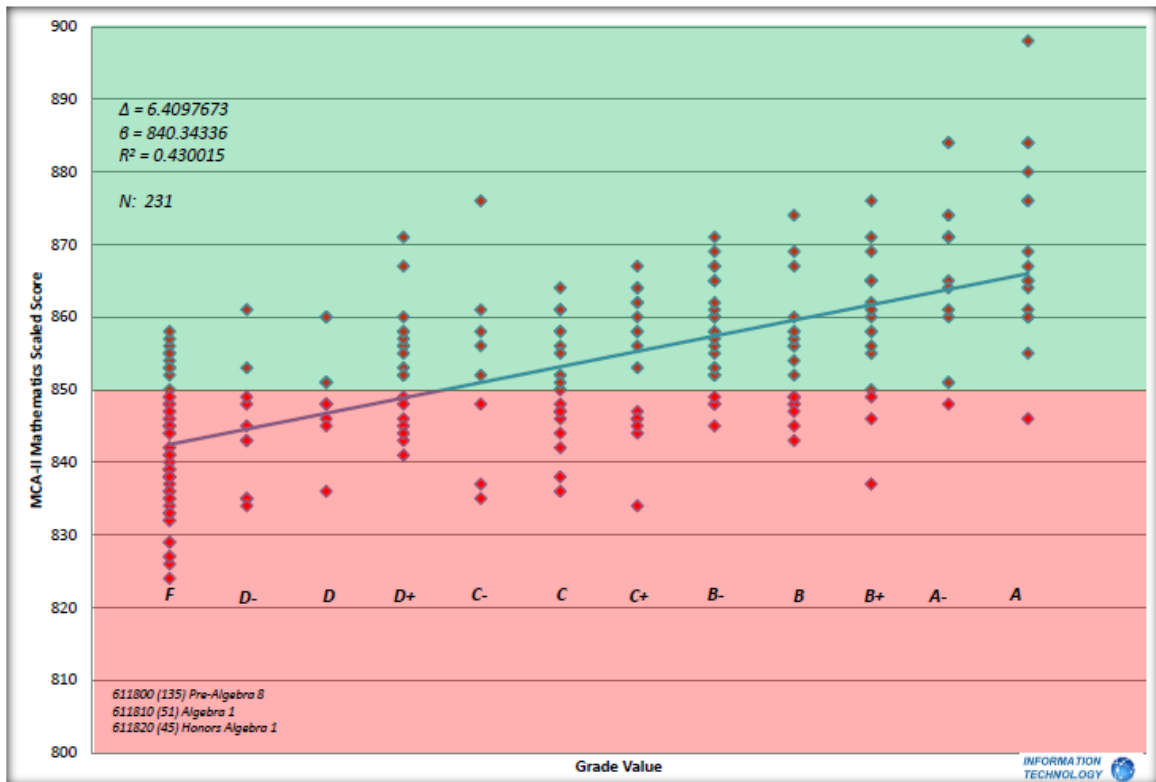


Figure 3. 2009-2010 grade 8 math correlation analysis scatterplot

Again, it should be noted that the comparison is not within the same cohort, so other factors beyond the grading system could have influenced the drop in the correlation coefficient. However, the relation in this year is still better than the initial correlation prior to any change.

In 2010-2011, Grading for Learning was instituted building-wide. At this point, students received the knowledge and life skills grades in all classes. The correlation study continued only for 8th grade math. The R^2 value for all 8th grade math students' math knowledge grade to the state standardized test was .470, a rise in correlation from the previous year. The visual representation of this analysis in the form of a scatterplot can be found in Figure 4.

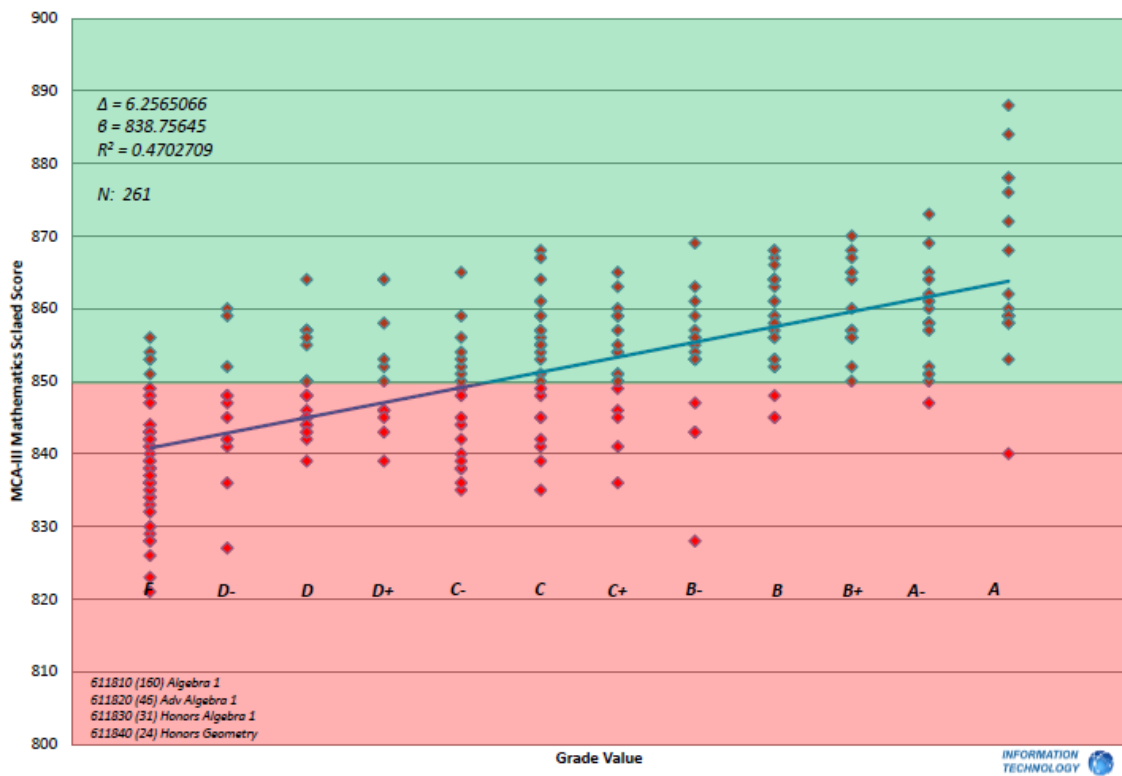


Figure 4: 2010-2011 grade 8 math correlation analysis scatterplot

Findings of Degree of Correlation of Grades to Standardized Tests for Study Year

The 2011-2012 school year was the year of the evaluation and the second year of the building-wide change. During the evaluation year a correlation study of students in each grade in all math classes was conducted. The R^2 for grade 8 was .584 as represented in Figure 5.

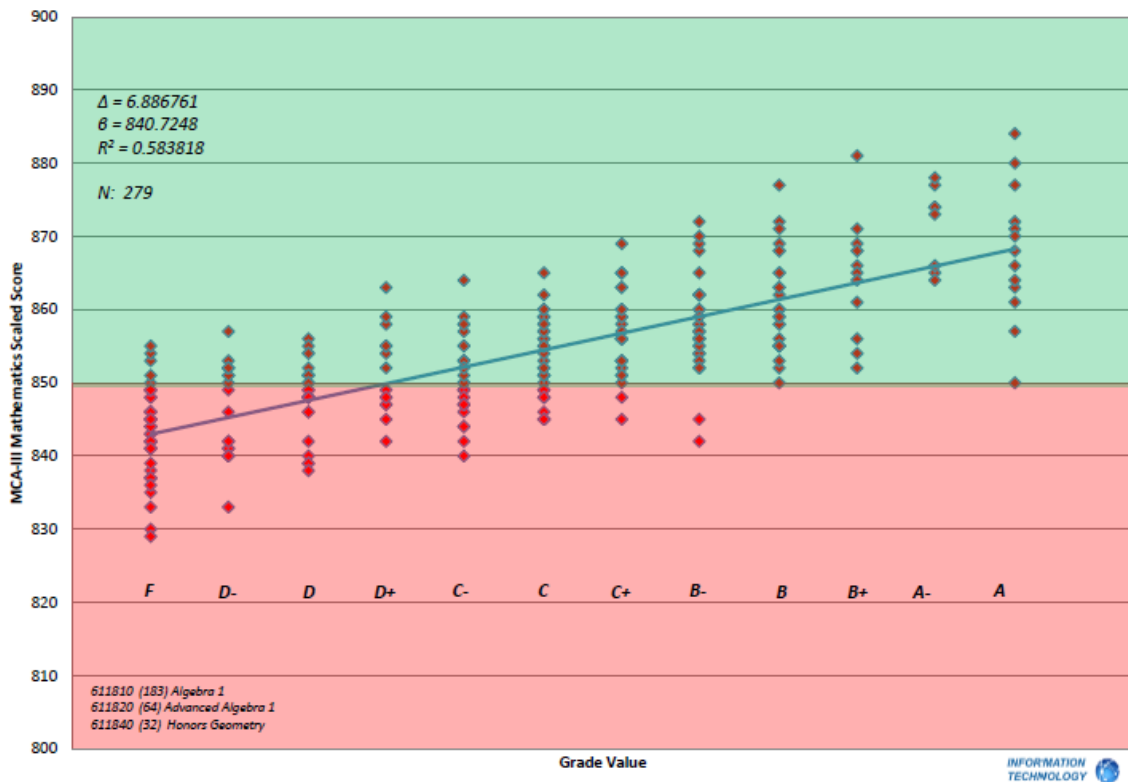


Figure 5: 2011-2012 grade 8 math correlation analysis scatterplot

The R^2 for grade 7 was .603 as seen in Figure 6.

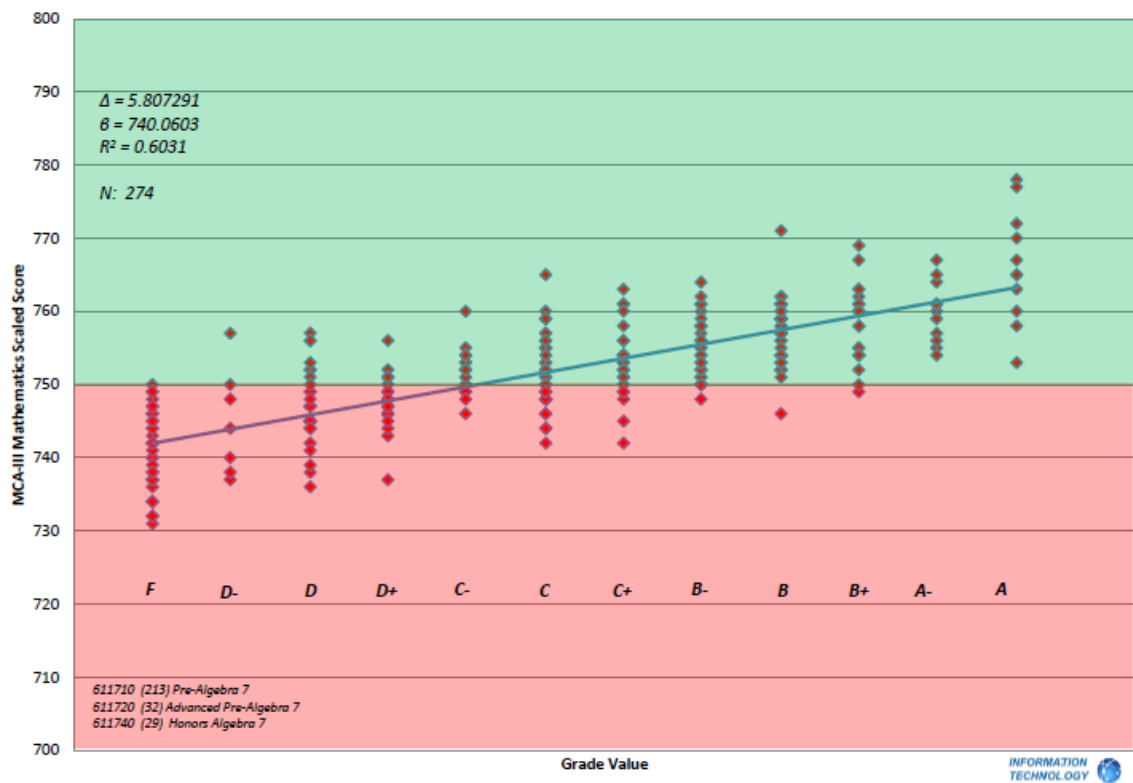


Figure 6: 2011-2012 grade 7 math correlation analysis scatterplot

The R^2 for grade 6 was .625 as seen in Figure 7. Comparison data to other years for grades 6 and 7 were not calculated.

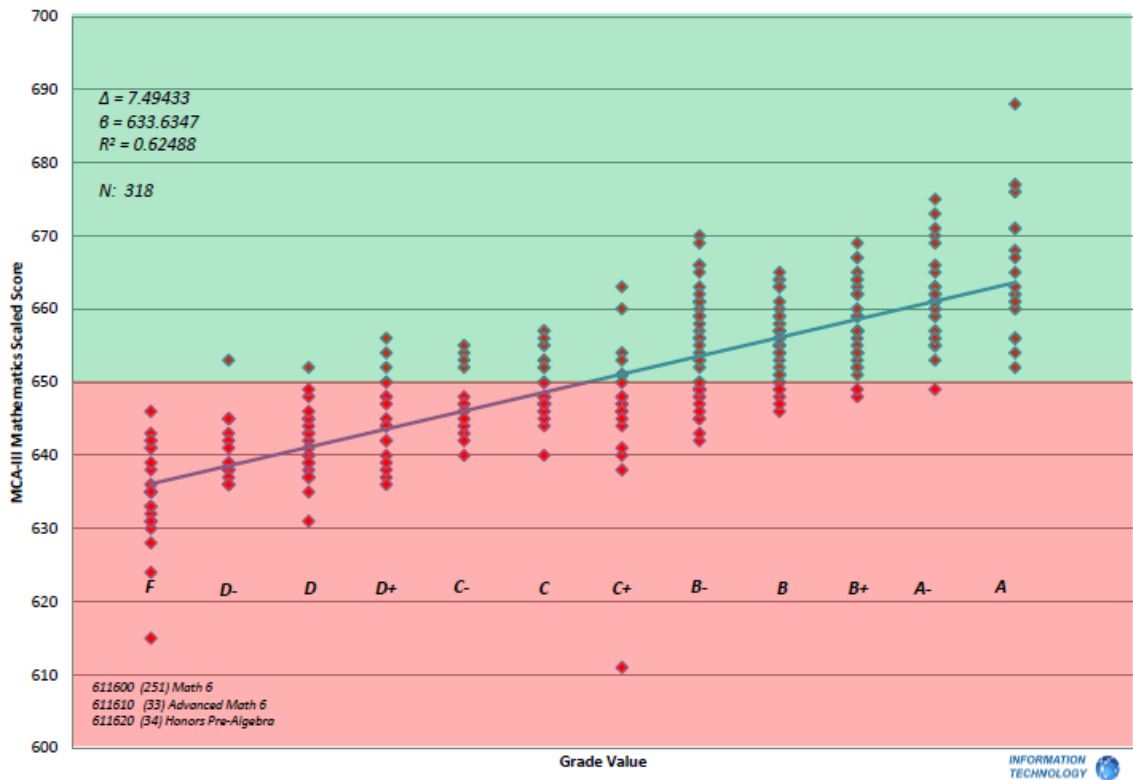


Figure 7: 2011-2012 grade 6 math correlation analysis scatterplot

Correlation coefficients of the three grades during the study year do show relative consistency among the three grades at .584 (8th grade), .603 (7th grade), and .625 (6th grade). This suggests that the standard approach to grading through the changes instituted in Grading for Learning had an impact on the correlation between the math course grade and the state standardized test by showing a positive correlation.

Measure 2: Correlation between knowledge grades and teacher perception of effort

Correlation of knowledge grades assigned by teachers and effort grades assigned by teachers was uneven by grade level. The correlations were closest at the extremes for all grade levels. Students who had very high or very low knowledge grades were also much more likely to receive corresponding high or low effort grades.

The change in the grading system to include the life skills grade was intended to give students and parents a clearer understanding of the other important aspect of student performance: mastery of life skills. Measure 2, the correlation between knowledge grades and teacher perception of effort as reported in the effort grade, was intended to evaluate if there was alignment or misalignment between the two. Misalignment meant either the effort that was needed to achieve a high knowledge grade was not actually needed or that the teacher's perception of effort that the student was putting forth to earn the high grade was not accurate.

Unlike the somewhat consistent numbers in the correlation coefficients of grades to standardized test scores seen in Measure 1 among the three grade levels in 2011-2012, there was more discrepancy when correlating the end of course math grade to the 4th quarter effort grade assigned by the math teacher.

In grade 8, the R^2 value for the correlation of math grade to the effort grade was .328. A visual representation of the correlation in the form of a scatterplot can be found in Figure 8 on the next page. While the number of students represented on the scatterplot is 279, each dot represents many students as there is overlap at each point.

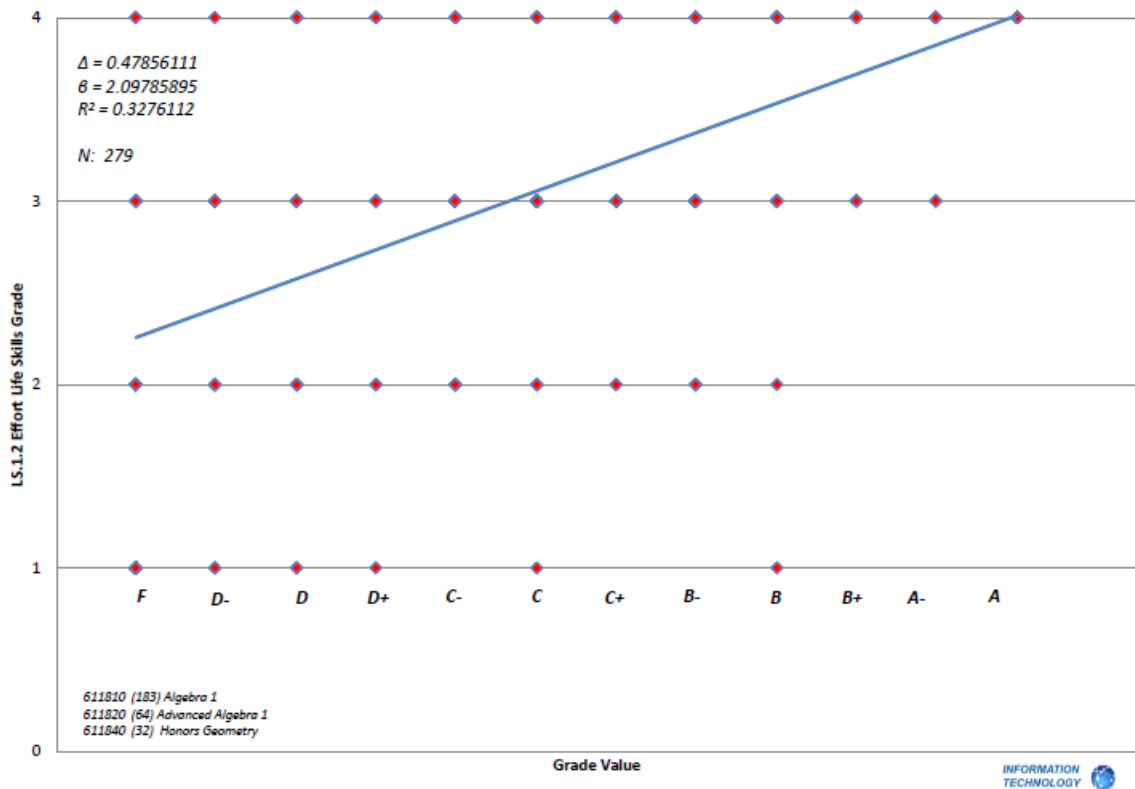


Figure 8: 2011-2012 grade 8 math grade and effort grade correlation analysis scatterplot

Unlike the other two grades, there were students in grade 8 who earned D- and Fs who also earned effort grades of a 4. The effort grade of a 4 stated that the student “consistently demonstrates initiative and self-direction.” It is also worth noting that no students in 8th grade who earned a B+, A- or A earned below a 3 effort grade. This discrepancy could mean that teachers were inflating a student’s effort grade. Conversely, it could mean that some students’ earning Fs were working extremely hard and not learning the content, which highlights the need for a different approach to instruction or remediation for those students.

In grade 7, the R^2 value for the math grade to the effort grade was .609, the highest coefficient for any of the grade levels. The visual representation of the correlation in the form of a scatterplot can be found in Figure 9.

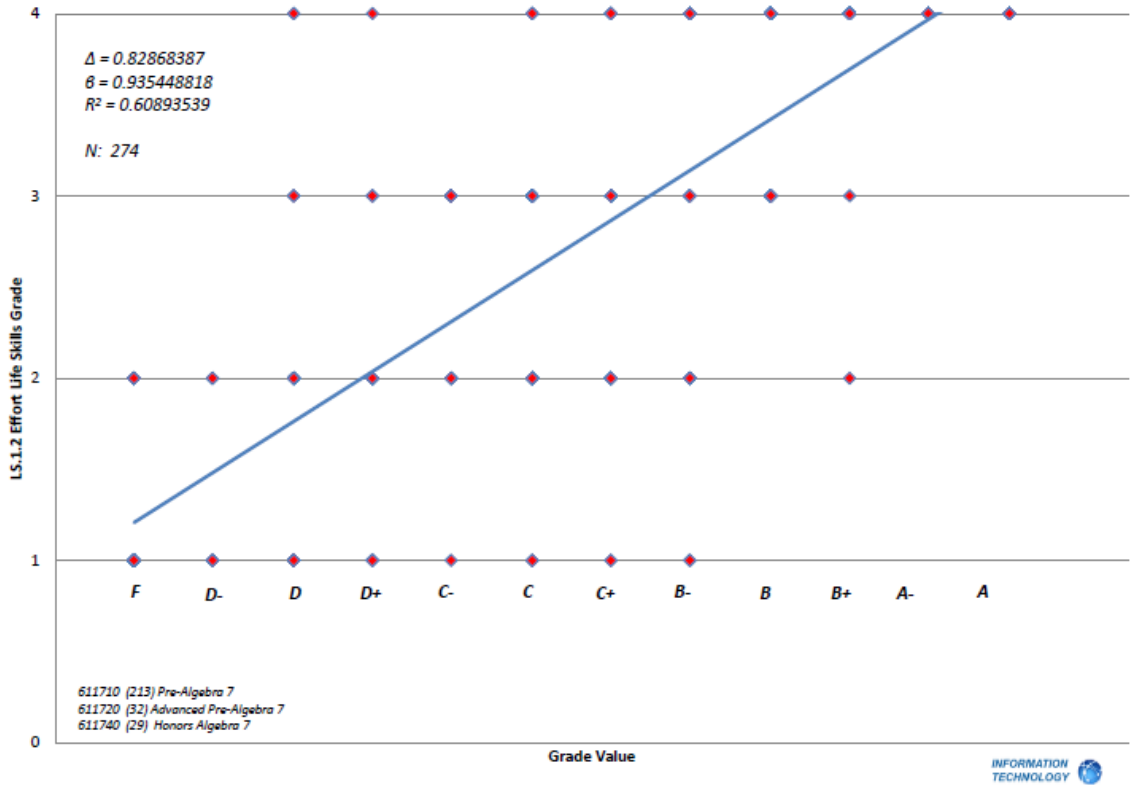


Figure 9: 2011-2012 grade 7 math grade and effort grade correlation analysis scatterplot

In 7th grade no students who received a D- or F knowledge grade received above a 2 effort grade. No student earning an A or A- received below a 4 effort grade. It could be argued that more effort or at least teacher perception of effort in 7th grade math correlated with higher math grades. This would seem like a rather obvious conclusion for most people, but given the lower correlation coefficient in 8th grade, what may be

considered an obvious conclusion by some is not necessarily the case if one believes that the teacher's perception of effort as evidenced in the assigned grade is accurate.

Similar to grade 7, the grade 6 correlation study found no students who earned an A in math received less than a 4 effort grade. However, the 6th grade R² value is lower at .444 as students who earned D- and F received 3 effort grades and, conversely, students who earned A-s received effort grades of 3. A visual representation of the correlation for 6th grade in the form of a scatterplot can be found in Figure 10.

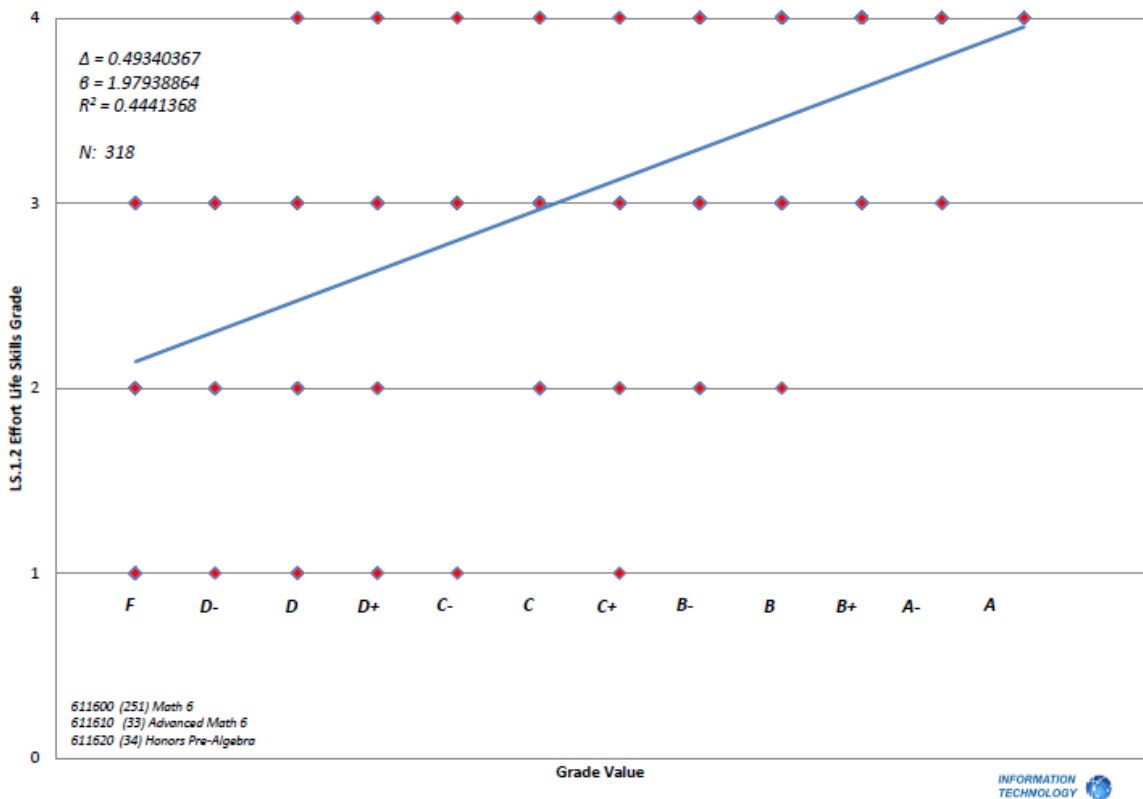


Figure 10. 2011-2012 grade 6 math grade and effort grade correlation analysis scatterplot

The lack of consistency among the correlations between math grades and effort grades among the three grades was not surprising as the majority of the work done in the

change process surrounding Grading for Learning was focused on the knowledge grade. Much staff development time was spent on the rationale of why certain factors would or would no longer be calculated in the knowledge grade. Less time was spent on establishing inter-rater reliability that would lead to consistency regarding the life skills grades. Anecdotally, where the highest correlation coefficient exists, grade 7, is the team of math teachers who also worked together very closely in terms of planning and assessment.

Measure 3: Relationship between teacher and student perceptions of effort

To see if students' perception of their own effort was matched with their teachers', an analysis of students' self report of effort and self-regulatory questions from the survey given by the University of Minnesota's college readiness program, Ramp-Up to Readiness™ were compared to the teacher effort grade. That analysis produced an R^2 of .500, a moderate positive relationship between student and teacher perceptions of effort.

Students on average reported higher rates of effort through their self-report of questions 22, 25 and 27 on the survey. A visual representation of the students' self-report via the mean response on the CAREI survey plotted against the teacher report of effort via the effort grade can be seen in Figure 11 on the next page.

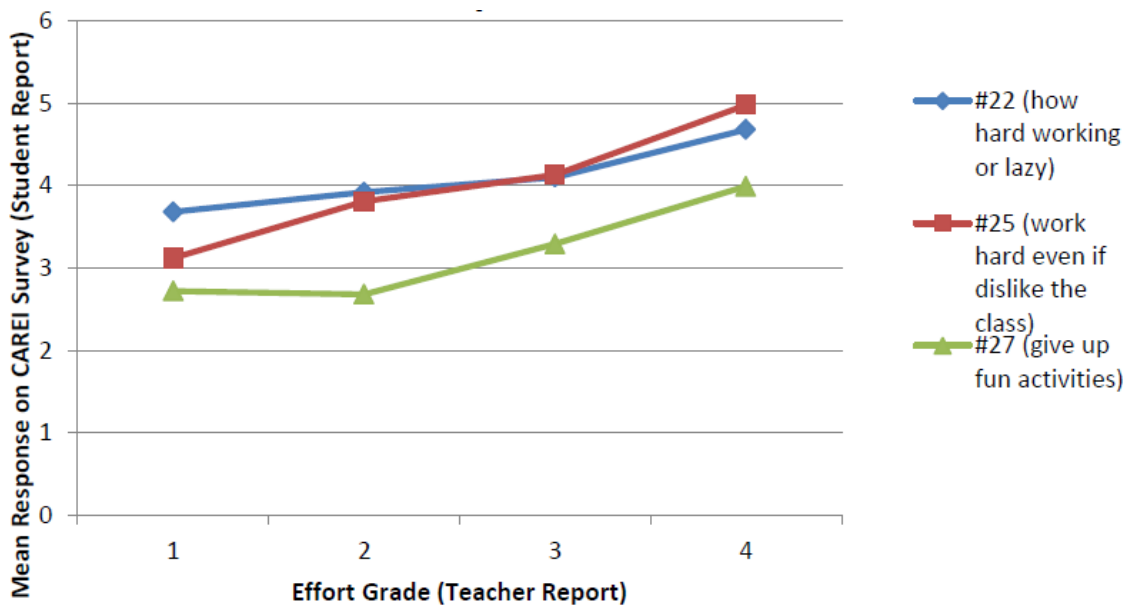


Figure 11: *Relationship between teacher and student perceptions of effort*

The discrepancy in student perception of effort versus the teacher perception of effort was reiterated by one student during the group interview process. She said, “Sometimes the teachers don't think the same as you.” This quotation illustrates the larger message that while the life skills grades are based on a rubric, they are still highly subjective.

Measure 4: Student perceptions of separating the knowledge and life skills grade

Student perceptions of the separation of the two grades as measured by a student survey was a key indicator in analyzing whether or not the goal of the initiative was achieved. A high response rate to the survey of 74.6%, 726 out of 973 eligible students took the survey as represented in these findings. Most students indicated that they perceived the way grades were given as fair (75.2% somewhat agreed, agreed or strongly agreed). Most also reported that they liked the separation of knowledge and life skills grades (73.84% somewhat agreed, agreed, or strongly agreed). There was contradictory

evidence on this measure, however, as many students also reported that they thought a final grade should include effort and behavior along with what is taught in class (78% somewhat agreed, agreed or strongly agreed).

When a similar question was asked in group interviews, the majority of students again agreed that it was a good idea to have two grades. One 6th grade student cited support of the two grades when she said, “Knowledge grades shows that you are paying attention. Life skills show that you are a responsible student, I guess.” It was clear that she understood the intent of the two grades.

Even given the discrepancy in the student response of what they felt should be included in a grade, when asked specifically about knowledge grades, 80.9% of students had some level of reported agreement that the knowledge grades that they received were usually good measures of what they have learned. Furthermore, 83.6% of students had some level of agreement that they felt the knowledge grade was made up of enough information (tests, quizzes, projects, assignments) to accurately represent what they know.

Students also reported that the life skills behavior grade was a good measure of the behavior they displayed in class (80.9% somewhat agreed, agreed, or strongly agreed). Also they agreed that the life skills effort grade was a good measure of the amount of effort they put into their classes (82.4% somewhat agreed, agreed, or strongly agreed).

In the group interviews of over 35 students in total, students broadly corroborated the data from the survey, identifying how well a student really knows a subject as the main factor in determining a knowledge grade. Responses to what a knowledge grade means and the frequency for each response were:

1. How well you really know the subject	11
2. What you need to work on	4
3. Reaffirms confidence	2
4. Level of personal responsibility	2
5. How well you listen in class	1
6. Progress	1
7. If I am going to get grounded	1
8. I don't know	1

Students said that studying more for quizzes was the main thing that changed in their behavior as a result of receiving a knowledge grade. Their responses about what behavior they changed as a result of a knowledge grade with the number of students who mentioned each were:

1. Study more for quizzes/tests	9
2. Spent more time on homework	6
3. Reviewed notes/work	5
4. Got help from a teacher	4
5. Paid better attention in class	3
6. Cut out lazy habits	2
7. Stayed after school to do homework	2
8. Kept up current study habits	2
9. Stepped it up	2
10. Didn't try as hard, already understood	2

11. Tried to get missing assignments in 1

Student comments during the group interviews provide an indication of the degree of influence that students think various factors have in determining life skills grades. The factors are listed in the declining order of frequency with which students cited them:

- | | |
|-------------------------------------|---|
| 1. Behavior – how you act | 7 |
| 2. Respectful to others and teacher | 4 |
| 3. Not a big deal | 4 |
| 4. Timeliness | 2 |
| 5. If you are participating | 1 |
| 6. “1” means talking | 1 |
| 7. How teacher thinks you are doing | 1 |
| 8. How prepared you are | 1 |

Measure 5: Student perceptions of grade calculation changes

Students overwhelmingly agreed that allowing them to retake tests was a positive change. Over 90% of students had some degree of agreement, with 52.9% of students strongly agreeing that students should be allowed to retest. They also agreed that their grade should not be lowered for turning in work late (75.9% somewhat agreed, agreed, or strongly agreed). Students were divided on the reform of not recording any score below 50%. Students agreed less enthusiastically with this practice at a rate of only 64.5%. A majority of those respondents (48.5%) fell somewhat evenly at 24.4% somewhat agreeing and 24.1% agreeing that students who completed a test or assignment should not receive

a score lower than 50% in order to insure that the student has a chance of passing the course.

While there was a technical problem with the item that evaluated student perceptions in changes in the use of extra credit as outlined below, it can be concluded that students were also divided in their perceptions of the value of including extra credit in the final grade. Less than half of the student respondents agreed extra credit should not be used (43.4% somewhat agreed, agreed, or strongly agreed).

There was an error in the response categories of the online survey for the question about the use of extra credit. In the entry of the survey questions into the electronic survey, the response “Strongly Agree” was accidentally entered into the position where strongly disagree was entered for all other questions on the survey.

In this question, respondents saw the following:

Students should not be allowed to earn extra credit in order to raise their grade in a course.

- Strongly Agree
- Disagree
- Somewhat Disagree
- Somewhat Agree
- Agree
- Strongly Agree

In the case of all respondent groups, students, staff or parents, there were responses recorded for the first instance of Strongly Agree and no responses recorded for the second instance, the position they were used to seeing the option of Strongly Disagree

in throughout the rest of the survey. This error made it impossible for a survey respondent to Strongly Disagree with the statement.

The response in Question 11 of “Strongly Agree” will be considered in these findings as “Strongly Agree” for the reasons outlined below; however, there is no way to verify the intent of every respondent, so these data should be viewed with caution.

- No responses from any of the respondents in the student, staff or parent surveys were entered for the second “Strongly Agree,” indicating that they may have felt the first occurrence was intended as “Strongly Agree”
- Extra credit was ranked 10/10 of factors students thought should be the most important in determining a final grade, and 4/11 of factors students thought should be least important in determining a final grade, indicating that students perhaps did not feel that extra credit is an important part or should be considered in grading
- Extra credit was ranked 10/10 with 0% responding of factors teachers thought should be most important when determining a grade, and 1/11 with 75% of teachers saying it should be considered the least important when determining a final grade
- Responses from staff would be expected to support not allowing extra credit, given their similar responses to the other grading practices that distort the achievement grade

Given the rationale stated above to consider the mistaken placement of the strongly agree category, the rates of response are calculated as “Strongly Agree” although, as noted above, the data should be viewed with a strong note of caution.

When students were asked to rank the most important factors in determining a final grade, they ranked ten factors in the following order:

1. Scores on tests
2. Scores on quizzes
3. Grades on essays and papers
4. Effort
5. Grades on homework assignments
6. Completion of homework assignments
7. Classroom behavior
8. Attendance
9. Participation in class
10. Extra credit

Conversely, they ranked the following factors as least important:

1. Attendance
2. Participation in class
3. Classroom behavior
4. Extra credit
5. Grades on group projects
6. Grades on homework assignments

7. Effort
8. Completion of homework assignments
9. Scores on quizzes
10. Grades on essays and papers
11. Scores on tests

Measure 6: Teacher perceptions of separating the knowledge and life skills grades

With a fairly high response rate at 68.8%, 44 out of 64 eligible staff responded to the survey as reported in these findings. Most staff reported that they felt that they perceive the way grades are given at the school as fair (80% somewhat agreed, agreed or strongly agreed). While most felt that giving a knowledge grade and a life skills grade was a good reform, their degrees of support for the reform varied (18.2% somewhat agreed, 25% agreed and 36.4% strongly agreed).

In general, staff felt the knowledge grade and the life skills grades were good measures of student mastery of content and mastery of life skills. Most felt that the knowledge grades they assigned were usually good measures of what the student had learned in class (84.1% somewhat agreed, agreed or strongly agreed). Staff reported some degree of agreement at similarly higher levels regarding their feelings that the life skills were good measures of effort (84.1% somewhat agreed, agreed or strongly agreed). Higher levels of agreement among staff regarding the behavior life skills grade were reported (91.1% somewhat agreed, agreed or strongly agreed).

Interviews and questionnaires broadly confirmed these findings. When the question of separating the grades was posed again in the interviews and questionnaires,

they overwhelmingly stated they thought it was a good idea to separate the knowledge and life skills grade with all but one of the 18 respondent saying “yes.” A number of quotes from the staff interviews and questionnaires provide nuanced data on the staff views of the changes. The perspective of both a guidance counselor and classroom teacher further illustrated staff perceptions on the separation of the two grades. The guidance counselor stated,

Knowledge grades give parents an idea about how their child is performing academically (based on standards at a particular time). Life skills give parents insight into additional factors that may influence (or tell a story about) a student’s knowledge grade. As a counselor, I rely on life skills [grades] to give me the complete picture of a child’s academic skills. Life skills [grades] tell a story that cannot be told with just a knowledge grade. A knowledge grade of a B with Life skills scores at 1s and 2s tells a completely different picture than a knowledge grade of a B with life skills scores at threes and fours.

The point of the two grades being useful in understanding the whole picture of a student illustrates that both pieces of information were important to the teachers and that when they were separated, teachers may have a better understanding of a student’s need.

To further illustrate this point, an 8th grade language arts teacher said,

I absolutely think it is a good idea to report the knowledge grade separate from the life skills grades. I have found over the past few years that I am more accurately reporting students’ knowledge of the grade level standards. Also, I am better able to diagnose the students’ weakness when looking at each skill set separately, even though often times there is a direct correlation between the two grades.

One teacher explained her perception of the importance of the life skills grade in this way:

I believe assigning a life skills grade is important. When I am asked to provide a life skills grade, it really gives me an opportunity to think about each and every student in my class and reflect upon their areas of strength

and areas where they need additional support. Keeping in mind that, for the most part, life skills grades are subjective, I find it useful information. If I have a student that is struggling in the area of timeliness, effort or behavior, I can check to see how his or her life skills are in other content areas. This information helps when having conversations with students about their strengths and areas of interest. I also use the life skills grade as a prompt for helping the students set SMART academic goals.

Of the 18 respondents to the teacher interview/questionnaire, 11 indicated that they had observed positive changes as a result, 6 indicated that they had observed negative changes, and one was neutral. Staff cited changes in study skills and behavior as the most positive change of the reforms in grading.

Again, several quotations provide nuanced information. One teacher expressed this change in terms of where he saw students placing their focus in their engagement in learning.

The main change I've seen in student behavior/attitudes is their attention to the content that is actually important to their knowledge grade and that which is not. For example, if a project was assigned where students were asked to design a portfolio cover, in the past, students may have spent an inordinate amount of time on the design of their cover without considering how much of their grade this part accounted for. Now students are more critical in their thinking about what is actually being evaluated when it comes to their knowledge grade.

The most frequently cited negative observation about Grading for Learning by the teachers was that as a result of homework only being weighted at 10% in the knowledge grade, students had a reason to rationalize not doing it because it did not dramatically impact their final grade. One school counselor agreed that while this could present an initial problem for some students, they eventually learned the connection of homework to learning in their final grade.

Because of the age of children I work with, it is only natural for students to want to rationalize reasons for not completing homework, and since practice work is weighted at 10% when rationalization are needed, students often jump to not completing homework because it is only weighted 10% and doesn't impact their grade that much. However, when questioned further, most students can identify that that argument is truly just a rationalization and will identify that practice work is for the most part vital to their learning and scoring well on assessments.

Teacher perceptions supported the separation of the two grades into one that communicates mastery of content--the knowledge grade--and one that communicates mastery of life skills.

Measure 7: Teacher perceptions of grade calculation changes

Staff were noticeably less supportive of technical changes in the ways that grades were calculated than the students. Staff were broadly supportive of allowing students to retake in order to raise their score (70.4% reported a degree of agreement). Students agreed with this practice at over 90%, with 52.9% strongly agreeing. In the case of the staff, they only strongly agreed at 13.6%. It could be assumed that staff might not have agreed as strongly with this practice because it does require more work for the teacher in terms of recreating a test for the student to take. They were notably less supportive of not lowering scores for work turned in late. Just over half of the teacher respondents somewhat agreed, agreed or strongly agreed at a total of 56.8%. Staff were evenly divided over the merits of not giving students a score lower than 50% for any completed test or assignment (50% reported some degree of agreement and some degree of disagreement).

Problems with the survey make it difficult to summarize staff perceptions of students not being allowed to earn extra credit to raise grades. The staff agreed that

students should not be allowed to earn extra credit to raise their scores (70.5% somewhat agreed, agreed, or strongly agreed). The data from the question on the survey asking about extra credit faced the same unfortunate error as the student survey. During the creation of the online survey the “Strongly Agree” option was mistakenly entered into the “Strongly Disagree” location and the “Strongly Agree” location on the options for response. This meant respondents did not have the option to “Strongly Disagree.” The responses are being used as Strongly Agree given the rationale cited for the use in measure 5.

Staff identified the following inputs as the most important through rank ordering the options presented in question 17 of the teacher survey in determining a student’s final grade:

1. Scores on tests
2. Scores on quizzes
3. Grades on essays and papers
4. Grades on homework assignments
5. Participation in class
6. Completion of homework assignments
7. Effort
8. Attendance

Staff identified the following inputs as least important through rank ordering the options presented in question 18 of the teacher survey in determining a student’s final course grade:

1. Extra credit
2. Attendance
3. Classroom behavior
4. Grades on group projects and participation in class
5. Effort
6. Grades on homework assignments
7. Completion of homework assignments and grades on essays and papers
8. Scores on quizzes
9. Scores on tests

Summary Section

The charts below summarize data collected for this evaluation of changes in grading practices at a middle school.

Measures of Statistical Correlation

Table 6 on the following page summarizes the R^2 correlation between data points on three of the seven measures and ranks them in order of their significance with the most significant first.

Table 11: *Summary of indicators of statistical correlation*

Indicator	Sample	Result
Correlation between knowledge grade and state standardized test scores	Grade 6 N=318	$R^2 = .625$
Correlation between knowledge grade and teacher effort grade	Grade 7 N=274	$R^2 = .609$
Correlation between knowledge grade and state standardized test scores	Grade 7 N=274	$R^2 = .603$
Correlation between knowledge grade and state standardized test scores	Grade 8 N=279	$R^2 = .583$
Correlation between student perception of effort from CAREI survey to teacher perception of effort from effort grade	Grade 8 N=247	$R^2 = .500$
Correlation between knowledge grade and teacher effort grade	Grade 6 N= 318	$R^2 = .444$
Correlation between knowledge grade and teacher effort grade	Grade 8 N= 279	$R^2 = .328$

Measure of Student Perceptions (Surveys)

Table 12 on the following page summarizes 12 questions from the student survey that were used as indicators to two of the measures in the evaluation of the intervention. General statements of the questions from the survey are followed by the sample size and the result of the percentage of students who responded in some form of agreement: somewhat agree, agree or strongly agree.

Table 12: *Summary of survey of student perceptions*

Indicator	Sample	Result
Should have opportunity to retake a test to improve grade	All students N=726	90.1%
Knowledge grade is made up of enough information to represent what they know	All students N=726	83.6%
Life skills grade is a good measure of effort they put into classes	All students N=726	82.4%
Knowledge grade is a good measure of what they have learned	All students N=726	80.9%
Life skills grade is a good measure of their behavior in class	All students N=726	80.9%
Grades should include effort and behavior in addition to what they have learned	All students N=726	78.1%
Grades should not be lowered for turning in work late	All students N=726	76%
The way grades are given at this school is fair	All students N=726	75.2%
Like the separation of knowledge grade and life skills grade	All students N=726	73.8%
Grades should be based only on what has been taught and what a student has learned	All students N=726	72.5%
Students should not receive below 50% if test or assignment is complete	All students N=726	64.5%
Extra credit should not be allowed	All students N=726	48.4%

Measure of Staff Perceptions (Surveys)

Table 13 on the following page summarizes 13 questions from the staff survey that were used as indicators to two of the measures in the evaluation of the intervention. General statements of the questions from the surveys are followed by the sample size and the result of the percentage of staff who responded in some form of agreement: somewhat agree, agree or strongly agree.

Table 13: *Summary of survey of teacher perceptions*

Indicator	Sample	Result
Life skills grade given is a good measure of student behavior in class	Staff N=44	91.1%
Knowledge grade is made up of enough information to represent what students know	Staff N=44	86.4%
Believe students understand what is measured by knowledge grade	Staff N=44	86.4%
Life skills grade is a good measure of effort students put into classes	Staff N=44	84.1%
Knowledge grade is a good measure of what students have learned	Staff N=44	84.1%
The way grades are given at this school is fair	Staff N=44	80.2%
Like the separation of knowledge grade and life skills	Staff N=44	79.5%
Grades should be based only on what has been taught and what a student has learned	Staff N=44	77.3%
Extra credit should not be allowed	Staff N=44	70.5%
Should have opportunity to retake a test to improve grade	Staff N=44	70.4%
Grades should not be lowered for turning in work late	Staff N=44	56.8%
Students should not receive below 50% if test or assignment is complete	Staff N=44	45.5%
Grades should include effort and behavior in addition to what they have learned	Staff N=44	33.6%

Measure of Student Perceptions (Group Interviews)

Student group interview data did not lend itself to charts or detailed comparisons, but major themes of the interviews can be highlighted:

- Students saw the use of two grades as useful
- Students reported that the course grades were “pretty accurate” or “accurate” at capturing and reporting what they knew and could do

- Students did not see the life skills grade as being as important as the knowledge grade
- Students saw knowledge grades as signals about their own knowledge of the subject
- Students report that changes regarding what they did in school were largely based on their knowledge grade
- Students saw the life skills grades as subjective with their accuracy dependent upon the teacher giving the grade

Summary of Measure of Staff Perceptions (Interviews and Questionnaires)

Staff interview and questionnaire data also did not lend itself to charts or detailed comparisons, but major themes can be highlighted:

- All but one staff member felt it was a good idea to give students two separate grades
- Staff reported the positives of giving two grades was the ability to communicate academic performance and the “other story” or factors that may influence the knowledge grade
- Staff reported observing both positive and negative changes in student behavior as a result of the intervention. The most cited positive behavior change was students developing better study skills. The most cited negative change was students rationalizing not doing homework because of its low weight in the knowledge grade

- The most frequently reported consideration when assigning a knowledge grade was if their assessments measured what was taught

Because multiple methods were utilized for this evaluation, it is not possible to combine the measures for a single conclusion regarding the changes in grading. The multiple measures and their indicators inform the following conclusions:

- Changes in the practices of calculating the knowledge grade appeared to increase the correlation of the knowledge grade and state standardized test scores
- There is a moderate correlation between the knowledge grades and life skills grades
- Student and teacher perceptions of the life skills effort grade are not highly correlated, although they both report that they are a good measure

Chapter 5

Discussion

Three research questions guided this mixed methods evaluation of how changes in grading practices in a middle school affect the perceptions of students and staff about grading and the degree of alignment between grades and test scores and other measures. The discussion provides observations about the change related to the key questions and also offers implications for practice and suggestions for future research. The study's three research questions were:

1. How did staff and students perceive separation of the traditional single grade into two grades?
2. Did staff and students find the changes in grading practices useful? If so, how? If not, why not?
3. Did the reform in grading and related practices promote closer alignment between:
 - a. Subjective teacher evaluation of student knowledge as measured by the knowledge grade and an objective measure of student knowledge as indicated by scores on a standardized achievement test
 - b. Teacher evaluation of student knowledge as measured by the knowledge grade and teacher evaluation of student effort as measured by the life skills effort grade; and

- c. Teacher perception of student effort as measured by the life skills effort grade and student perception of effort as measured by a survey?

Observations on Question 1- Awarding a Knowledge Grade and a Life Skills Grade

A clear majority, 73.8% of students and 79.5% of staff, agreed that they liked the separation of the knowledge grade and the life skills grade. Given that a major premise of this study is that students, staff, and parents should have a shared understanding of grades, staff and student agreement on the issue is significant.

The benefits of awarding two grades appear to be a clearer understanding of what students know and how they behave, a deliberate focus on instruction and assessment on the part of the teacher, and a focus on students learning academic content rather than amassing points. Separating the knowledge grade from the factors that made up the life skills grade appeared to eliminate the white noise that was previously masking student understanding of content. With this information reported separately, students, parents, and educators were able to easily identify not only if a student had mastered the content knowledge, but also if the student's life skills, specifically behavior, effort and timeliness, were supporting or detracting from his or her learning of content. Staff interviews provided only support for the separation of the knowledge and life skills grades. Many staff members, such as the science teacher quoted below, argued that a major benefit of this reform was that it provided students and families with more accurate information:

I think it's a great idea to give both types of grades because it's a more sound way of reporting. Students' knowledge grades are no longer

inflated and better reflect what state wide assessments show. Also, it is easier to diagnose problems and offer solutions to parents wanting better performance out of their children.

Similarly, a 7th grade language arts teacher said that a benefit of the reform was that, “...it paints a larger picture of them [students] as a learner and communicates more than just knowledge.”

A deliberate focus on instruction and assessment on the part of the teacher was an unintended outcome of the separation of the two grades. Whereas before the reforms of Grading for Learning teachers could reward students for effort and good study habits through practices such as offering extra credit, including points for participation, or calculating daily homework as a more significant part of the grade, now the knowledge grade limited the content of the grade to an assessment of what the student had actually learned. This caused teachers to think more deliberately about what their goal for student learning was at each point in the curriculum, prompted them to craft assessments to measure achievement of that goal, and pushed them to plan instruction to insure student understanding of content. Teacher collaboration increased in the form of more and better curriculum planning, writing of common assessments, and analyzing student outcomes for the need for remediation and ways to improve improved instruction.

Prior to Grading for Learning, students, parents, and even some teachers focused their discussions of student progress on the student’s ability or willingness to amass points to achieve a desired grade, instead of a student’s mastery of curricular content. Following the adoption of Grading for Learning, parent-teacher conferences became longer, more students accessed teachers for help outside of class, and numbers in the after

school academy exploded when students and parents realized that the knowledge grade now represented understanding of content that had to be mastered to be promoted and eventually pass state graduation tests.

There were drawbacks to the separation of the grades as well. This approach to grading is more work for the teacher, and the life skills grade was generally not viewed as important as the knowledge grade. Not only was there an additional grade to determine, but it consisted of three separate marks -- one for effort, one for behavior, and one for timeliness. While a consistent approach to knowledge grade calculation across the school was a positive for students and parents, the transparency increased pressure on teachers with students and parents' desire of teachers to score and post grades in a similar and timely manner.

As the student interviews revealed, the life skills grade was seen as less important than the knowledge grade. One student who was interviewed plainly said, "Life skills grades are not as important." Some saw the life skills grades as, one student said, "subjective." Another student described the drawbacks to the subjectivity of life skills grades this way, "Some teachers take a first impression on a student, and if it is a bad one they will just consider them not a good student."

When asked more directly in the group interviews if they felt it was a good idea to give both a knowledge grade and a life skills grade, there were mixed responses from the students. One 7th grader said that the use of two grades, "...lets your parents know how you are doing in school." More specifically, a 6th grade student explained that, "The knowledge grade shows that you are paying attention and

life skills show that you are being a responsible student I guess.” Another student said that, “Kids know if they are good or not, so I don’t think we need life skills grade[s].” While the interview data provided only a limited perspective of the students’ more specific thoughts on the separation of the two grades, they do illustrate a significant drawback to the separation of the two grades and the need for further understanding. A much more thorough study of student perceptions of the individual grades could inform better use of the grades in the future.

The reform of installing two grades could have been implemented differently to enhance the positives and reduce the negatives described above. The benefit of providing a clearer understanding of what students know and how they behave as a result of the two grades could have been enhanced if parent input in crafting of the reform had been sought. In addition, more explanation for parents of the rationale for the separation of the two grades, clearer information on how the knowledge grade was calculated and why, and more background on the importance and criteria of the life skills grade might have proved valuable.

The drawback of increased teacher work load that this evaluation identified could be addressed through increased and dedicated time for staff collaboration for the development and analysis of common assessments. Another alternative approach would be to bring staff together to develop a shared understanding of the life skills rubric indicators and inter-rater reliability in assigning life skills grades.

Additionally, the drawback of the life skills grade being considered less important than the knowledge grade could potentially be addressed through systemic work of

developing and promoting a shared understanding of the grades among teachers, students, and parents. This could be done through the use of student self-rating on the life skills grade rubric, two-way conversations between student and teachers about the scores, and ongoing analysis of the scores over time.

If these drawbacks to the effectiveness of Grading for Learning were effectively addressed, reporting student progress using two separate measures of content knowledge and life skills might have even greater benefits than those identified through this study.

Observations on Question 2 - Technical Changes in Grading Practices

The ultimate original objective of Grading for Learning was to more accurately represent what a student knew about the math content. Four technical changes were made so that the knowledge grade would better represent knowledge of content.

Those changes were: (1) allowing students the opportunity to retest; (2) not lowering a grade for work turned in late; (3) using 50% as the lowest point of calculation for assigned tasks; and (4) not allowing for extra credit. The study provides evidence that these technical changes helped to remove the white noise that could have been affecting a final grade based on student behavior, timeliness, or unequal distribution of the percentage grading system.

While the technical changes that were made were done to more accurately measure, calculate, and communicate a student's understanding of content, an unintended benefit was that teachers felt that even though they no longer had the ability to inflate a grade, they were now not only accurate in their communication, they were providing students with multiple opportunities to demonstrate learning

when it happened. Allowing for retesting and not lowering scores for late work are examples of this. Implementing these technical changes also contributed to a team-based approach to assessing student work and addressing student needs.

Observations on Question 3 - Promoting Alignment

A key premise to this study is that alignment matters. Students need to know that the ways they are judged by standardized tests are connected to the ways they are judged by their teachers in their classes. This study showed that alignment is possible, though not always easy, and most importantly that alignment does not necessarily mean agreement. A knowledge grade can be very different from an effort grade or a standardized test score like the results of the Minnesota Comprehensive Assessments. However, the study clearly demonstrates that where misalignment exists, we need to explain that information to students and parents and seek their feedback, along with the feedback of the other educators who work with that student. We need to help young people and the adults in their lives to make sense of these discrepant indicators.

In addition, a shared understanding of the indicators and the discrepancy among them can be an effective way to begin to develop strategies and solutions for mitigating this misalignment. Aligning indicators of knowledge grades, life skills grades and standardized assessments is a starting point. The communication of these and other key indicators in an aligned manner is needed for students and parents to understand where students are and what they need to do to be college and career ready by the end of high school.

Observations on Grading and Instruction

This reform started as a result of two incredible teachers wanting to fix grades and fix students, but it became a way to examine curriculum, assessment, and instruction. As such, many of the benefits of the reform cannot be captured within a focused evaluation such as this one. Change brings about more change. Changes in grading led to changes in assessment. Changes in assessment led to changes in instruction. As students and parents questioned the new grading practices, teachers began to collaborate out of a sense of needing to feel secure about what they were doing. As they did so, they began to discuss the types of assignments and assessments that they entered into their grade books. Teachers found that the number and diversity of assessments must increase if a student's grade was going to be based primarily on tests, quizzes, and projects. Teachers of the same subjects began to discuss the types of assessments they were giving. Those conversations eventually led to teachers writing common assessments together.

The quality of those assessments also became an important ongoing topic of conversation. A move to rubric-based assessments allowed teachers to feel more confident that they knew what they intended to assess and that students knew the expectations of performance. In order to communicate performance expectations to students, assessments needed to be designed based on standards before instruction took place. This intentional crafting of assessments encouraged teachers to design their instruction so that students were exposed to what was going to be assessed. It is

worth noting that this was largely the hope of Outcome Based Education in the early 1990s in Minnesota. Twenty years later it appears that it may be happening.

In summary, while this evaluation does not directly report on change in instruction, what began as a strategy for improving grading became a strategy for improving instruction. One sixth grade social studies teacher captured this unanticipated but very positive outcome well when she noted that for her the biggest outcome of Grading for Learning was this: “I have had to question my teaching practices, evaluate what’s important and alter my assessment practices.” As educators are faced with so many opportunities and avenues to improve student learning, a clear and focused change in grading proved to be the lever that began a change in collaboration, assessment and instruction.

One of the math teachers who initiated the grading change process in the first place made the same point in greater detail in response to the often asked question, “*Are grades improving?*”:

No, not yet, but we have done a better job of more accurately reporting student grades. When we first started investigating grading reform our frustrations were with student behavior. We wanted to fix kids. What we found out was that our instruction needed fixing. We studied the standards and the test specs. We did away with sections in the text that didn't address the standards so we could spend more time on those that did. We wrote our tests based on the standards and then designed our lessons based on the tests. We were trying to make sure our lessons prepared students for our tests, which in turn prepared them for the state assessments. I just know I feel better doing it this way because I feel like I am doing the right thing. The grades that I give are based on what the student knows and has learned and what behaviors they have shown during the process.

This educator's comment highlights a final conclusion of this study: *teacher leadership matters*. The teacher as a professional must be at the core of a change initiative of this sort. Grading changes can be mandated, but shared understanding cannot. Grading for Learning was teacher-initiated and largely teacher-led. The teachers at this school were committed to student learning. They were faced with a challenge to their past practices and were offered opportunities for growth by their colleagues. The shared understanding of grading that was created by their own professional study and dialog led to better grading, better assessments, and better instruction through professional collaboration.

Future Research

This evaluation of a change in grading practices at a large middle school in rural Minnesota can be used not only for improvement of the grading program at the study site, but also to suggest directions for future efforts to strengthen grading in middle schools:

1. This study was unable to report on parent perspective or perceptions due to extremely low response rates of parents. Further studies should attempt to understand parents' opinions and perspectives on grading in general and before and after changes to approaches in grading are made in their child's school.
2. The focus of the professional development activities at the study site emphasized calculation of the knowledge grade. In the future, studies should

examine the implications of focusing instead or equally on the calculation of the life skills grade.

3. Future studies should include sub-group analyses by race and gender. The scope of this study was focused on the overall impact of the changes in grading. However, a closer look at the effects of the grading system on sub-populations could allow for educators to see if the alignment of survey responses, grades, standardized test scores, and teacher and student perceptions of effort are different given students' race or gender. If the alignment of these factors is different, further qualitative and descriptive study of the sub populations of students could be done to see if there are biases that exist in grading based on race or gender.
4. To see if classroom grades were aligned with an objective measure, analyses performed for this study examined the degree of correlation between the knowledge grades in mathematics assigned by teachers and scores on the state standardized math test (Minnesota Comprehensive Assessment). For future studies to be replicated outside of Minnesota and to the extent possible to compare schools, classroom math grades should be compared to a nationally normed test like the EXPLORE.
5. This study attempted to see if there was alignment between the knowledge grade and the life skills grade. The outcome of the correlation found moderate alignment, but this is clearly an issue of great importance as the importance of non-cognitive skills increases in schools, postsecondary education, and the

workplace. Further studies of the links between measures of knowledge and life/academic schools should be a high priority for both researchers and practitioners in the years ahead.

6. This study focused on one middle school as an evaluation of the school's initiative of a significant change in grading practices. In future research, more study sites should be included to enable more powerful statistical conclusions and comparisons across different types of school structure and culture. The inclusion of high schools would also be beneficial and allow for a study as it would make it possible to follow students over time to see if the changes they encountered in grading in middle school have any impact on their high school performance.

Bibliography

- ACT. (2005). *Are High School Grades Inflated?* Retrieved November 9, 2010, from ACT Resesarch and Policy Briefs: <http://www.act.org/research/policymakers/briefs.html#2005>
- Airasian, P. W. (1988). Symbolic Validation: The Case of State-Mandated, High Stakes Testing. *Educational Evaluation and Policy Analysis* , 10 (4), 301-313.
- Allen, J. D., & Lambating, J. (2001). *Validity and Reliability in Assessment and Grading; Perspectives of Preservice and Inservice Teachers and Teacher Education Professors*. Seattle: Paper presented at the Annual Meeting of the American Educational Research Association.
- Austin, S., & McCann, R. (1992). "Here's Another Arbitrary Grade for your Collection": A Statewide Study of Grading Policies. *Paper presenetd at the annual meeting of the AERA* (pp. 1-41). San Fransisco: Research for Better Schools, Inc.
- Baron, P. A. (2000). *Consequential Validity for High Schol Grades: What is the Meaning of Grades for Senders and Receivers?* New Orleans: Paper presented at the Annual Meeting of the American Educational Research Association.
- Brookhart, S. M. (2003). Developing Measurement Theory for Classroom Assessment Purposes and Uses. *Educational Measurement: Issues and Practice* , 5-12.
- Brookhart, S. M. (1993). Teachers' Grading Practices: Meaning and Values. *Journal of Educational Measurement* , 30 (2), 123-142.
- Brookhart, S. M. (1994). Teachers' Grading: Practice and Theory. *Applied Measurement in Education*, 279-301.
- Brookhart, S. M., Walsh, J. M., & Zientarski, W. A. (2006). The Dynamics of Motivation and Effort for Classroom Asessments in Middle School Science and Social Studies. *Applied Measurement in Education* , 19 (2), 151-184.
- Conley, D. T. (2000). *Who Is Proficient: The Relationship between Proficiency Scores and Grades*. New Orleans: Paper presented at the Annual Meeting of the American Educational Research Association.
- Cross, L. H., & Frary, R. B. (1999). Hodgepodge Grading: Endorsed by Students and Teachers Alike. *Applied Measurement in Education* , 12 (1), 53-72.
- Docan, T. N. (2006). Positive and Negative Incentives in the Classroom: An Analysis of Grading Systems and Student Motivation. *Journal of the Scholarship of Teaching and Learning* , 21-40.

- Duncan, C. R., & Noonan, B. (2007). Factors Affecting Teachers' Grading and Assessment Practices. *The Alberta Journal of Educational Research* , 53 (1), 1-21.
- Erickson, J. (2010, March). Grading Practices: The Third Rail. *Principal Leadership* , pp. 22-26.
- Friedman, S. J., & Frisbie, D. A. (1995). The Influence of Report Cards on the Validity of Grades Reported to Parents. *Educational and Psychological Measurement* , 5-26.
- Guskey, T. R. (2004, November). Are Zeros Your Ultimate Weapon? *Principal Leadership* , pp. 31-35.
- Guskey, T. R. (2009). *Bound by Tradition: Teachers' View of Crucial Grading and Reporting Issues*. San Fransisco: Paper presented at the Annual Meeting of the American Educational Research Association.
- Guskey, T. R. (2000, December). Grading Policies that Work Against Standards...and How to Fix Them. *NASSP Bulletin* , pp. 20-27.
- Guskey, T. R. (2007, Spring). Multiple Sources of Evidence: An Analysis of Stakeholders' Perceptions of Various Indicators of Student Learning. *Educational Measurement: Issues and Practice* , pp. 19-27.
- Guskey, T. R. (2002). *Perspectives on Grading and Reporting: Differences Among Teachers, Students, and Parents*. New Orleans: Paper presented at the Annual Meeting of the American Educational Research Association.
- Guskey, T. R., & Bailey, J. M. (2010). *Developing Standards-Based Report Cards*. Thousand Oaks: Corwin.
- Howley, A., Kusimo, P. S., & Parrott, L. (1999). *Grading and the Ethos of Effort*. EDRS.
- Kvale, S. (1996). *Inter-Views: An Introduction to Qualitative Research Interviewing*. Thousad Oaks, CA: Sage Publications, Inc.
- Labaree, D. F. (1997). *How to Succeed in School Without Really Learning: The Credentials Race in American Education*. New Haven: Yale University Press.
- Limor, G.-E., Shavit, Y., & Yaish, M. (2009). Curricular Choice: A Test of the Rational Choice Model of Education. *European Sociological Review* , 26(4), 447-463.
- Marzano, R. J. (2000). *Transforming Classroom Grading*. Alexandria: ASCD.
- McDonough, P. M. (1997). *Choosing Colleges: How Social Class and Schools Structure Opportunity*. Albany: State University of New York Press.
- McMillan, J. H. (2001, Spring). Secondary Teachers' Classroom Assessment and Grading Practices. *Educational Measurement: Issues and Practice*, pp. 20-32.

- McMillan, J. H. (2003). Understanding and Improving Teachers' Classroom Assessment Decision Making: Implications for Theory and Practice. *Educational Measurement: Issues and Practice*, 34-43.
- McTighe, J., & O'Connor, K. (2005, November). Seven Practices for Effective Learning. *Educational Leadership* , pp. 10-17.
- Morgan, S. L. (2005). *On the Edge of Commitment*. Stanford , CA: Stanford University Press.
- O'Connor, K. (2007). *A Repair Kit for Grading 15 Fixes for Broken Grades*. Portland: Educational Testing Service.
- Pilcher, J. K. (1994). The Value-Driven Meaning of Grades. *Educational Assessment* , 2 (1), 69-88.
- Powell, A. G., Farrar, E., & Cohen, D. K. (1985). *The Shopping Mall High School: Winners and Losers in the Educational Marketplace*. Boston: Houghton Mifflin Company.
- Randall, J., & Engelhard, G. (2010, March). Examining the grading practices of teachers. *Teaching and Teacher Education* , pp. 1372-1380.
- Ravid, R. (2005). *Practical Statistics for Educators*. (3rd Ed.) New York: University Press of America.
- Reeves, D. B. (2004, December). The Case Against Zero. *Phi Delta Kappan* , pp. 324-325.
- Savitz-Romer, M., & Bouffard, S. M. (2012). *Ready, Willing and Able: A Developmental Approach to College Access and Success*. Cambridge, MA : Harvard Education Press.
- Teddlie, C., & Tashakkori, A. (2009). *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences* . Thousand Oaks, CA: Sage Publications.
- Tyack, D., & Tobin, W. (1994). The "Grammar" of Schooling: Why Has It Been So Hard To Change? *American Educational Research Journal*, 31 (3), 453-479.
- Vatterott, C. (2009). *Rethinking Homework Best Practices That Support Diverse Needs*. Alexandria: ASCD.
- Wilburn, K. T., & Felps, B. C. (1983). *Do Pupil Grading Methods Effect Middle School Students' Achievement: A Comparison of Criterion-Referenced Versus Norm-Referenced Evaluation*.

Willingham, W. W., Pollack, J. M., & Lewis, C. (2002). Grades and Test Scores: Accounting for Observed Differences. *Journal of Educational Measurement* , 39 (1), 1-37.

Winger, T. (2009, November). Grading What Matters. *Educational Leadership* , pp. 73-75.

Wormeli, R. (2006, Summer). Accountability: Teaching Through Assessment and Feedback, Not Grading. *American Secondary Education* , pp. 14-26.

Appendix A: Grading for Learning Overview for Students and Parents

Grading for Learning – Ellis Middle School – Parent/Student Overview

Rationale

Ellis Middle School recognizes that grading, homework, make-up opportunities, report cards, and reporting practices in general should be consistent, accurate, fair, meaningful, research-based, and connected to state and national standards.

Students will receive a **knowledge grade** and a **life skills grade**.

Knowledge grades will reflect achievement. Other important student characteristics such as behavior and effort should be reported separately. Therefore an additional score will be reported at Ellis Middle School. The Life Skills grade will reflect behavior, effort and timeliness.

Grading

Knowledge Grade will consist of two categories:

Practice – This is commonly known as homework and is weighted at 10%. Practice includes class work and assignments.

Assessments – This is commonly known as tests, quizzes and projects. Assessments are weighted at 90%.

Life Skills Grade will be assigned based on the following rubric:

KEY TO TERMS	Acceptable		Unacceptable	
	4	3	2	1
Behavior	Consistently contributes to a productive learning atmosphere by following expectations	Usually contributes to a productive learning atmosphere by following expectations	Inconsistently contributes to a productive learning atmosphere by following expectations	Rarely contributes to a productive learning atmosphere by following expectations
Effort	Consistently demonstrates initiative and self-direction	Usually demonstrates initiative and self-direction	Inconsistently demonstrates initiative and self-direction	Rarely demonstrates initiative and self-direction
Timeliness	Consistently meets established deadlines for all assigned tasks	Usually meets established deadlines for all assigned tasks	Inconsistently meets established deadlines for all assigned tasks	Rarely meets established deadlines for all assigned tasks

Grading Parameters

- Parents are informed regularly (at least 4 times per year) of their child's achievement and progress. Additionally, a parent portal will be available to parents so they can access, via computer, their child's grades and scores. Parents will not receive paper copies of mid-quarter reports mailed to the home. These mid-quarter grades can be accessed online via the portal.
- Students are taught how to utilize their student portal account and will be provided a regular opportunity to access their grades.
- The professional judgment of teachers should be respected.
- Generally, work submitted late will not receive a reduced score. Teachers will provide support for the learner. However, in order to receive credit, all work must be turned in by the end of the quarter.
- Missing work will be reflected in the grade book as a zero.
- Individual assessment and practice scores below 50% will be adjusted to 50% in the grade book with the earned score noted as a comment in the grade book.
- Students will have the opportunity to re-take Tests if all of their practice work is completed prior to the date of the original test. Re-takes must be completed within the quarter and will cover the same content but may be in a different format. Students will be expected to initiate a re-test. Re-assessment beyond tests in the cases of quizzes and projects are left to the discretion of the academic department. Individual department guidelines can be found in each teacher's syllabus available on their teacher web page.
- Teachers will not offer extra credit in order to improve a knowledge grade.
- Academic dishonesty will be addressed with behavioral consequences. Teachers will reassess and determine the actual level of achievement when initiated by the student.
- Scores on group projects should have a minimal impact on a student's overall grade. Group projects should have an individual score included as part of the assessment.
- Teachers will provide clear explanation of all assessments and practice within the description section of Infinite Campus.

- Students who have incomplete assessments will be expected to stay for Assessment make-up sessions at the end of each quarter. If a student does not complete an assessment they will receive an “I” Incomplete for a report card grade. At the end of the academic year any student with an Incomplete will be *required* to attend summer school.

Dates that mid-quarter progress grades will be posted and available on the portal are as follows:

October 7, 2011

December 16, 2011

February 24, 2012

May 4, 2012

Appendix B: CAREI Analysis of Relationship Between Teacher and Student Perceptions of Effort

Correlations Between Key Variables (Number of students used to calculate the correlation is in parentheses)

	Q3 Effort Grade	Q3 Math Grade	Effort & SR Strategies
Third Quarter Effort Grade		.698 (283)	.500 (247)
Third Quarter Math Grade			.399 (234)
Effort & Self-Regulatory Strategies (see below)			
Q25. How often do you work hard in class and on assignments, even when you don't like the class?	.478 (245)	.416 (232)	.748 (666)
Q27. How often do you give up TV, video games, Facebook, texting, or time with friends to study for a test or do an assignment for school?	.376 (247)	.279 (234)	.670 (675)
Q22. If, on a scale of 1-6, a one is the laziest student you know, and a six is the hardest working, what number are you?	.367 (247)	.346 (234)	.700 (673)
Q19. I've learned to manage my time and meet deadlines.	.365 (246)	.366 (233)	.680 (671)
Q8. I take notes in class to help me learn.	.350 (247)	.209 (234)	.657 (671)
Q26. How often do you ask for help when you don't understand something in class?	.294 (246)	.205 (233)	.669 (671)
Q7. I am willing to take hard classes even though I may get a lower grade.	.224 (247)	.242 (234)	.385 (672)
Q18. Succeeding in school depends more on how hard you work than on how naturally smart you are.	.205 (247)	.108 (234)	.468 (673)

Items in the Effort & Self-Regulatory Strategies factor:

Q25. How often do you work hard in class and on assignments, even when you don't like the class?

Q22. If, on a scale of 1-6, a one is the laziest student you know, and a six is the hardest working, what number are you?

Q19. I've learned to manage my time and meet deadlines.

Q27. How often do you give up TV, video games, Facebook, texting, or time with friends to study for a test or do an assignment for school?

Q26. How often do you ask for help when you don't understand something in class?

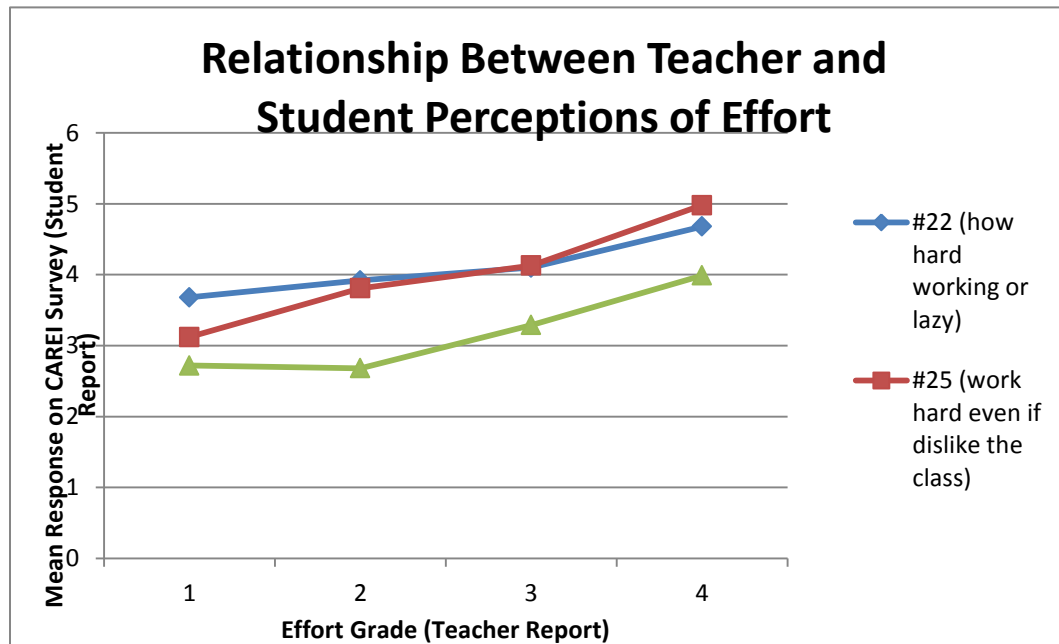
Q8. I take notes in class to help me learn.

Q17. My teachers have high expectations for me.

Q18. Succeeding in school depends more on how hard you work than on how naturally smart you are.

Key Findings

- ❖ Student perceptions of their own effort, as measured in the Effort & Self-Regulatory Strategies dimension, explained 25% of the variation in teacher perceptions of student effort.
- ❖ The strongest single-item predictor on the CAREI survey of teacher perceptions of effort was item #25. The degree to which students responded that they work hard on assignments for even for classes they didn't like explained 23% of the variation in teacher perceptions of student effort, and 17% of the variation in third quarter math grades.
- ❖ Eighth grade student perceptions of how hard-working or lazy they were (CAREI survey item #22) explained 13% of the variation in effort grades and 12% of the variation in third quarter math grades.
- ❖ Math teacher perceptions of student effort explained 49% of the variation in the grades that their students earned in the third quarter, or vice-versa



- ❖ Overall there is a moderate, positive relationship between student and teacher perceptions of student effort.
- ❖ Students on average reported higher self-reports of effort (#22, #25, and #27) compared to their teacher reports of effort (Effort Grade). For a student who received an effort grade of 1 from the teacher report, they self-reported to have a mean effort response of three, which relates to an “average” amount of self-reported effort.
- ❖ Asking students how often they work hard on school work even when they don’t like the class might be a slightly better measure of how hard-working or lazy they are than asking them directly, especially at the extremes of effort grades

Appendix C: Student Survey

SCREEN 1

Dear Ellis Student:

Attached you will find a survey regarding grading in general and Grading for Learning, the grading system used at Ellis Middle School. This is a voluntary survey. The purpose of this survey is to gain feedback from students on their understandings and perceptions of grading and reporting in general and at Ellis Middle School as it relates to your experiences as a student.

The following screen is an informed consent form that I would like you to read while your teacher reads it aloud and click yes or no indicate if you agree to participate in the study. This form confirms that you understand the purpose of the study and are willing to participate in it.

- Proceed to informed consent
-

SCREEN 2

STUDENT ASSENT FORM

Grading for Learning Study

A study of Grading for Learning, the grading and reporting system used at Ellis Middle School is being conducted to see if *Grading for Learning* is accurately communicating to students and parents what you know in terms of your learning in classes like math and language arts. Content knowledge is represented in your Knowledge Grades. Academic behaviors are represented in your Life Skills Grade.

Procedure:

If you agree to participate, you will be asked to respond to the survey in Infinite Campus that will take approximately 10-15 minutes to complete. Please understand that there are no right or wrong answers and that different students will respond differently to the questions on the survey.

Risks and Benefits of Being in the Study:

There are no known risks and no direct benefits of participation in the study. However, participations will help us better understand grading at Ellis and middle school grading in general.

Confidentiality:

The records of this study will be kept private and will be stored securely so that only researchers will have access to the records. Survey responses will be anonymous. A summary report of how all students in the school responded will be shared with the school district, but students' individual responses are confidential and will not be shared with your teachers, parents or anyone else.

Voluntary Nature of the Study:

Participation is voluntary and there will be no negative consequences for choosing not to participate. Your decision whether or not to participate will not affect current or future grades or relations with the school. If you decide to participate, you are free to not answer specific questions and to withdraw at any time.

Contacts and Questions:

You may ask questions you have now. If you have questions later, you are encouraged to contact Ms. Berglund at school or at katie.berglund@austin.k12.mn.us or (507)-440-1501. If you have any questions or concerns regarding the study and would like to talk to someone other than the researcher(s), you are encouraged to contact the Fairview Research Helpline at telephone number 612-672-7692 or toll free at 866-508-6961. You may also contact this office in writing or in person at Fairview Research Administration, 2433 Energy Park Drive, St. Paul 55108.

Statement of Consent:

I have read the above information and agree to participate in the study.

- I agree to take the survey
 - I do not want to take the survey
-

SCREEN 3

Please respond to the following according to the degree to which you agree or disagree with each statement.

Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
-------------------	----------	-------------------	----------------	-------	----------------

1. When I get a grade in a class at Ellis Middle School, I understand what the grade says about my performance in school
2. When I get a grade in a class at Ellis Middle School, it tells me something important about how I am doing in school
3. The same grade – such as a “C” – can mean very different things in different teachers’ classes at Ellis Middle School
4. The way that grades are given at Ellis Middle School is fair
5. It is possible to get a good grade at Ellis Middle School even if a student hasn’t really learned what was taught in the class
6. A student’s grade should be based only on how well he or she has learned what was taught in the class
7. A student’s grade should include things like effort and behavior along with how well the student has learned what was taught in the class
8. A student who completes a test or assignment should not receive a score lower than 50% on that test or assignment in order to ensure that the student always has a chance of passing the course
9. Students should always have the opportunity to retake a test in order to raise their score
10. Students should not be allowed to earn extra credit in order to raise their grade in a course
11. A student’s grade should not be lowered for turning in work late
12. I like the approach of giving students separate Knowledge and Life Skills grades that is now being used at Ellis Middle School.
13. The Knowledge Grades that I have received at Ellis Middle School are usually good measures of how well I learned what was taught in a class

14. The Knowledge Grade at Ellis Middle School is made up of enough information (tests, quizzes, projects, assignments) to accurately represent what I know.
15. The Life Skills Grades that I have received in my classes at Ellis Middle School are a good measure of the amount of effort I put into those classes
16. The Life Skills Grades that I have received in my classes at Ellis Middle School are a good measure of my behavior in those classes
-

SCREEN 4

17. Which of the following factors should be most important in determining a student's final grade in a course (Please choose 3):
- Attendance
 - Classroom behavior
 - Completion of homework assignments
 - Effort
 - Extra credit
 - Grades on homework assignments
 - Grades on essays and papers
 - Grades on group projects
 - Participation in class
 - Scores on quizzes
 - Scores on tests
18. Which of the following factors should be least important in determining a student's final grade in a course (Please choose 3):
- Attendance
 - Classroom behavior
 - Completion of homework assignments
 - Effort
 - Extra credit
 - Grades on homework assignments
 - Grades on essays and papers
 - Grades on group projects
 - Participation in class
 - Scores on quizzes
 - Scores on tests
-

SCREEN 5

Demographic items

19. What grade are you in?

- a. 6th
- b. 7th
- c. 8th

20. Are you:

- a. Male
- b. Female

21. Please indicate your race/ethnicity (Mark all that apply):

- a. African
- b. Black or African American
- c. American Indian
- d. Asian or Asian American
- e. Hispanic or Latino
- f. White
- g. Other _____

Appendix D: Protocol for Student Group Interviews

Introduction to the Student Interviews

Thank you for agreeing to participate in this interview. I am interested in this subject so I can understand and share with others the understanding and perceptions of grading, specifically, Grading for Learning at Ellis Middle School as they relate to students, parents, and teachers so that students and parents are well informed about their child's education.

I have prepared an informed consent form that I would like you to sign if you agree to participate in the study. This form confirms that you understand the purpose of the study and are willing to participate in it. I will be recording our conversation to be sure I can quote your comments accurately and in detail.

Interview/Focus Group Questions

1. What things do you think teachers consider when they give a course grade here at Ellis Middle School?
2. How accurately do you think the course grades you receive at Ellis Middle School capture what you know and can do?
3. What do you think is the purpose of giving grades?
4. When you are given a Knowledge Grade in a course here at Ellis Middle School, what do you think that grade means?
5. Have you ever changed what you do in school based upon a Knowledge Grade that you received? If so, please tell me about that.
6. When you are given a Life Skills Grade in a course here at Ellis Middle School, what do you think that grade means?
7. Have you ever changed what you do in school based upon a Life Skills Grade that you received? If so, please tell me about that.
8. Do you think it is a good idea to give students both Knowledge Grades and Life Skills Grades? Why?

Additional questions if time is available:

9. Why do you think we give students here at Ellis Middle School the opportunity to retake tests in order to raise their scores? What do you think of this approach?
10. Why do you think we do not allow students here at Ellis Middle School to earn extra credit to raise their course grades? What do you think of this approach?
11. Why do you think we do not lower students' grades for turning work in late here at Ellis Middle School? What do you think of this approach?

Appendix E: Staff Survey

SCREEN 1

Dear Ellis Teacher:

Attached you will find a survey regarding grading in general and Grading for Learning, the grading system used at Ellis Middle School. This is a voluntary survey. The purpose of this survey is to gain feedback from teachers on their understandings and perceptions of grading and reporting in general and at Ellis Middle School as it relates to their students.

The following screen is an informed consent form that I would like you to read and click yes or no sign if you agree to participate in the study. This form confirms that you understand the purpose of the study and are willing to participate in it.

- Proceed to informed consent
-

SCREEN 2

TEACHER CONSENT FORM

Grading for Learning Study

A study of Grading for Learning, the grading and reporting system used at Ellis Middle School is being conducted this spring at Ellis Middle School. The purpose of this research is to examine if *Grading for Learning* is more accurately communicating to students and parents what students know in terms of content (Knowledge Grade) and what students are doing in terms of their academic behaviors (Life Skills Grade). This study is being conducted by Katie Berglund, the principal of Ellis Middle School and doctoral student at the University of Minnesota.

Procedure:

In this study, you will be able complete a 21 item survey about your understanding and perceptions of the grading system, perceived students' learning of content, effort, motivation and persistence, and knowledge of your students' academic progress. The survey will be administered via Infinite Campus. It will take approximately 15 minutes to complete. A small sample of students participating in the survey will be randomly selected to participate in a group interview approximately two weeks after the survey. Parents and randomly selected teachers of the students participating in group interviews will also be invited to participate in a group interview.

Risks and Benefits of Being in the Study:

There are no known risks and no direct benefits of participation in the study. However, participation will inform grading at Ellis and middle school grading in general.

Confidentiality:

The records of this study will be kept private and will be stored securely so that only researchers will have access to the records. Survey responses will be anonymous. A summary report of responses will be shared with the school district, but individual responses are confidential and will not be shared with anyone.

Voluntary Nature of the Study:

Participation is voluntary and there will be no negative consequences for choosing not to participate. Your decision whether or not to participate will not affect current or future relations with the school. If you participate, you are free to not answer specific questions and to withdraw at any time.

Contacts and Questions;

If you have any questions, you are encouraged to contact Katie Berglund at katie.berglund@austin.k12.mn.us or (507)-440-1501 or David Krenz, Superintendent of Schools. If you have any questions or concerns regarding the study and would like to talk to someone other than the researcher(s), you are encouraged to contact the Fairview Research Helpline at telephone number 612-672-7692 or toll free at 866-508-6961. You may also contact this office in writing or in person at Fairview Research Administration, 2433 Energy Park Drive, St. Paul 55108.

Statement of Consent:

I have read the above information and agree to participate in the study.

- I agree to take the survey
 - I do not want to take the survey
-

SCREEN 3

Please respond to the following according to the degree to which you agree or disagree with each statement

Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
-------------------	----------	-------------------	----------------	-------	----------------

1. When I give a grade for a class I have taught at Ellis Middle School, I am confident that my students understand what the grade says about their performance in school
2. I believe that the course grades I give at Ellis Middle School provide my students and their parents with important information about the student's performance in school
3. The same grade – such as a “C” – can mean very different things in different teachers' classes at Ellis Middle School
4. The way that grades are given at Ellis Middle School is fair
5. It is possible to get a good grade at Ellis Middle School even if a student hasn't really learned what was taught in the class
6. A student's grade should be based only on how well he or she has learned what was taught in the class
7. A student's grade should include things like effort and behavior along with how well the student has learned what was taught in the class
8. A student who completes a test or assignment should not receive a score lower than 50% on that test or assignment in order to ensure that the student always has a chance of passing the course
9. Students should always have the opportunity to retake a test in order to raise their score
10. Students should not be allowed to earn extra credit in order to raise their grade in a course
11. A student's grade should not be lowered for turning in work late
12. I like the approach of giving students separate Knowledge and Life Skills grades that is now being used at Ellis Middle School.

13. The Knowledge Grades that I assign at Ellis Middle School are usually good measures of how well a student has learned what was taught in a class
 14. The Knowledge Grade at Ellis Middle School is made up of enough information (tests, quizzes, projects, assignments) to accurately represent what students know and are able to do
 15. I believe that students understand what is being measured by the Knowledge Grades I assign at Ellis Middle School
 16. I believe that parents understand what is being measured by the Knowledge Grades I assign at Ellis Middle School
 17. The Life Skills Grades that I assign at Ellis Middle School are a good measure of the amount of effort that students put into my classes
 18. The Life Skills Grades that I assign in my classes at Ellis Middle School are a good measure of students' behavior in my classes
 19. I believe that students understand what is being measured by the Life Skills Grades I assign at Ellis Middle School
 20. I believe that parents understand what is being measured by the Life Skills Grades I assign at Ellis Middle School
-

SCREEN 4

21. Which of the following factors should be most important in determining a student's final grade in a course (Please choose 3):
 - Attendance
 - Classroom behavior
 - Completion of homework assignments
 - Effort
 - Extra credit
 - Grades on homework assignments
 - Grades on essays and papers
 - Grades on group projects
 - Participation in class
 - Scores on quizzes
 - Scores on tests

22. Which of the following factors should be least important in determining a student's final grade in a course (Please choose 3):

- Attendance
 - Classroom behavior
 - Completion of homework assignments
 - Effort
 - Extra credit
 - Grades on homework assignments
 - Grades on essays and papers
 - Grades on group projects
 - Participation in class
 - Scores on quizzes
 - Scores on tests
-

SCREEN 5

23. What grade to you primarily teach?

- a. 6th
- b. 7th
- c. 8th

24. Are you:

- a. Male
- b. Female

25. Please indicate your years of experience in teaching.

- a. 1-5 years
- b. 5- 10 years
- c. 10-15 years
- d. 15- 20 years
- e. 20-30 years
- f. 30+ years

Appendix F: Protocol for Individual Staff Interviews/Questionnaires

Introduction to the Staff Interviews

Thank you for agreeing to participate in this interview. I am interested in this subject so I can understand and share with others the understanding and perceptions of grading, specifically, Grading for Learning at Ellis Middle School as they relate to students, parents, and teachers so that students and parents are well informed about their child's education.

I have prepared an informed consent form that I would like you to sign if you agree to participate in the study. This form confirms that you understand the purpose of the study and are willing to participate in it. I will be recording our conversation to be sure I can quote your comments accurately and in detail.

Interview/Focus Group Questions

1. What do you think is the purpose of giving grades?
2. What factors do you consider when you assign a Knowledge Grade for a course here at Ellis Middle School?
3. When students receive a Knowledge Grade in your classes here at Ellis Middle School, how do you think they interpret that grade? What conclusions do you think they draw from it?
4. When parents and guardians receive their child's Knowledge Grade for one of your classes here at Ellis Middle School, how do you think they interpret that grade? What conclusions do you think they draw from it?
5. What factors do you consider when you assign a Life Skills Grade for a course here at Ellis Middle School?
6. When students receive a Life Skills Grade in your classes, how do you think they interpret that grade? What conclusions do you think they draw from it?
7. When parents and guardians receive their child's Life Skills Grade for one of your classes, how do you think they interpret that grade? What conclusions do you think they draw from it?
8. Do you think it is a good idea to give students both Knowledge Grades and Life Skills Grades? Why?
9. Have you observed any changes in students' behavior and/or attitudes that you believe were significantly influenced by the adoption of Grading for Learning at Ellis? If so, what are they?

Additional Questions if time is available: (if time is available ask the following):

10. What do you think students at Ellis think is the reason we allow them to retake tests?
11. What do you think students at Ellis think is the reason we do not allow them to earn extra credit to raise their course grades?
12. What do you think students at Ellis think is the reason we do not lower their grades for turning work in late?

Appendix G: Results of Student and Staff Surveys

Question 2

Student: When I get a grade in a class at Ellis Middle School, I understand what the grade says about my performance in school

Staff: When I give a grade for a class I have taught at Ellis Middle School, I am confident that my students understand what the grade says about their performance in school

Q1	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n= 723)	2.2	0.8	4.1	19.0	44.7	28.6	.4
Staff (n= 42)	2.2	2.2	6.8	4.5	50.0	29.5	4.5

Question 3

Student: When I get a grade in a class at Ellis Middle School, it tells me something important about how I am doing in school

Staff: I believe that the course grades I give at Ellis Middle School provide my students and their parents with important information about the student's performance in school

Q3	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n= 720)	2.0	1.9	4.1	17.4	46.1	27.4	.8
Staff (n=42)	2.2	2.2	4.5	9.0	43.1	34.0	4.5

Question 4

Student/Staff: The same grade – such as a “C” – can mean very different things in different teachers’ classes at Ellis Middle School

Q4	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n= 715)	3.5	7.1	11.2	30.5	33.7	12.1	1.5
Staff (n=42)	2.2	31.8	11.	25	20.4	4.5	4.5

Question 5

Student/Staff: The way that grades are given at Ellis Middle School is fair

Q5	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n=719)	7.4	4.6	11.7	28.3	32.5	14.3	.9
Staff (n=42)	2.2	9.0	4.5	11.3	50	18.1	4.5

Question 6

Student/Staff: It is possible to get a good grade at Ellis Middle School even if a student hasn't really learned what was taught in the class

Q6	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n= 715)	9.0	18.8	16.2	28.1	18.0	8.1	1.5
Staff (n=42)	18.1	40.9	13.6	15.9	4.5	2.2	4.5

Question 7

Student/Staff: A student's grade should be based only on how well he or she has learned what was taught in the class

Q7	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n=712)	4.5	7.5	13.5	26.8	30.	15.2	1.9
Staff (n=42)	4.5	4.5	9.0	18.1	36.3	22.7	4.5

Question 8

Student/Staff: A student's grade should include things like effort and behavior along with how well the student has learned what was taught in the class

Q8	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n=708)	5.2	5.9	9.3	25.6	32.5	18.8	2.2
Staff (n=42)	18.1	34.0	15.9	11.3	9.0	6.8	4.5

Question 9

Student/Staff: A student who completes a test or assignment should not receive a score lower than 50% on that test or assignment in order to ensure that the student always has a chance of passing the course

Q9	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n=714)	9.0	10.7	14.0	24.3	24.1	15.9	1.6
Staff (n=42)	18.1	18.1	13.6	13.6	15.9	15.9	4.5

Question 10

Student/Staff: Students should always have the opportunity to retake a test in order to raise their score

Q10	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n=712)	1.5	2.0	4.4	11.9	25.2	52.8	1.9
Staff (n=41)	6.8	11.3	4.5	20.4	36.2	13.6	6.8

Question 11

Student/Staff: Students should not be allowed to earn extra credit in order to raise their grade in a course

Q11	Strongly Agree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n=720)	25.4	36.0	14.7	10.6	12.2	0	.8
Staff (n=42)	27.2	18.1	6.8	4.5	38.6	0	4.5

Question 12

Student/Staff: A student's grade should not be lowered for turning in work late

Q12	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n=715)	5.6	7.7	9.2	18.4	25.9	31.5	1.5
Staff (n=42)	4.5	20.4	13.5	9.0	34.0	13.6	4.5

Question 13

Student/Staff: I like the approach of giving students separate Knowledge and Life Skills grades that is now being used at Ellis Middle School

Q13	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n=707)	7.5	6.2	9.7	31.9	30.7	11.1	2.6
Staff (n=42)	6.8	0	9.0	18.1	25	36.3	4.5

Question 14

Student: The Knowledge Grades that I have received at Ellis Middle School are usually good measures of how well I learned what was taught in a class

Staff: The Knowledge Grades that I assign at Ellis Middle School are usually good measures of how well a student has learned what was taught in a class

Q14	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n=712)	2.3	3.8	10.3	27.9	39.3	13.6	1.9
Staff (n=42)	2.2	4.5	4.5	4.5	56.8	22.7	4.5

Question 15

Student: The Knowledge Grade at Ellis Middle School is made up of enough information (tests, quizzes, projects, assignments) to accurately represent what I know.

Staff: The Knowledge Grade at Ellis Middle School is made up of enough information (tests, quizzes, projects, assignments) to accurately represent what students know and are able to do

Q15	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Student (n=715)	3.3	3.5	7.9	23.9	42.2	17.3	1.5
Staff (n=42)	2.2	2.2	4.5	9.0	54.5	22.7	4.5

Question 16 – Student; Question 18 - Staff

Student: The Life Skills Grades that I have received in my classes at Ellis Middle School are a good measure of my behavior in those classes

Staff: The Life Skills Grades that I assign in my classes at Ellis Middle School are a good measure of students’ behavior in my classes

Q16/18	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Q16 -Student (n=718)	4.9	3.9	12.4	29.3	42.0	19.0	1.0
Q 18 -Staff (n=42)	0	4.5	4.5	20.4	56.8	15.9	4.5

Question 17 – Student; Question 18 – Staff

Student: The Life Skills Grades that I have received in my classes at Ellis Middle School are a good measure of the amount of effort I put into those classes

Staff: The Life Skills Grades that I assign at Ellis Middle School are a good measure of the amount of effort that students put into my classes

Q17/18	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree	Not Answered
Q 17- Student (n=717)	4.82	5.10	9.09	26.58	43.94	18.46	1.24
Q 18- Staff (n=42)	0	4.55	11.36	25	43.18	15.91	4.55

Question 18 – Students; Question 22 – Staff

Students/Staff: Which of the following factors should be the most important in determining a student’s final grade in a course in your opinion? (Please choose 3):

Factor	Students (n=717)	Staff (n=42)
Attendance	23.7% (n=174)	2.2% (n=1)
Classroom behavior	28.1% (n=204)	0%
Completion of homework assignments	35.4% (n=257)	11.3% (n=5)
Effort	50.1% (n=364)	9.0% (n=4)
Extra Credit	23.2% (n=169)	0%
Grades on homework assignments	42.4% (n=308)	38.6% (n=17)
Grades on essays and papers	3 - 53.9% (n=392)	3 - 52.2% (n=23)
Participation in class	23.5% (n=171)	25% (n=11)
Scores on quizzes	2 - 55.9% (n=406)	2 - 81.8% (n=36)
Scores on tests	1 - 78.6% (571)	1 - 93.1% (n=41)
Not Answered	.9% (n=7)	4.5% (n=2)

Question 19 – Students; Question 23 – Staff

Students/Staff: Which of the following factors should be least important in determining a student’s final grade in a course in your opinion? (Please choose 3):

Factor	Students (n=709)	Staff (n=42)
Attendance	1 – 53.7% (n=390)	2 - 50% (n=22)
Classroom behavior	3 - 39.8% (n=289)	3 – 47.7% (n=21)
Completion of homework assignments	20.6% (n=150)	6.8% (n=3)
Effort	23.6% (n=172)	20.4% (n=9)
Extra Credit	39.1% (n=284)	1 – 75% (n=33)
Grades on homework assignments	24.7% (n=180)	18.1% (n=8)
Grades on essays and papers	13.7% (n=100)	6.2% (n=3)
Grades on group projects	28.8% (195)	22.7% (n=10)
Participation in class	2 - 44.7% (n=325)	22.7% (n=10)
Scores on quizzes	15.1% (n=110)	4.5% (n=2)
Scores on tests	11.5% (n=84)	2.2% (n=1)
Not Answered	2.3% (n=17)	4.5% (n=2)

DEMOGRAPHICS

Students = 726 total Respondents out of 980

Grade

6	248	34.1%
7	238	32.7%
8	239	32.9%
Not answered	13	1.7%

Gender

Male	322	44.3%
Female	399	54.9%
Not answered	17	2.3%

Ethnicity

African	23	3.1%
Black or African American	47	6.4%
American Indian	19	2.6%
Asian or Asian American	36	4.9%
Hispanic or Latino	124	17.8%
White	507	69.8%
Other	36	4.9%
Not Answered	29	3.9%

STAFF = 44 Total Respondents out of 55

Grade Primarily Taught

6 th Grade	16	36.3%
7 th Grade	15	34.0%
8 th Grade	15	34.0%
Not Answered	4	9.0%

Gender

Male	15	34.0%
Female	25	56.8%
Not Answered	4	9.0%

Years of Experience in Teaching

1-5 years	2	4.5%
5-10 years	7	15.9%
10-15 years	9	20.4%
15-20 years	7	15.9%
20-30 years	13	29.5%

Appendix H: Unused Parent Survey

SCREEN 1

Dear Ellis Parent/Guardian:

Attached you will find a survey regarding grading in general and Grading for Learning, the grading system used at Ellis Middle School. This is a voluntary survey. The purpose of this survey is to gain feedback from parent/guardians on their understandings and perceptions of grading and reporting in general and at Ellis Middle School as it relates to your son/daughter.

The following screen is an informed consent form that I would like you to read and click yes or no if you agree to participate in the study. This form confirms that you understand the purpose of the study and are willing to participate in it.

- Proceed to informed consent
-

SCREEN 2

PARENT CONSENT FORM

Grading for Learning Study

A study of Grading for Learning, the grading and reporting system used at Ellis Middle School is being conducted this spring at Ellis Middle School. The purpose of this research is to examine if *Grading for Learning* is more accurately communicating to students and parents what students know in terms of content (Knowledge Grade) and what students are doing in terms of their academic behaviors (Life Skills Grade). This study is being conducted by Katie Berglund, the principal of Ellis Middle School and doctoral student at the University of Minnesota.

Procedure:

In this study, you will be able complete a 21- item survey about your understanding and perceptions of the grading system, perceived child's learning of content, effort, motivation and persistence, and knowledge of your child's academic progress. The survey will be administered via Parent Portal within Infinite Campus. It will take approximately 15 minutes to complete. A small sample of students participating in the survey will be randomly selected to participate in a group interview approximately two weeks after the survey. Parents of the students participating in group interviews will also be invited to participate in a group interview.

Risks and Benefits of Being in the Study:

There are no known risks and no direct benefits of participation in the study. However, participation will inform grading at Ellis and middle school grading in general.

Confidentiality:

The records of this study will be kept private and will be stored securely so that only researchers will have access to the records. Survey responses will be anonymous. A summary report of responses will be shared with the school district, but individual responses are confidential and will not be shared with anyone.

Voluntary Nature of the Study:

Participation is voluntary and there will be no negative consequences for choosing not to participate. A parent's or student's decision whether or not to participate will not affect current or future grades or relations with the school. If you participate, you are free to not answer specific questions and to withdraw at any time.

Contacts and Questions;

If you have any questions, you are encouraged to contact Katie Berglund at katie.berglund@austin.k12.mn.us or (507)-440-1501 or David Krenz, Superintendent of Schools. If you have any questions or concerns regarding the study and would like to talk to someone other than the researcher(s), you are encouraged to contact the Fairview Research Helpline at telephone number 612-672-7692 or toll free at 866-508-6961. You may also contact this office in writing or in person at Fairview Research Administration, 2433 Energy Park Drive, St. Paul 55108.

Statement of Consent:

I have read the above information and agree to participate in the study.

- I agree to take the survey
 - I do not want to take the survey
-

SCREEN 3

Please respond to the following according to the degree to which you agree or disagree with each statement.

Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
-------------------	----------	-------------------	----------------	-------	----------------

1. When my son or daughter gets a grade in a class at Ellis Middle School, I understand what the grade says about his or her performance in school
2. When my son or daughter gets a grade in a class at Ellis Middle School, it tells me something important about how he or she is doing in school
3. The same grade – such as a “C” – can mean very different things in different teachers’ classes at Ellis Middle School
4. The way that grades are given at Ellis Middle School is fair
5. It is possible to get a good grade at Ellis Middle School even if a student hasn’t really learned what was taught in the class
6. A student’s grade should be based only on how well he or she has learned what was taught in the class
7. A student’s grade should include things like effort and behavior along with how well the student has learned what was taught in the class
8. A student who completes a test or assignment should not receive a score lower than 50% on that test or assignment in order to ensure that the student always has a chance of passing the course
9. Students should always have the opportunity to retake a test in order to raise their score
10. Students should not be allowed to earn extra credit in order to raise their grade in a course
11. A student’s grade should not be lowered for turning in work late
12. I like the approach of giving students separate Knowledge and Life Skills grades that is now being used at Ellis Middle School.
13. The Knowledge Grades that my son or daughter has received at Ellis Middle School are usually good measures of what he or she knows and is able to do

14. The Knowledge Grade at Ellis Middle School is made up of enough information (tests, quizzes, projects, assignments) to accurately represent what my son or daughter knows and can do
 15. The Life Skills Grades that my son or daughter has received at Ellis Middle School are usually good measures of the amount of effort he or she puts into school
 16. The Life Skills Grades that my son or daughter has received at Ellis Middle School are usually good measures of his or her behavior at school
-

SCREEN 4

17. Which of the following factors should be most important in determining a student's final grade in a course in your opinion? (Please choose 3):
 - Attendance
 - Classroom behavior
 - Completion of homework assignments
 - Effort
 - Extra credit
 - Grades on homework assignments
 - Grades on essays and papers
 - Grades on group projects
 - Participation in class
 - Scores on quizzes
 - Scores on tests
 18. Which of the following factors should be least important in determining a student's final grade in a course in your opinion? (Please choose 3):
 - Attendance
 - Classroom behavior
 - Completion of homework assignments
 - Effort
 - Extra credit
 - Grades on homework assignments
 - Grades on essays and papers
 - Grades on group projects
 - Participation in class
 - Scores on quizzes
 - Scores on tests
-

SCREEN 5

19. What grade is your student in? (Check all that apply)

- a. 6th
- b. 7th
- c. 8th

20. Are you:

- a. Male
- b. Female

21. Please indicate your race/ethnicity (Mark all that apply):

- a. African
- b. Black or African American
- c. American Indian
- d. Asian or Asian American
- e. Hispanic or Latino
- f. White
- g. Other _____

Appendix I: Unused Template for Parent/Guardian Interviews

Introduction to the Parent/Guardian Interviews

Thank you for agreeing to participate in this interview. Your son or daughter has already participated in the study in the form of a survey, interview, or both. I am interested in this subject so I can understand and share with others the understanding and perceptions of grading, specifically, Grading for Learning at Ellis Middle School as they relate to students, parents, and teachers so that students and parents are well informed about their child's education.

I have prepared an informed consent form that I would like you to sign if you agree to participate in the study. This form confirms that you understand the purpose of the study and are willing to participate in it. I will be recording our conversation to be sure I can quote your comments accurately and in detail.

Interview/Focus Group Questions

1. What things do you think teachers consider when they give a course grade at Ellis Middle School?
2. How accurately do you think the course grades your son or daughter has received at Ellis Middle School capture what your son or daughter knows and can do?
3. What do you think is the purpose of giving grades?
4. When you receive your son's or daughter's Knowledge Grade for a course at Ellis Middle School, what do you think that grade means?
5. When you your son's or daughter's Life Skills Grade for a course here at Ellis Middle School, what do you think that grade means?
6. Do you think it is a good idea to give students both Knowledge Grades and Life Skills Grades? Why?

Additional questions if time is available:

7. Why do you think we give students at Ellis Middle School the opportunity to retake tests in order to raise their scores? What do you think of this approach?
8. Why do you think we do not allow students at Ellis Middle School to earn extra credit to raise their course grades? What do you think of this approach?
9. Why do you think we do not lower students' grades for turning work in late at Ellis Middle School? What do you think of this approach?