

Middle School Common Grading Expectations and Teacher Perceptions of Grading

Pedagogy

By

Jennifer Carson

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Education

Major: Educational Leadership

Under the Supervision of Dr. Tamara J. Williams, Ed.D.

Omaha, Nebraska

December, 2017

Supervisory Committee:

Dr. Tamara J. Williams, Ed.D.

Dr. Kay Keiser, Ed.D.

Dr. Jeanne Surface, Ed.D.

Dr. Elliott Ostler, Ed.D.

Dr. Phyllis Adcock, Ph.D.

Abstract

Middle School Common Grading Expectations and Teacher Perceptions of Grading
Pedagogy

Jennifer Carson, Ed.D.

University of Nebraska, 2017.

Advisor: Dr. Tamara J. Williams, Ed.D.

Grades in schools can be subjective in nature. Yet, grading remains a key piece in the American education system. Beliefs around grading practices are deeply held and can vary from teacher to teacher. This study aims to explore how teachers arrive at their own grading practices and to analyze teachers' perceptions of implementing common grading expectations at one middle school in the Midwest. The implementation was grounded in research surrounding ethical change leadership. In year two of implementation, certified staff responded to a survey of Likert scale and open-ended questions providing feedback on implementing common grading expectations. The medians of responses were analyzed and compared to subgroups of grade level taught, core teachers, exploratory teachers, other certificated staff, years of experience, content area, leadership experience, and experience teaching in more than one district. Then the open-ended qualitative responses were used to explain or justify the quantitative data. Overall findings indicated no significant differences in perceptions of the subgroups. However, qualitative data showed concerns with retakes and how authentic conversations and relationships are at the core of effective change in schools. Though this study is not intended to be replicated across other schools, it may serve as a learning opportunity when considering a change initiative in grading practices.

Acknowledgements

I would like to thank my committee for your professional knowledge and feedback along the way. Feeling comfortable asking questions and seeking your guidance was a big help. Dr. Williams, this would not have been possible without you. You knew when to help and when to let me figure it out on my own. The weekend texts, emails, and phone calls were greatly appreciated. You provided a perfect balance of pushing me to think, laughter, sarcasm, and expertise in the process.

To my other colleagues along the way, you all proved it takes a village. Dr. Darin Kelberlau, thank you for allowing me to continue my research through a job change. More importantly, thank you for being there on one of my most frustrating days. You dropped everything on a Saturday afternoon to come help me for a few hours with your statistical expertise. For that, I am grateful. For all my MPS colleagues that offered words of encouragement and gave their time to my research, thank you for being part of the journey.

Finally, there are no words to express my gratitude to my biggest supporter, TJ. Thank you for always backing me through my educational and professional pursuits. The extra work and burdens you take on to allow me to reach my goals is nothing short of amazing. This research is a reflection of both of our time and work. You made me laugh and offered “suggestions” when I needed it most. To my kids Zander, Maclin, Kinzie, and Tegan, thank you for sacrificing many weekends without me to allow me to work. I am excited to finally get back to normal family weekends.

Table of Contents

Chapter One: Introduction.....	1
Causes of Grading Variation.....	2
Conceptual Framework.....	9
Research Questions.....	13
Definitions of Terms.....	13
Limitations.....	16
Delimitations.....	17
Assumptions.....	17
Significance of the Study.....	17
Chapter Two: Research Literature Review.....	20
Why Similarity is Important.....	20
Grading Practices.....	27
Zeros.....	28
Grade inflation.....	29
Non-achievement Factors Included in Grades.....	30
Expand Assessment Options.....	32
Homework.....	34
Retakes.....	36

Grades as Rewards and Punishments.....	37
Minimum Grade.....	38
Initiatives Involve Change.....	39
Evidence of Change.....	48
Process of Initiating Change in Grading Expectations at the Research Site.....	52
Chapter Three: Methodology and Design.....	56
Study Design.....	56
Population.....	57
Instrumentation.....	57
Data Collection.....	58
Research Questions.....	59
Data Analysis.....	60
Chapter Four: Results.....	63
Research Question #1.....	64
Research Question #2.....	67
Research Question #3.....	74
Chapter Five: Conclusions.....	95
Conclusions.....	95

Recommendations for Further Research.....	100
Summary.....	101
References.....	103
Appendix.....	111

List of Multimedia Objects

Figure 1: Michael Fullan’s (2004) Framework for Leadership	9
Table 4.1 Overview of Median Frequencies by Teacher Type (Core, Exploratory, and Other) on Perceptions of Implementation Fidelity.....	67
Table 4.2 Kruskal-Wallis by Teacher Type (Core, Exploratory, and Other) on Perceptions of Implementation Fidelity.....	68
Table 4.3 Overview of Median Frequencies by Grade Level on Perceptions of Implementation Fidelity.....	69
Table 4.4 Kruskal-Wallis by Grade Level on Perceptions of Implementation Fidelity...	69
Table 4.5 Overview of Median Frequencies by Years of Experience on Perceptions of Implementation Fidelity	70
Table 4.6 Kruskal-Wallis by Years of Experience on Perceptions of Implementation Fidelity	70
Table 4.7 Overview of Median Frequencies by Years of Content on Perceptions of Implementation Fidelity	71
Table 4.8 Kruskal-Wallis by Content on Perceptions of Implementation Fidelity	71
Table 4.9 Mann-Whitney Test Comparing Teachers Who Have District Leadership Experience and Teachers Who Do Not Have District Leadership Experience on Perceptions of Implementation Fidelity	72

Table 4.10 Mann-Whitney Test Comparing Teachers Who Have Worked in Other Districts and Teachers Who Have Not Worked in Other Districts on Perceptions of Implementation Fidelity	73
Table 4.11 Overview of Median Frequencies by Teacher Type (Core, Exploratory, and Other) on Perceptions of Common Grading Expectations	75
Table 4.12 Kruskal-Wallis by Teacher Type (Core, Exploratory, and Other) on Perceptions of Common Grading Expectations	76
Table 4.13 Overview of Median Frequencies by Grade Level on Perceptions of Common Grading Expectations	78
Table 4.14 Kruskal-Wallis by Grade Level on Perceptions of Common Grading Expectations	79
Table 4.15 Overview of Median Frequencies by Years of Experience on Perceptions of Common Grading Expectations	81
Table 4.16 Kruskal-Wallis by Years of Experience on Perceptions of Common Grading Expectations	82
Table 4.17 Overview of Median Frequencies by Content on Perceptions of Common Grading Expectations	84
Table 4.18 Kruskal-Wallis by Content on Perceptions of Common Grading Expectations	85

Table 4.19 Mann-Whitney Test Comparing Teachers Who Have Held District Leadership
Positions and Those Who Have Not on Perceptions of Common Grading

Expectations87

Table 4.20 Mann-Whitney Test Comparing Teachers Who Have Worked In Other
Districts and Those Who Have Not on Perceptions of Common Grading

Expectations89

Chapter 1: Introduction

Interpreting a grade is like interpreting an abstract piece of art. A grade may be displayed as a percentage such as 95%. Or it may be displayed as a letter, such as A, B, C, D, or F, to represent some given percentage breakdown. Either way, it is left to be interpreted by the viewer. Each person that views it sees what is important to him and takes away whatever he chooses. One viewer may look and see every detail, analyze the way in which it was produced, and internalize the way it makes him feel. The next viewer may instead give it a glance, decide it is not individually important, and find no emotional tie. Grades are the art that are left to interpretation.

Grades are defined as a compilation of marks from multiple assignments and assessments that lead to a final judgement about students' achievement. Grades are a teacher's numeric or alpha evaluation of a student's demonstration of knowledge and are used to communicate a variety of information. Just as an artist chooses the medium and color scheme, teachers are left to make a professional judgment regarding what will be included in a grade. When individual teachers are making decisions specific to each classroom, students' grades can encompass many sub-categories. The problem is that when teachers try to report so many factors in a grade, they generally end up achieving none of them very well (Guskey, 2004).

Though grades can be subjective in nature, grading remains a key piece in our American education system. In fact, Lynn Olson (1995) observed that grades are "one of the most sacred traditions in American education...The truth is that grades have acquired an almost cult-like importance in American schools" (p. 24). Educators rely on grades to

help make key decisions about students. Grades hold value within schools as they are one of the factors used to determine a student's promotion, athletic eligibility, need for additional support, and qualification for graduation. Outside the K-12 education system, colleges, the armed forces, employers, and scholarship-funders rely on grades to make decisions about students that can open or close doors to their futures (Feldman, 2014).

The importance of grades and the implications of utilizing grades holds value in the United States K-12 education system. After exploring possible causes of grading variation, the literature review will investigate why similarity is important as well as the pros and cons of different grading practices used in schools today. The research on similarity, grading practices, implementing change initiatives, evidence of change, and Fullan's framework of ethical change leadership provides the structures for implementing the common grading expectations change initiative at the research site.

Causes of Grading Variation

According to Wiliam (2016), in the United States, inconsistencies in classroom expectations far outweigh inconsistent expectations at the building level. When looking at a building as a whole, most times individuals can agree on the big rocks surrounding grading practices. For example, all staff may agree grades need to represent student demonstration of knowledge. However, grading variation occurs based on how the statement is interpreted or put into practice in the classroom. It is dependent on individual teacher's beliefs and experiences. This causes variation from classroom to classroom. One teacher may believe "student demonstration of knowledge" consists of grades showing knowledge acquisition at specific times in a student's learning. If a

student earns a failing grade on an assignment, it tells a story about the student's knowledge at that particular point in time and should be set in stone. On the other hand, the teacher next door may believe "student demonstration of knowledge" has no time frame and allows multiple redo opportunities to demonstrate a higher level of understanding with the grade being replaced with each redo attempt. Both teachers believe they are meeting the overarching goal of using grades to represent student demonstration of knowledge but are providing unequal opportunities for how students reach the end goal of a final grade. Due to the conflict from vision to implementation, it generally does not matter which school a student attends but it does matter which classroom a student is in (William, 2016).

Another cause of variation stems from the belief that teachers need a healthy dose of independence. For leaders, giving autonomy can be a double-edged sword. On one hand, Sieh (2016) agrees on the importance for leaders to grant teachers autonomy. Fullan (2006) acknowledges educators appreciate getting to use their own judgment, to exercise discretion, and to determine the conditions of their own work in classrooms and schools. The positive side of promoting autonomy is it provides a sense of professionalism. Fullan (2006) believes leaders should communicate the importance of autonomy and then step back to allow teachers to run their classroom as they choose. Teachers create their own guidelines and expectations according to their own beliefs as long as they fall under the general rules and policies.

However, if leaders are providing autonomy, one study's findings do not support teachers feeling the autonomy. Sparks (2015) reports teachers feel they have less autonomy in 2012 than they did in 2003. Veteran teachers with more than 10 years of

experience showed the biggest sense of loss of control (Sparks, 2015). This may solely be due to trends in education over the last several decades where veteran teachers have experienced various levels of autonomy over time. As schools have become more structured with policies and accountability, it has forced a lessened sense of autonomy.

On the negative side, when teachers feel deprived of independence, they lack sufficient understanding of and rationale for the vision and goal given by leaders. The lack of understanding creates a silent resistance to share the same vision. Teachers often then avoid voicing their opinions in hopes of continuing on with their own goals. This may contribute to job dissatisfaction and frustration, making the consistency gap even wider (Griffiths & Portelli, 2015). With a variety of educational training programs, years of experience, content specialties, student needs, and leadership expectations, swaying too far to one side of autonomy creates varying grading practices and disruption to a shared vision.

Another contributing factor to inconsistencies is comfort. Teachers with a fixed mindset might become comfortably numb and therefore lack continuous growth. Their job is, and has always been, to impart their knowledge (Dweck, 2006). They have stayed status quo for so many years because the surrounding systems have gone untouched. For status quo schools, Schneider, Feldman, and French (2016) found “Each system has continued on, not because it represents the best we can do, but because it’s all that most of us have ever known” (p. 65). While it seems impossible to stop learning and changing, it is the resistance to being challenged and the threat to long-established beliefs being disturbed that holds teachers and schools back (Griffiths & Portelli, 2015). Good teachers continue learning with students and look to grow professionally. This includes

keeping up with current research on grading practices and standards. When teachers get comfortable with grading practices and refuse to change with the times, there are varying degrees of expectations with multiple levels of inconsistencies.

However, the level of comfort felt by teachers is a direct result of the leadership style. A transformational leader will seek to take others out of their comfort zone to encourage continuous learning. To push others out of their comfort zones, leaders challenge ideas, ask hard, uncomfortable, uninviting questions, and refuse to accept the way things have always been done (Griffiths & Portelli, 2015). On the other hand, a leader who lives in the comfort zone may allow others to do the same. This can inhibit a school from progressing forward since the leader is modeling a stagnant state of mind with no growth mindset.

Inconsistencies in implementing change can also be tied back to the leadership of a school or district based on how initiatives are presented to stakeholders. Initiatives may be implemented too quickly without all stakeholders understanding the change. Quick learners will excel and thrive on the consistent grading practices while others will struggle to learn the expectations. Thus, it is important for leaders to direct their attention on one main focus that will have the greatest sustainable impact on improving academic achievement for students (Fullan & Quinn, 2016). It is important to clearly communicate the main focus while acknowledging it may not have been a priority for previous leaders. If a leader focuses on consistent, equitable grading practices, the change will be slow and steady but eventually the varying practices will converge to one common goal...higher student achievement.

Fullan (2006) recommends constantly come back to the one focus until the new systems become the new continuity. Create parameters that all stakeholders understand and work towards without taking away all autonomy. Those parameters should include input and collaboration. If rules are arbitrary or issued for the sake of asserting authority, the process will result in chaos (Griffiths & Portelli, 2015). Teachers need to understand the content first, then the process. The content is what teachers will change and the evidence to support the change. The process is the how, including small steps, accountability, and support (Wiliam, 2016). The goal is to create consistency and shared responsibility, without pushing too hard or too fast and losing people along the way

Variation is also caused by teachers' perceptions of unfair grading done to them. Teaching is a unique profession where teachers spend at least sixteen years as the consumer before becoming the expert. Living on the other side of the teacher's desk for so many years, the experiences teachers have as students have an influence on who they become when they are the teacher. If teachers in training have obtained grades based on factors other than academic achievement, they will have a difficult time accepting theoretical principles that do not match with their personal experience (Allen, 2005). Teachers "typically reflect back on what they experienced as students and use strategies that they perceived to be fair, reasonable, and equitable...Teachers do what was done to them" (Guskey, 2004, p. 31).

Also, many times teachers' preconceptions and expectations conflict with grades (Alm & Colnerud, 2015). A teacher may have had a sibling in the past that did not perform well. If the younger sibling is then in the class, the teacher may have a preconceived notion that he will not perform well either. As the student puts forth great

effort and does not perform well, the student may perceive it as a reflection of what the teacher felt about the older sibling. The younger student can see this as unfair. When the younger sibling goes into the teaching profession, he may use effort and participation in grading as he found he was treated unfairly and wants to make a change. This teacher is now grading on grounds that do not reflect student knowledge simply due to how he was treated in school.

Though past experiences weigh heavily when making grading decisions, teachers' own belief systems tend to weigh more. Friedman (1998) concluded that teachers tend to find a way to have their grading policy reflect their own deeply held beliefs (not necessarily knowledge) about how students should be graded, even if building or district-level policies exist. When individual teacher beliefs regarding grading are implemented, accuracy and validity of each grade can be argued. As Allen (2005) has found, it frequently happens within the same school building that teachers hold very different views about the purpose of grades. The problem is compounded when teachers fail to communicate with their colleagues about their grading practices and each classroom is operating with different expectations.

If teachers and leaders communicate grading philosophy and practices, it is possible to have the purpose of grades move beyond numerical computations and allow for teacher judgment of proficiency (O'Connor, 2008). Communication can lead to collaboration and a shared vision. Teachers who put individual beliefs aside have some shared, accurate meaning of what grades represent (Allen, 2005). Switching to a systematic method of grading may contribute to a collegial responsibility for every

decision in the education system. A systematic approach holds every teacher and administrator accountable to the same standards.

O'Connor (2016) agrees that grading is complicated, subjective, emotional, and seems to have become inescapable in schools. However, the excessive time and energy placed on grading is not necessary. O'Connor (2016) has found grading is not essential for learning. When much value is placed on the grade itself, learning may be lost in translation. Grades have become an expectation within schools without necessarily finding a clear, consistent connection to student learning. This conflict forces educators to make a distinction between the importance of learning versus the importance of grading. The interpretation of these two topics leads to variation based on individuals' experiences. When grading practices are dependent on so many variables tied directly back to human emotion and beliefs, finding common ground is complicated.

Ultimately, according to Fullan (2006), the key is to reduce bad variation by increasing consistency. The process of changing a system by implementing common grading expectations to create consistency ties directly back to the main focus of increasing student achievement. How teachers perceive the change and perceive the lasting impact on their own pedagogy will have future implications for all students. Therefore, the purpose of this study is to explore how teachers arrive at their own grading practices and analyze teachers' perceptions of implementing common grading expectations.

Conceptual Framework

The conceptual framework to support this study is Fullan's (2004) work on change leadership. He identifies five components that work cohesively to bring about positive change. The five components of moral purpose, understanding change, building relationships, creating and sharing knowledge, and creating coherence were the foundation for implementing common grading practices at the research site. There is also a set of personal characteristics a leader should possess in order to support the five previous components. Leaders display energy, enthusiasm, and hope. When leaders live and breathe the five components while also possessing the three characteristic traits, they can draw in other members to generate long-term internal and external commitments. Fullan (2004) emphasizes when all these systems work together, the outcomes tend to be more positive and the bad outcomes are reduced or minimized (see fig. 1).

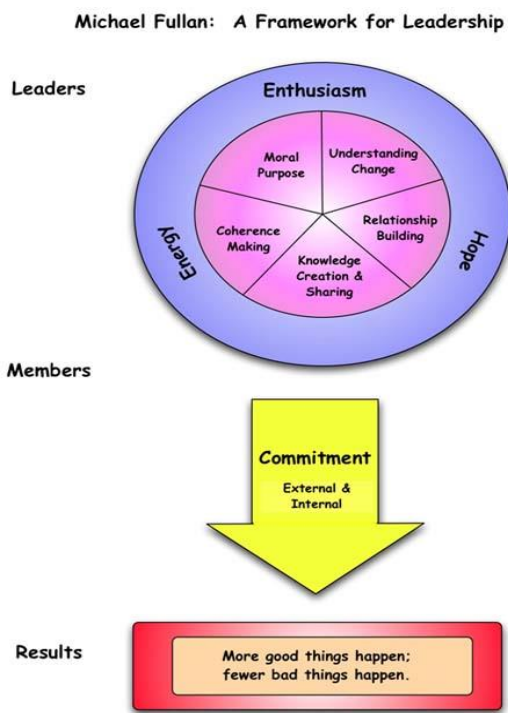


Figure 1: Michael Fullan's (2004) Framework for Leadership

The first component of moral purpose can be demonstrated in a variety of ways. Authentic leaders who have gained the trust of employees make moral decisions that benefit the larger group. Leaders create strategies to bring the moral purpose to life and motivate others to pursue the same vision (Fullan, 2004). Though moral purpose is focused on positivity, living the moral purpose also causes conflict. The vision and goals can conflict with interests or purposes of other groups. Starratt (2007) believes each human being is born with an internal moral agenda that cannot be negotiated or handed over to someone else. Therefore, the goal in creating a moral purpose is to align the vision with the motivators that already intrinsically live within each individual. Connecting purpose with others creates cooperative groups that thrive and move toward a common goal. Fullan (2007) reminds leaders that “although moral purpose is natural, it will flourish only if leaders cultivate it” (p. 25).

When leaders combine moral purpose with a thorough understanding of change, it tends to lead to a deeper moral purpose. In leading others through change, we have to unlock what has previously been accepted, form new ideas, and then lock those new ideas into place. However, leaders should use caution in being innovative. If leaders are constantly implementing new initiatives, others can feel overwhelmed or lack focus on one task at a time. It is a fine balance between facing resistance head-on and overwhelming others to the point that it destroys the climate and culture of the building. Understanding change also includes appreciating bumps along the way, also known as an implementation dip (Fullan, 2007). Leaders should not expect immediate results. In fact, it is more likely to anticipate a slide backwards before seeing forward progress. This dip can be seen as a natural consequence of change and evidence that change is occurring

(otherwise data would remain status quo). Being comfortable experiencing a dip demonstrates the leader is able to stay focused on the long-term goal and let go of negative consequences that come with leading change.

Allowing failed attempts to be part of the process also helps build relationships with stakeholders. Building positive relationships seems to be the one key component of successful change leadership. The “fake it ‘til you make it” philosophy is no longer effective in getting results (Fullan, 2007). Relationships are extremely powerful in creating a shared vision and internal motivation. This power can be used for good or to move an organization backwards. When the leader is an equal member in a positive community, the organization will move forward. However, when there is a strong sense of community and the whole building is not part of it, members tend to reinforce and support ineffective practices (Fullan, 2007). This strains other relationships and takes away from the shared moral purpose.

The fourth component of creating and sharing knowledge ties directly back to fostering positive relationships. When schools have data without positive relationships, it tends to lead to misinterpreted and misused data (Fullan, 2007). Different groups will view the data from one perspective, the one that best fits their model of thinking. Therefore, new data will not lead to new ideas. People are more likely to change if their emotions are tied to the change in behavior. The environment must be one of a shared moral purpose where people volunteer to share new ideas knowing they will be heard. Starratt (2007) reinforces the idea that new knowledge is often rejected because it goes against previously accepted ideas, philosophies, and frameworks. New knowledge emerges from changing circumstances, new technology, and needs that arise. Without

questioning existing beliefs and collective sharing of background knowledge, organizations will become stagnant.

Finally, creating coherence will provide the structure for productive change. In an organization where coherence is reached, the leader has let go of having full control and embraces ambiguity and the natural complexities of problem-solving processes. When coherence is lacking, the leader will often times resolve all the problems on his own. Or he will delegate decision-making tasks to live in the comfort of having someone else take responsibility for decisions (Fullan, 2007). It does not build capacity nor does it tie back to creating a shared moral purpose. It also lessens the effect of peer pressure. Others often thrive from a healthy dose of peer pressure to meet expectations and be part of finding solutions. Ambiguity is messy and in that mess hides the creative solutions stakeholders need to seek when the status quo is disrupted (Fullan, 2007).

When Fullan's five components of moral purpose, understanding change, building relationships, creating and sharing knowledge, and creating coherence are supported by energy, enthusiasm, and hope, the possibilities are endless. Unavoidable change will occur but the "how" and "why" it occurs will differ based on the vision of the leader. When systems, structures, and a positive culture are laser-focused on the moral purpose, the organization is bound to move forward. According to Fullan (2007), "Information is on paper. Knowledge is in people" (p. 115). Maximizing people as the most important resource, and supporting them along the way, will continuously build coherence and maximize capacity building. Coherence and capacity building tie to the purpose of this study as both can impact teachers' perceptions of their own grading pedagogy.

Research Questions

The central research question to be addressed is how do teachers formulate their own grading practices and what are their perceptions of implementing common grading expectations? Within the study, the following sub-questions will be answered:

1. What factors do teachers report impact their current grading philosophy and how do the findings vary among subgroups including grade level taught, core teachers, exploratory teachers, other certificated staff, years of experience, content area, leadership experience, and experience teaching in more than one district?
2. What are teachers' perceptions of implementation fidelity and how do the findings vary among subgroups including grade level taught, core teachers, exploratory teachers, other certificated staff, years of experience, content area, leadership experience, and experience teaching in more than one district?
3. What are teachers' perceptions of common grading expectations and how do the findings vary among subgroups including grade level taught, core teachers, exploratory teachers, other certificated staff, years of experience, content area, leadership experience, and experience teaching in more than one district?

Definitions of Terms

Assessment: Gathering information about students' level of understanding (Marzano, 2000). Knowledge may be demonstrated through a physical skill, project, reflection, test, quiz, essay, or other means deemed appropriate by the teacher.

Capacity Building: Skills, competencies, and knowledge needed in order to effectively accomplish goals (Fullan & Quinn, 2016).

Coherence: Consists of the shared depth of understanding about the purpose and nature of the work (Fullan & Quinn, 2016).

Collaboration: Development of social capital (Fullan & Quinn, 2016).

Common Grading Expectations: Teams were given topics to reach consensus on, one of which was the topic of grading. Teachers discussed philosophies and individual practices to agree upon common expectations around late work, incomplete assignments, grading or not grading homework, point values and weights of grading categories, point deductions, weights of extension activities, timelines of accepting work, and retake opportunities on assessments. Each core team within a grade level was expected to implement the agreed upon practices with all students, no matter the content of the course.

Core Teacher: Teaches Math, Science, Social Studies, Reading, English, or a combination of these subjects. This group also includes all special education teachers.

Ethical Educational Leadership: Values, virtues, and decision-making focused on justice, critique, care, and the profession. It places students at the center of the decision-making process always reflecting on the question, “What should the professional educator take into account to consider the best interest of the students?” (Portelli & Griffiths, 2015).

Equality: Providing equal opportunities and equal education systems. Equality ensures everybody has the same or is treated the same (Griffiths & Portelli, 2015).

Equity: Ensures procedures are fair without bias or favoritism. Everybody may not be treated the same but all are treated fairly within the constraints of the overarching fair procedures (Griffiths & Portelli, 2015).

Exploratory Teacher: Teaches non-core classes, often referred to as elective courses. These include Art, World Language, Physical Education, Family and Consumer Science, Industrial Technology, Computers, and Music.

Grading: A compilation of marks from multiple assignments and assessments that lead to a final judgement about students' achievement (Tomlinson, 2005).

Grading Pedagogy: Beliefs held and practices utilized by teachers regarding grading procedures.

Leadership: The ability to define the reality of others (Griffiths & Portelli, 2015).

Moral Purpose: Acting with the intention to bring a positive change (Fullan, 2004).

Other Certified Staff: Includes Counselors, Speech Language Pathologist, Media Specialist, Social Worker, and School Psychologist.

Similarity: Using the same grade calculating method (total points versus weighted categories) as well as comparable opportunities for students to submit work that will be recorded as a grade.

Social Capital: The quality and quantity of interactions and relationships with the group including teachers' commitment to work towards a common goal (Fullan & Quinn, 2016).

Systemness: A small number of ambitious goals pursued, while reducing distractors, growing capacity, using student and other data transparently for developmental purposes. It includes building in strategies for stakeholders to learn from each other and make progress with feedback and supportive intervention when needed (Fullan & Quinn, 2016).

Limitations

Though the research site is similar in size and demographics to other district schools, it is still unique in its own way and does not represent every other school. Findings may not be generalized to other populations. Due to the small scope of participants and analyzing participant perceptions, the findings may not represent causality.

The role of the researcher is also a limitation. The researcher is an assistant principal at the building where research is conducted.

The survey posed natural limitations as any survey would pose. It was used to gauge and gather perceptions of people that may be influenced by other factors on the specific day and time the subjects responded. There were also the following limitations:

1. 50/75 certificated staff returned the survey.
2. The characteristics of staff that answered, versus those that did not answer, may skew data (age, gender, years of experience, etc.).
3. There is a different number of respondents in each category assigned to staff (6th, 7th, 8th, exploratory, other certificated) and may not be proportional to the whole population.

Delimitations

The boundaries of the study include (a) using perceptions instead of quantitative data to measure occurrence of change, (b) results come from one building in one school district, (c) using on-the-shelf data from a survey used to give administrative team feedback on change in grading expectations, (d) the survey was given one time (November, 2016) throughout the first two years of implementation.

Assumptions

This study is conducted under three assumptions:

1. Staff answered the survey questions honestly and accurately;
2. Staff understood all the terminology on the survey;
3. All staff played an equal part in creating their common grade level expectations.

Significance of the Study

Educational leaders focus on creating equitable opportunities for students. In order to lead change towards equity, educational leaders demonstrate a clear understanding of the “why”, the process, and how to sustain change over time. Creating change is not solely about changing behaviors. To sustain long-term change, a shift in emotional and investment climate will occur to create a new motivation (Kanter, 2004). Change is seen as a process, not an event. Educators do not arrive once and for all at an end goal. There are always new developments and innovations. It is a continuous process of making and reshaping mindsets and culture (Fullan & Quinn 2016). To

redefine others' reality, leaders display a self-confidence that models an open mind and willingness to change. Demonstrating a readiness to grow (and doing whatever it takes to get others to demonstrate the same), will keep the school moving forward with change (Dweck, 2006).

This study will help educational leaders focus on the process of change and relying on accountability. Leaders acknowledge accountability has a moral component which can be defined differently for individuals. It is associated with integrity and passion with a strong sense of equity and justice. It justifies every decision made (Griffiths & Portelli, 2015). However, not all staff members may come with this inner passion. It is inevitable that additional supports will be needed for some and those supports provide a solid foundation for those needing additional assistance. In providing the supports, leaders will model the expected behaviors. Fullan (2006) shows cultures do not change by mandate. They change by displacing existing norms, structures, and processes. A change in culture depends fundamentally on modeling the new values and behaviors that are expected to displace the existing ones. For effective change to occur, the leader holds himself accountable to the results just as he expects accountability of all others.

Finally, this study will test the need to create consistency while intentionally building systems. Many times in education, the planning takes so much time and energy that implementation becomes the second priority. Fullan (2006) urges leaders to reduce the distance between planning and action. The actual documents are far less important than implementation, execution, and monitoring. Knowing the idea of change is difficult for some, leaders will model expectations and stretch teachers outside their comfort zones

to encourage continuous learning. As research on effective grading practices has evolved, educators too shall evolve.

The ultimate goal for educational leaders is to create a culture of equity in grading expectations to match pedagogy with current research-based best practices. Not only is equity a positive for students, it has a positive effect on the overall climate of the school. Fullan (2006) has found that cultures who focus on equity have environments that tend to be less violent, more supportive and inclusive, and are marked by better health. Similarity in all aspects of the school contribute to a more positive culture. Closing the gap in inequitable grading expectations will better support all students' educational opportunities. While attempting to close the gap by implementing common grading expectations, this study will also help to identify trends in groups of teachers whose perceptions of grading pedagogy have changed. Whether leaders agree or disagree with teachers' perceptions from the survey, those results will not be investigated. Perceptions drive reality and therefore, teachers' perceptions will be the driving force behind investigating common grading expectations and the possibility of a shift in pedagogy. Following the literature review in chapter two and methodology in chapter three, chapter four will display the quantitative and qualitative results from the survey while chapter five will discuss the findings and further implications.

Chapter 2: Research Literature Review

Considering the purpose of this study is to explore how teachers arrive at their own grading practices and to analyze teachers' perceptions of implementing common grading expectations, the literature review will focus on five main components: why similarity is important, grading practices, implementing a change initiative, evidence needed to show successful and sustainable change has occurred, and the steps taken specifically at the research site to implement common grading expectations.

Why Similarity is Important

Similarity in grading practices is important for both students and teachers. Though there are multiple perspectives, the priority is doing what is best for students. Just as principals are accountable to their teachers and staff, teachers and staff are accountable to maintain good practices and equitable and fair treatment for students (Griffiths & Portelli, 2015).

When looking from the student perspective, consider whether or not grading practices provide equality and equity for students. Griffiths & Portelli (2015) identify the differences between equity and equality. Equality is providing equal opportunities and equal education systems. Equality ensures everyone has the same and is treated the same. Equality poses a problem in the current United States education system as not all students need the same. If all students were treated the same, teachers would be unable to meet individual students where they are on their learning journey.

On the other hand, equity challenges the idea that it is even possible to distribute things in an equal manner. Educational equity focuses on raising the achievement of all students while narrowing the gaps between the highest and lowest performing students. Equity also aims to eliminate the disproportionality of racial groups within the highest and lowest achieving student groups. Equity ensures that procedures are fair and individuals receive equal treatment, without bias or favoritism (Griffiths & Portelli, 2015). Creating equity for students forces teachers to step out of a fixed mindset and have a willingness to step into a student's world. Wolter (2016) believes it is not the deficits students bring to school that limits their growth. Instead it is their lack of equitable opportunities at school.

Not only is equity a positive for students, it has a positive effect on the overall climate of the school. Fullan (2006) has found cultures who focus on equity have environments that tend to be more supportive and inclusive. Similarity in all aspects of the school contribute to a more positive culture. When the quality of social relations improves, people are more likely to trust each other. In general, the community life is stronger. A school can have a whole staff of amazing individuals. However, Fullan & Quinn (2016) understand it is still the collective culture and climate that are more important. Culture will eat up individuals faster than one can produce them. Therefore, the focus is on culture and individuals simultaneously. The goal is for individuals to come together as a team and work in a cohesive manner to level the playing field for all students.

Specifically in middle school, lack of similarity in grading practices can have a negative effect on student achievement. When the work becomes more challenging and

grading policies become more complex, the transition to middle school can be a challenge for many students. A study done by Randall and Engelhard (2008) shows how grading practices can change from elementary to middle school. They compared the grading practices of elementary teachers to middle school teachers. The results showed teachers in the two different levels made different determinations of what should be included in a grade. Middle school teachers had more strict grading procedures while elementary was more lenient. Elementary teachers felt grades were more personal since they spend more time with students each day. They reported feeling the need to protect the self-esteem of their students. Middle school teachers felt more pressure to prepare students for high school and employ grading practices more similar to high school. This change in grading philosophy contributes to a negative effect on student achievement.

Beyond academic achievement, lack in similarity of grading practices can also negatively impact students' self-confidence. While students are experiencing the physical and emotional changes of adolescence, they are trying to balance learning a new school system of having multiple classes along with multiple teachers. When the developing adolescent's mind cannot organize all the moving parts, he experiences frustration and a feeling of inadequacy. In turn, grades can begin to suffer, but not everyone's grades suffer equally (Dweck, 2006). Some students are more aware of the inconsistencies than others. When students become aware of the different grading expectations, the level of confusion increases. Students realize the learning opportunities provided vary and the opportunities to earn a grade equal to a classmate becomes unbalanced. A student's grade can largely depend on the luck of the draw of whichever teacher's class he is

assigned (Dweck, 2006). The lack of similarity has now impacted the student's grades as well as his self-confidence in being able to achieve the same level of success as his peers.

Teachers having a shared understanding and practice of grading expectations can benefit students and it can also be a solution to educators feeling overwhelmed.

Educators complain of confusion and overload through initiative fatigue and working in silos. Confusion and overload can result in distrust and lack of collaboration (Fullan & Quinn, 2016). When a group of professionals come together, collaborate, and reach consensus on grading practices, they are no longer working in isolation. The level of fatigue will lower as each knows they only have to pull their own weight in meeting the common expectations. It no longer becomes a game of popularity or trying to keep up with others.

When the achievement expectations are the same for all, teachers then hold students to the same level of expectations instead of altering them according to groups of students or content areas taught. Lowering standards only "leads to poorly educated students who feel entitled to easy work and lavish praise" (Dweck, 2006, p. 193). Instead of lowering standards, teachers will aim to find innovative ways to reach students and help them achieve the high standards set for all. Without changing the level of expectation, teachers can still exercise professional judgement and allow for some flexibility for individual circumstances.

Professional judgement also plays a large part in the group conversation as teachers are creating their common grading expectations. Since all stakeholders are involved in the process, teachers no longer have to fear the top down policies, confusion, or distrust. Each person has the opportunity to contribute to reach a consensus. Being a

part of the plan also heightens the level of accountability for all. Griffiths and Portelli (2015) define accountability as “being proactive in examining, questioning, and going beyond expectations that inadvertently work against the well-being of students” (p. 41). Reaching consensus before implementation opens the conversation to question current practices and decide on common grading expectations that work towards the well-being of students. Taking the time to reflect on practices and reach consensus forces all to get away from the way it has always been, which differs for each individual (Griffiths and Portelli, 2015).

Similarity in grading practices also increases the accuracy and validity of assessments. According to Allen (2005), "The most fundamental measurement principle related to meaningful assessment and grading is the principle of validity" (p. 218). If an assessment is not valid, it will not measure the intended outcome. The grade stemming from that assessment then holds no value. One of the first steps to improving accuracy and validity in grading is for leaders to set clear expectations regarding grading practices.

When beginning the conversation about grading practices, principals anticipate it evoking fairly powerful reactions from teachers. Teachers' grading practices are rooted in deeply held beliefs, some of which are learned through education while others are learned behaviors from being a student. Friedman (1998) acknowledges there is a variety of reasons why teachers have not generally been held accountable for the grading practices they use. One reason may be the principal does not require teachers to submit a copy of their grading policy. Requiring teachers to submit their policy allows the principal to review practices and reference them throughout the year. It also allows for comparisons across a building to close the gap in equity concerns. The principal can then

determine if grading practices are fair, consistent, and have set high expectations. Schools want to set the standard for high expectations since Bonesronning (2010) found students, especially those that are high achieving, perform better when graded against rubrics that reflect high expectations. Leaders that set high expectations and hold teachers accountable to those expectations can expect that to transfer to students and thus raise student achievement.

When principals review teachers' grading practices, fairness and equity are two components to check. Friedman (1998) believes if teachers do not begin with the expectation that all students will be treated fairly (to the extent this is humanly possible), it will be impossible to know what grades mean. It is the responsibility of the principal to monitor grading practices and assist in ensuring teachers know how to do the calculations and understand what they mean. For example, do teachers have a clear understanding of the differences between weighted grades and grading strictly on points? If teachers do not understand the calculating process, they are highly unlikely to follow what is written in their policies.

Also, having an awareness of the general beliefs in different subject areas will help principals evaluate fairness and equity. One study by Proitz (2013) showed the subject matter being taught does have an impact on the evidence collected and tools used for grading. For example, math teachers emphasize academic enabling behaviors (ability level, student effort, paying attention, improved performance and work habits) less than teachers in other content areas (Duncan & Noonan, 2007). When content teachers perceive grading practices differently, it is left to the principal to decide where to draw the line on accurate, consistent, and fair grading practices for all students across

content areas. The current paradigm of grades can open up or close down important future learning opportunities for students. Invalid grades that understate the student's knowledge may prevent a student with ability to pursue certain educational or career opportunities (Allen, 2005).

The assumption that student grades are to be distributed across a scale also contributes to lack of similarity. Kalthoff's (2013) study investigated the grading procedures of teachers at German high schools in regards to one common assessment. Among other factors, student scores were determined by teacher biases and the expectations that students would be distributed across grades similar to a bell curve. They worked under the assumption that there is no possibility all students' work could be considered for the highest grade. Each student's paper was placed in relation to peers so teachers could distinguish the learning gaps from paper to paper. It created a classification system to put students at all levels of the spectrum. Equality in grading practices was not the goal, but rather forced variance, to make the grades socially valid and acceptable (Kalthoff, 2013). Grades have a lasting impact on future opportunities for students. Ensuring fair and equitable grading practices, free of bias and forced variance, provides students equitable opportunities for future endeavors.

Another step to improve similarity is to educate teachers regarding valid measurements and a common purpose of grades. Internationally, less than one-third of teacher education programs require an assessment course (Allen, 2005). Many of those required courses are focused on informal assessments, or standardized assessment of students with special needs, and not focused on classroom assessment and grading. Fewer than half of the fifty states require specific coursework on assessment for their

initial certification (Allen, 2005). This missing coursework leads to teachers in training lacking sound measurement principles relevant to their classroom procedures. College instructors are tasked with providing a solid theoretical base supported by sound mathematical measurements to replace future teachers' notions of assessment and grading.

Ultimately, having similar grading practices comes down to ethics. The basis of ethics leads to one belief: human behavior has consequences for the welfare of others. We are capable of acting toward others in such a way as to increase or decrease the quality of their lives. We are capable of helping or harming. What is more, we are theoretically capable of understanding when we are doing one and when we are doing the other (Elder & Paul, 2011). The choices educators make on how they grade students has a lasting effect on the possible future opportunity paths for students. Making a conscious, research-based, ethical decision on shared grading practices leads back to the goal of doing what is best for students.

Grading Practices

Grades are used for a variety of purposes, depending on each district's policies and expectations. Some teacher-assigned grades include aspects outside of academic achievement and are used to communicate a variety of aspects of a student's experiences in school. However, Allen (2005) says the sole purpose of grades is to accurately communicate to others the level of academic achievement that a student has obtained. In order for grades to be used by all stakeholders to make important decisions on students'

futures, grades need to communicate the same pieces of information in an accurate fashion.

Zeros. Assigning zeros for missing assignments has a direct impact on student grades. Guskey (2004) and Marzano (2006) agree that a zero is seldom an accurate reflection of what a student has learned or is able to do and should not be used in grading systems. It does not tie back to the main purpose of reporting grades which is to communicate the level of achievement a student can demonstrate. The impact of a zero on an overall grade is magnified if teachers use average scores to determine final grades (Guskey, 2004). A student who has demonstrated a high level of proficiency on assessments throughout a semester can drop to less than proficient due to one zero. Thus, the final grade is not a reflection of student knowledge. Teachers can consider using other measures of central tendency beyond the mean. O'Connor (2008) offers the suggestion that the median may be a better representation of student achievement, depending on the number of grades assigned.

Some teachers also use zeros as instruments of control (Guskey, 2004). They may believe assigning a zero will motivate students to comply and meet teacher expectations. According to Guskey (2004) and Vatterott (2009), the use of zeros or low grades do not serve as effective punishments. Zeros do not motivate students to improve, nor do they teach students to take responsibility for their own learning. The reverse is often true. Assigning a zero can cause students to withdraw from learning and fail to progress. This is especially true for students with barriers at home. If students are required to work after school, care for younger siblings, or take care of the household while parents work, completing work at home becomes an impossible task. Recording a

zero ultimately punishes the student for living in poverty (Neason, 2017; Vatterott, 2009). The compounding zeros become stigmatizing and continue to lower the hopes of the student of ever catching back up.

Grade inflation. Stanley and Baines (2010) acknowledge a teacher that sets nonnegotiable standards will likely experience some resistance from students and parents. On the opposite side, when the majority of students are succeeding with A's and B's in a course, a teacher's life becomes easier. With no angry parents or confrontational phone calls, teachers gain back time and lower their own stress and anxiety. These desirable results can cause teachers to inflate grades. Teachers may also inflate grades to boost students' self-esteem. However, the opposite effect is generally true. Inflating grades prevents students from learning skills needed to succeed in the real world. And where school funding depends on high pass rates, the likelihood to encounter grade inflation is higher (Stanley and Baines, 2010). This can lead to lasting negative effects. Students who are rewarded with grades that were unearned lose respect for their teachers and the subject. This also fosters an attitude of entitlement among students (Stanley and Baines, 2010). Students with inflated grades will enter college with unrealistic expectations and have not learned the skills to handle dealing with a lower earned grade that will not be changed in a post-secondary course.

Besides funding, there is pressure to alter grades coming from parent expectations. Where there is school choice, parents will seek the school with the highest performing students. Grade inflations, accurate or not, will serve as evidence of higher student achievement and attract parents with choice in schools. Smith and Fleisher (2011) report grade inflation is tied to levels of experience of educators. In a study

throughout colleges and universities, they found professors without tenure reported higher grades. The prediction was that grade inflation was due to the need for higher student evaluations to keep their job and work towards tenure. However, the study found the non-tenured staff reported higher grades due to inexperience and not understanding the scope of assigning grades at the college level. This finding could be extended to K-12 education where new teachers entering the profession simply do not have enough background knowledge, experience, or support to submit grades that are reflective of students' demonstration of knowledge. Or, if teachers do have enough training, they could utilize grades as a self-fulfilling prophecy. Teachers may assign grades more on expected performance instead of actual performance and have overall grades reflect what they believe to be true about the quality of their own teaching (Zoeckler, 2007).

Non-achievement factors included in grades. Research by Allen (2005) supports the misconception regarding including non-achievement factors in grades by confirming that two out of three teachers believe effort, student conduct, and attitude should influence final grades of students. Depending on district and teacher grading practices, some grades reflect items such as tardies/attendance, behavior, and effort. There have been court cases where these aspects have been challenged. Friedman (1998) references *Wermuth v Bernstein* (1965), where the New Jersey commissioner of education wrote:

The use of marks or grades as deterrents or as punishment is likewise usually ineffective in producing the desired results and is not educationally defensible. Whatever system of marks and grades a school may devise will have serious

inherent limitations at best, and it must not be further handicapped by attempting to serve disciplinary purposes also (p. 82).

Additionally, in *Dorsey v Bales* (1976), the court of appeals upheld a lower court ruling that reducing the grades of a student who had been absent due to a suspension for misbehavior was invalid. The suspension negated the school board's authority to impose additional sanctions, in this case, grade reductions (Friedman, 1998).

In regards to effort, Friedman (1998) recognizes that if teachers are allowed to mix something like attitude with course grades, grades will be highly influenced by teacher opinions and will no longer be objective. Student effort is subjective and teachers will have a difficult time justifying and accounting for varying levels of effort. Qualifying the levels of effort and making judgments on effort can greatly change from day to day. Judging effort ignores the inherent difference that not all students need to exude the same level of effort to achieve at the same level. Academic ability plays a role in the level of effort a student needs in order to demonstrate mastery of content.

Rewarding or consequating individual students for the level of effort exuded contributes to inconsistencies. On the other side of the coin, Cross and Frary (1999) urge teachers to grade on three criteria: product (what students know and can do), process (effort, work habits, participation, attendance), and progress (improvement). They believe it comes down to personal preference of which criteria should be used. Therefore, effort and other non-achievement data is factored into a final grade. This option could contribute to the lack of consistency in what a grade communicates to stakeholders.

Resh (2009) took a different approach to looking at grading effort. His study found teachers rely more on performance while it is the students who were the ones that tended to emphasize their investment in effort. The disparity between the two perspectives was a possible cause of students feeling like they were graded unfairly. The argument of whether or not to grade on effort can have multiple perspectives depending on the stakeholder. Some students may depend on being graded on effort as a result of cultural influences, parental pressure regarding priorities, or an internal motivator. When teachers, students, and other stakeholders have different levels of value placed on effort, the choice to include or remove effort in a grade will create conflicting emotions between individuals.

Expanding assessment options. Moving beyond paper/pencil assessments, teachers expand on assessment options by using other means deemed appropriate for the content. These include strategies such as a probing discussion where the teacher meets with the student and asks questions at varying depths of understanding (Marzano & Heflebower, 2011). The teacher continues to ask questions at different levels until she is confident about the student's level of proficiency. Teachers can also use unobtrusive assessments where they observe skills taught and can be assessed on a proficiency scale when the student may not be aware he is being assessed. For example, a physical education teacher may observe a student at recess throwing a ball using the skills taught in class. This can be recorded as an unobtrusive assessment and assessed in the moment the student is using the skills. Teachers may also use student-generated assessments. A student can propose what he will do to demonstrate a specific level of performance. The teacher and student come to an agreement of the alternative means to demonstrate

knowledge (Marzano & Heflebower, 2011). The number and variety of assessment measures employed affect the validity of a student's grade. Using a variety of assessment options takes creativity. Carlson (2003) urges teachers to be creative in finding other assessment options.

One such school, the Hong Kong International School (HKIS), has taken a creative approach to assessing students. Students are highly involved in assessment practices. Students are required to reflect and assign themselves a grade they feel is fair. Class grades are assigned using a “majority rules” approach. After a student assigns himself a grade, two teachers also select the grade they feel is fair. The three separate grades are evaluated and a consensus is reached regarding which grade seems fair. It is an inclusive approach that removes the power of the teacher to either validate or bring down a student. It removes the barrier that is often established between students and teachers. Collaborating with students on grades and assessment options creates a stronger sense of community (Wiggins, 2016). Collaborating also builds students’ ownership of the result. It shifts the sole power away from the teacher so students understand grading is not something that is “done to them.” When students understand the rubric and grading criteria, they become more aware of what quality work looks like and are more inspired to reach their full potential (Kallick & Zmuda, 2017).

Hong Kong is not the only example of moving assessments beyond tests. Boaler (2016) examines some other countries and their assessment practices. Finland is one of the highest-scoring countries in the world on international mathematics assessments. Yet, their students do not take any tests during school. They spend their school hours diving deep into rich topics and focus on problem-solving skills. In England, students

spend three years (ages 13-16) working on open-ended projects in mathematics. They do not take tests and their work is not graded. Though students do not practice test-taking skills and they are unfamiliar with working through a timed assessment, they score significantly higher than a cohort group of students who do prep for the national exam. The students who score higher show they are willing to attempt questions where they have no knowledge of the topic. They score similarly to their cohort on procedural questions and outscore the cohort on conceptual questions (Boaler, 2016). These results show it is possible to move students away from standardized tests and frequent grades in the classroom to project-based assignments that allow students to think and synthesize at a deeper level.

Homework. At some point, teachers will administer a test of some kind to see if students have learned from the mistakes that they were free to make while doing homework. However, many teachers grade every assignment, which pressures students to get it right the first time (Friedman, 1998). If students are not encouraged to try homework without fear of being penalized, many will simply choose not to attempt it. Learning from errors is how students find clarity in their new learning. If students are penalized on homework and all assessments, they are never given the opportunity to fail and learn from mistakes. Students are ultimately being held accountable to do everything right the first time.

Using homework as a means to learn instead of simply earning a grade will also help prepare students for the future. Psychologist Jason Moser studied what happens in the brain when people make mistakes. The brain has one of two responses. An ERN response causes an increase in electrical activity because the brain knows there is a

conflict between the correct response and an error made. A Pe response is a brain signal forcing a conscious decision to pay attention to the mistake. With both responses, a synapse has grown in the brain. Some teachers will argue the brain will not grow the synapse unless the error is corrected. However, Moser's study shows the synapse occurs whether the error is corrected or not (Boaler, 2016). Schools in China use this research to encourage mistakes with students. Chinese teachers often call on students who they know have the wrong answer. Students proudly explain their wrong answer and welcome feedback to correct the error. It is not a social stigma in China to give a wrong response as it is in American schools (Boaler, 2016). Each opportunity students have to make a mistake is a chance for their brains to grow. When students are turned off to doing homework because of possibly earning a bad grade, they have missed the chance to grow their minds and form new synapses. Moser's work also extended outside of schools into the real world. He was able to show that the most successful business people have experienced the most mistakes over time (Boaler, 2016). If students have the chance to make mistakes without penalty, this mindset may continue into their adult years where they will take chances and continue to grow their minds through making errors.

Vatterott (2009) suggests to move away the old paradigms of grading homework. While some parents argue homework is a sign of a high-achieving school and teaches responsibility, other parents believe it is the (cult)ure of homework that is separating students into winners and losers. Instead, Vatterott suggests shifting from grading homework to checking for understanding and focusing on feedback. Formative assessment drives instruction and encourages student ownership of learning. This shift will take time to replace old paradigms. In a study of 50 countries, U.S. teachers led the

world in grading homework (Vatterott, 2009). Furthermore, PISA looked at data from 13 million students around the world and announced homework only increased inequities in education. It did not positively contribute to student achievement. It either had a negative impact or no impact (Boaler, 2016). The concept of not grading homework is one that will die hard.

Retakes. There is research on both sides of the argument regarding allowing students to retake assessments and/or resubmit assignments. Boaler (2016) says teachers should always allow students to resubmit for a higher grade. It demonstrates a growth mindset and shows students that teachers care more about learning and less about reporting grades. O'Connor (2002) refers to retakes as "improvement grading". Grades reflect the student's most current demonstration of a skill. If a student retakes an assessment, O'Connor cautions to avoid applying mathematical algorithms to change the level of understanding demonstrated. For example, a first and second attempt should not be averaged together. The second attempt will provide the most current information and thus be used to report the grade. It also negates other factors that did not allow for a student to perform at his best on a given assessment day dictated by the teacher. Whether those factors are academic, physical, or emotional, a second chance provides students an opportunity to demonstrate learning. O'Connor (2002) recognizes the time constraints and additional work for teachers to provide second chances but ties the practice to the real world saying, "As life provides second (and more) chances, so should school" (p. 131).

On the other hand, allowing students to retake assessments may not always be a positive option. Some teachers argue allowing retakes unintentionally creates a habit for

students to not give their best effort the first time, knowing there is always a second chance. It is a vicious cycle of a student constantly playing catch up. Time is also a concern. If students demonstrate a lack of understanding on the first assessment, clearly these students are the ones needing the most instructional time. However, the more they retake, the less instructional time they have in the classroom to learn or relearn material (Kamenetz, 2014). Allowing multiple retakes during the school day robs students of the additional instruction needed to grasp concepts for the sake of understanding material instead of simply to earn a better grade.

Grades as rewards and punishments. To keep the focus on learning, grades cannot be the motivation for students. Teachers succeed only if they help students want to learn. Dorothy De Zouche, an experienced junior high teacher, said "If I can't give a child a better reason for studying than a grade on a report card, I ought to lock my desk and go home and stay there" (Winger, 2005). An alternative to assigning a zero or low grade is to assign an incomplete with explicit requirements for completing the work. No excuses accepted. A consequence of a Saturday school or some other gift of time for not completing the work is direct, immediate, and academically sound (Guskey, 2004). Assigning a low grade or a zero does not provide motivation for the student to learn the content and improve their depth of knowledge.

In addition, when grades are used as a punishment, it creates a compounding problem. For example, when averages are used to compute final grades, there will be students who experience despair in the last few days of a semester. This leads to an increase in student discipline problems. The increased discipline issues take teachers' and administrators' time away from effective instruction. Resources are directed toward the

discipline issues. On the contrary, grading practices can fuel further improvements without being used directly as rewards. If school systems improve their grading systems, their work on curriculum, instruction, assessment, and leadership will improve. Using grading practices that are backed by research is a high-leverage strategy that can boost every other element of the educational system (Reeves, 2011).

Minimum grade. One possible solution to reduce failure and attrition rates of students is to implement a minimum grade policy. Before implementing the practice, the student population is considered to see if it would benefit a group of students. It is meant to assist students who swing from grading period to grading period within 20 percentage points. A student who consistently scores below 50% would not benefit from the system. The practice prohibits teachers from assigning anything below a minimum threshold, often somewhere around a 50%. This helps to remedy the unfairness of a 100-point grade scale where a disproportionate three-fifths of the scale has failing scores. The minimum grade would also lessen the effects of a zero on an overall grade (Carifio & Carey, 2009).

Implementing a minimum grade system can also increase student motivation. Carifio & Carey (2009) found assigning minimum grades enhances student learning by encouraging students to stay engaged. Students will view failure as temporary. They can earn a low grade early in the semester and still be able to work their way to a more desired grade. If the minimum grade was not in place, the mathematics of calculating grades would impede the student from earning a more desirable grade throughout the rest of the grading period. It is one step towards creating a more compassionate and caring culture throughout the whole school.

Friess (2008) and Richmond (2008) present the other side and negate proponents of implementing a minimum grade. They believe it lowers students' expectations for reaching a minimum level of competency on the content. It may also be viewed as unfair by helping low-performing students have a grade that is perceived to be unearned. A minimum grade may also provide a false sense of security to students. Students believe they have a certain level of understanding and then later realize they lack the skills they once thought they had. Implementing minimum grades also cause grade inflation (see section on inflation) or social promotion, neither of which are an intended purpose of grades.

Initiatives Involve Change

Any initiative involves change and change is uncomfortable for most people. To successfully implement a change initiative, school leaders consider multiple factors. First, identify the problem. Stakeholders conduct a thorough review to pinpoint any possible weaknesses and generate strategies to correct them. When the plan is devised and ready for implementation, monitoring the progress of the plan is crucial. Regular reviews of the progress and discussions around any changes need to be transparent and clearly communicated. Also, throughout the change process, all members are clear about everybody's role. Just as the plan is carefully devised, the behaviors, tasks, and targets for all should also be clearly spelled out (Fullan, 2006).

Clear communication can be a key to success or the barrier to success. Once communication begins to break down, it can cause a chain reaction of other related negative consequences. Kanter (2004) identified nine pathologies working together as an

emotional and behavioral chain reaction. Once “communication decreases, criticism and blame increase, respect decreases, isolation increases, focus turns inward, rifts widen and inequities grow, initiative decreases, aspirations diminish, and negativity spreads” (pp. 97-98). Each of these reactions pose their own challenges and each one would need to be addressed individually. Therefore, not only is communication upfront at the start of a change initiative important, but frequent check-ins and discussions along the way are necessary to avoid the negative chain reaction.

As frequent reviews are held, it is important to remember that change takes time. Leaders provide time for plans to be executed and not be quick to change direction. Each initiative is to be tied to the long-range vision and mission of the school, thus allowing a long-term plan to unfold in stages. If a leader is looking for quick fixes, the leader and the school are likely to fail. Hargreaves and Fink (2006) acknowledge quick-fix changes often exhaust the teachers and/or leadership and improvement efforts are not sustained over time. Quick fixes may quickly promote a principal but the overall result for the school is not positive. Teachers will regress in efforts due to feeling abandoned by their leader or go back to old ways when the pressure to change is off. Either way, the quick fixes do not promote long-term success of the school. Collaboration takes time and time is worth the investment. It creates a shared vision where ownership of the vision does not rest with one single leader but instead with the whole community (Association for Middle Level Education, 2012).

To encourage the long-term success of any initiative, leaders take into consideration the mental well-being of staff and focus on reaching the emotions of others. Fullan (2006) says behavior change happens most effectively when we speak to people’s

feelings. Help to find ways to see the problems or solutions that influence emotions, not just thoughts. Thoughts can be changed or swayed more easily than human emotions. To begin reaching others at the emotional level, leaders increase internal accountability. Fullan (2006) explains “Internal accountability precedes external accountability and is a precondition for any process of improvement” (p. 26). When teachers can see the need for change and have internal buy in, they are more likely to commit to the change process over a long period of time. Schools do not succeed when they respond to external cues or pressures until they have internalized the organization’s pedagogy. Schools need clear, strong internal focus on issues of instruction, student learning, and expectations for teacher and student performance. There is a high degree of alignment among individual teachers about what they can do and about their responsibility for the improvement of student learning. Changing the thought process of individuals will change the behaviors as well. Covey (2006) says “We judge ourselves by our intentions and others by their behavior” (p. 13). Effective leaders ensure intentions and behavior are closely aligned.

Some leaders may go the other direction and tap into others’ emotion of fear. Fullan (2006) says fear is not a powerful motivator. It only creates an initial immediate effect. Leaders use their power and authority to get compliance through fear. This leads to short-term change, but does not create the internal accountability. Teachers might work against fear of discipline, disappointment, or fear of failure. Each can elicit a different response. Though Fullan (2006) says fear is not a motivator, he says the possibility of failure is the most uncomfortable phenomenon in American life. When you look at human responses to rewards, there is no room for failure. We work for the reward system and leave no gray area to fail. Fear can elicit a response, but it cannot be the sole

driver to lead to long-term change. Instead, leaders focus on relationships built upon more positive emotions such as trust. If leaders rely on fear to change behavior, fear can become overwhelming and it can lead to more resistance. The discomfort of fear heightens the sense of losing control and increases reasoning as to why the change is not needed. Eventually, individuals' fear will lead them to the conclusion that the new will not be any better than the old (Weber, 2016).

Covey (2006) uses the analogy "fish discover water last" to make a point about creating trust (p.273). Humans discover trust last. Although trust plays a large part in the functioning of society, it is taken for granted unless it becomes polluted or destroyed. Yet it is trust that will make or break a leader in times of change. To build trust, leaders focus on character and competence. Character is comprised of integrity, motive, and intent with people. Competence includes capabilities, skills, results, and track record. To build high trust, both components are necessary. These two functions can be viewed from a variety of different theoretical frameworks. For example, Covey (2006) ties them to Leadership Theory through what a leader is (character) and what a leader does (competence). He also places them in Ethics Theory through doing the right thing (character) and getting the right thing done (competence). No matter what lens is looked through, trust can be applied through character and competence. Therefore, it is imperative leaders build trusting relationships through observed behaviors of character and competence.

If trust is broken or fails to be established during any change initiative, time and energy will need to be focused on rebuilding the relationships first, before the initiative can proceed. When trust is broken, it is often the hardest and most time-consuming

foundation to rebuild. Restoring trust involves changing someone else's feelings and their confidence in you. It is out of the leader's control how long it will take to restore the trust. It is not something that can be forced and will happen at different rates for each individual. Leaders make and keep commitments to individuals and the school as a whole. Commitments build hope but when those commitments are kept, it builds trust (Covey, 2006).

Though trust is a heavy lifter in times of a change initiative, respect, positivity, and listening to understand rather than to reply are also key factors (Peeples, 2016). The best plans can be developed with the best of intentions but interactions with stakeholders will impact their attitudes towards the change. John Kotter has said "The central issue is never strategy or structure. [It] is always about changing the behavior of people" (Fullan, 2006, p.36). In order to change behavior, people's opinions are heard and respected. Leaders focus on building capacity and utilizing the strengths of all involved. This will raise the bar and close the gap of student learning for all students. Surowiecki (2004) says four conditions must be met in order to activate the collective wisdom of a group: members of the organization must feel independent of one another, members need to be diverse to represent the range of people in the larger population group, members need to specialize and draw on independent knowledge, and independent judgment and information has to be transformed in collective decisions. All four conditions can be tied to honoring each individual's talent for the collective purpose of strategic decision-making.

To be effective, the collective decision-making of the group must be laser-focused on the good of the whole. Leaders are often comfortable with being a change agent.

They are willing to leap before anyone else to take a risk with trying something new for a necessary change (Peeples, 2016). But to be successful as a change agent, leaders demonstrate the ability to get others on board. True meaningful change will happen only if leaders begin to give up control and establish a culture built on trust and respect (Sheninger, 2016). The main goal of a change initiative comes down to coherence. Working collaboratively as one unit is the key driver in shifting behavior. Once members are emotionally ready, leaders can build coherence through purposeful action and interaction, clarity, precision of practice, transparency, monitoring progress, and continuous correction. It is a fine balance of pressure and support (Fullan & Quinn, 2016).

One way to honor members' strengths is to acknowledge the critical role teachers play in the school's decision-making process. Griffiths & Portelli (2016) encourage a high "leadership density". A high density means a larger number of people are involved in the work of others, are trusted with information, are involved in decision-making, are exposed to new ideas and participate in the knowledge creation and transfer. A large number of individuals play a key role in the success of the school and a collective culture. It is imperative teachers feel empowered to speak about the school and change initiatives. When teachers come together to collaborate, it provides the opportunity for others to point out blind spots. Isolation makes it difficult to see blind spots (Fiarman, 2016). Ultimately, teachers are the gatekeepers when it comes to change and there is a current shift to enable them to be on the front lines of change. It is no longer about passing responsibility from the top down. Meaningful change always begins at the individual

level. This is also where it is sustained so that it becomes embedded in the culture (Sheninger, 2016).

Beginning at the individual level to build cohesiveness, there are stakeholders that avoid change. Not all people live in a constant state of improvement. Knowing that change is uncomfortable for many, preparing for and preventing the negative consequences of important, necessary changes will better support students (Weber, 2016). Covey (2006) addresses such individuals, by using a teacher of fifteen years as an example. Instead of having fifteen years of experience, she has only one year of experience repeated fifteen times. She is not adapting to the changes required to stay relevant and often becomes obsolete. Others outgrow her. Covey says, “Nothing fails like success” (p. 102). In other words, she keeps doing the things that made her successful in yesterday’s market. It is the responsibility of the leaders to help those individuals rekindle their moral purpose and work to become agents of change. Hargreaves (2003) urges these individuals to extend their professional responsibility beyond their own classrooms to school-wide and district-wide efforts that improve the system as a whole. To expect improvements in student outcomes, initiate changes in practices (Weber, 2016). If the resistance to change goes deeper into the leadership team, Covey (2006) reminds us that if one does not like change, he is going to like irrelevance even less.

Though the thought of building coherence is positive, it can also bring about additional challenges. One of the main challenges to building coherence is leadership turnover. When new leaders come in, they bring their own vision and agenda. It is not the turnover that challenges the school’s coherence but more the discontinuity of

direction. This is why true coherence has to be achieved at the receiving end, not the delivery end. Even with leadership turnover, the school should be able to sustain change initiatives due to relentlessly pursuing ambitious goals and reducing the number of distractors. Student and other data is shared transparently and strategies are built in for implementers to learn from each other. If members or leaders change, the path to the goals should be clearly visible and still on track. These are signs of a true systemness in place where coherence can be sustained (Fullan & Quinn, 2016).

Besides leadership turnover and change of vision, there are other barriers that can prevent a change initiative from moving forward. Fullan (2006) cautions to avoid five main barriers to action. First, ensure talking and planning are not substitutes for action. It is not infrequent that many meetings are held with little action to follow. Plans cannot be effective unless they are put into action. Second, avoid memory being a substitute for thinking. When teachers say “we have always done it this way”, it is likely the purpose has been lost and never re-evaluated. Next, do not let fear prevent acting on knowledge. Though all information may be pointing in a clear direction, the fear of change can prevent one from using current knowledge to make a better, informed decision. This also leads to the next barrier of ensuring measurement does not obstruct good judgment. Data can be misleading and not tell the whole story. Good decisions are made when data, judgment, and the whole picture are taken into consideration. Finally, avoid internal competition turning friends into enemies. The focus is on the good of the group instead of gains of individuals.

Fullan provided five barriers to action. When analyzing the coherence framework from Fullan & Quinn (2016), there are four wrong and four right drivers to whole system

change. Strategies to avoid are: focusing on individuals, employing punitive accountability, using technology only at the acquisition level, and having ad hoc policies create silo problems (each ad hoc is working on solving their own problems individually instead of working together). Instead, focus on the four right drivers to achieve whole system change: capacity building, collaboration, pedagogy (what good teachers like to do every day, remembering that technology is not a pedagogy, it is an accelerator to pedagogy), and systemness.

Ultimately, an effective leader will lead change through an ethical lens. He needs a solid handle on the school's stability and vulnerability. He must weigh the options and determine if a decision will create some change to overcome complacency (Griffiths & Portelli, 2015). Leading change is about shaping and reshaping good ideas. It is about evaluating the quality of an idea and the quality of the process. It is realizing that the best way to truly understand something comes from trying to change it (Fullan & Quinn, 2016). Ethical change leadership focuses on reducing bad variation within schools (Fullan, 2006).

Another lens to view ethical leadership is through Starratt's (2009) framework. Instead of focusing on one single ethic, he focuses on three that are interconnected: ethic of care, ethic of justice, and ethic of critique. The ethic of care is concerned with dignity and worth for all. It is acknowledging each person for who they are and displaying a loyalty to the relationship. The core is the relationships between school leaders, students, teachers, and the community members. All voices are heard and valued. The ethic of justice addresses fair and equitable treatment of people. There is a shared sense of community with all involved. There is also a focus on the legal aspects since fair and

equitable practices are contained in laws, rights, and policies. Finally, the ethic of critique encourages questioning current policies and practices in an attempt to uncover injustice and challenge inequality. The school community embraces a sense of social responsibility to review all aspects of the school to ensure equity for all.

Ultimately, leaders of ethical change take the necessary steps to create the vision while using collaboration as the means to keep the change initiative moving forward. Motivation, perseverance, and tenacity will feed the success of the team and create confidence in not only the leadership team, but also all stakeholders. Confidence influences the willingness to invest and create team coherence (Fullan, 2006). As with any change, leaders anticipate obstacles and plan for ways to overcome those obstacles. Be clear on what the change represents and what the change does not represent. Give time to study the why and what of change, provide opportunities to voice opinions, plan for professional development, and measure success (Weber 2016). The role of the change agent is to provide relevancy, meaning, and authenticity (Sheninger, 2016).

Evidence of Change

In order to determine if change has occurred, stakeholders rely on results. Results are clearly defined before any processes or procedures are put into place. Stakeholders collectively decide how and when to measure the effectiveness of the change (Weber, 2016). Covey (2006) uses a tree as an analogy to achieving results. “Results are the fruits of a tree....the tangible, measurable end purpose and product of the roots, trunk and branches. To have everything but results is like having a barren tree. It won’t create credibility (for the leader) no matter how strong other forces are. And it won’t inspire

confidence because the tree doesn't produce what it was intended to produce" (p. 100). Results are measurable, tangible, and prove the change has occurred.

One way to seek proof of change is to observe staff's behaviors. Campbell (2005) believes teachers simply accept that the best way to keep out of trouble in schools is to avoid challenging colleagues on matters of competence and ethical conduct. They learn to live with the guilt over their inaction. Instead, leaders focus on changing this quiet, hidden behavior and focus on stakeholders speaking up to contribute to the change. An atmosphere of psychological safety must exist so teachers feel they can speak up without fear of retribution. Decreasing the amount of lag time between someone identifying a problem and actually speaking up about it, helps to address the true problem in a timely manner. The longer those problems go unaddressed, the more time others will join the negativity (Miller, 2016). Empower others to speak up in a productive manner regarding change will collectively focus on addressing the issues that will have the greatest impact on students. When others are heard and problem behaviors are addressed, all stay on the path of working towards the shared vision.

However, simple compliance is not enough for change to occur. Leaders look beyond basic behaviors of compliance and seek for the inner change within individuals. There may be a teacher who is compliant and performs each task required to meet the standards of a new initiative. To move beyond tasks, it is vital for individuals to take responsibility for the results, not just the activities (Covey, 2006). Many times change includes lengthy planning and creating of documents and processes. Teachers get wrapped up in the planning instead of focusing on doing. Fullan (2006) urges leaders to

reduce the distance between planning and action. The actual documents are far less important than implementation, execution, and monitoring.

This leads to the key component of accountability. When evidence of results is being discussed in the beginning stages, so should accountability. Leaders make it clear what is expected and what will happen if the commitment does not occur. They also clearly communicate the supports that will be in place to hold others accountable (Weber, 2016). Accountability has a moral component associated with integrity and passion with a strong sense of equity and justice. It justifies what we all do (Griffiths & Portelli, 2015). However, not all staff members may come with this inner passion. It is inevitable that additional supports will be needed for some and those supports should provide a solid foundation for those needing additional assistance. In providing the supports, leaders model the expected behaviors. Cultures do not change by mandate. They change by displacing existing norms, structures, and processes. A change in culture depends fundamentally on modeling the new values and behavior expected to displace the existing one (Fullan, 2016). Association for Middle Level Education (2012) agrees with Fullan in stating, “The real difficulty in changing the course of any enterprise lies not in developing new ideas but in escaping from old ones” (p. 214).

Another way to seek proof of change is to collect qualitative data. Monitoring individuals’ word choice is a form of qualitative data. When individual teachers stop speaking about “my classroom” and start referring to “our school”, it is an indication teachers’ mindsets are shifting in the right direction (Fullan, 2006). Collect qualitative data on how people feel. Asking others to reflect on their thoughts and practices can provide meaningful feedback to leaders. Ultimately, the leader’s toughest job is to change

the reality of others. The leader has to ensure others believe in a new reality and have acquired a new belief system (Griffiths & Portelli, 2015). Dweck (2006) also supports the change in belief systems. The new beliefs take their place alongside the old ones. As those new beliefs become stronger and become the new norm, they start to change the way everyone thinks, feels, and acts. Behaviors will support the new belief system but simply asking teachers about their core beliefs can contribute to data to measure effectiveness.

If data does not show a change in shared vision, beliefs, and behaviors, it is important to stay the course. The initial gains in a turnaround situation are unsteady and unclear (Fullan, 2006). To expect an immediate shift during an uncomfortable time of change is setting all up for disappointment and failure. Professionals in education live in a constant state of improving. Dweck (2006) states, “Becoming is better than being” (p. 25). Leaders understand it can take a great deal of time to change the behaviors of people.

More important than changing behaviors, to sustain long-term change requires shifting the emotional and investment climate to create a new definition of motivation (Kanter, 2004). Change needs to be seen as a process, not an event. Educators cannot plan to arrive once and for all at that end goal. There are always new developments and innovations. It is a continuous process of making and reshaping mindsets and cultures (Fullan & Quinn 2016). Earl (2003) refers to the continuous process as “creeping incrementalism” (p. 15). Teachers make small changes at each step instead of making one major sweeping change. All the small changes that lead to big, sustained changes over time come back to the self-confidence in a leader’s skills. Leaders have the courage

to be open and welcomed to change. Demonstrating a readiness to grow (and doing whatever it takes to get others to demonstrate the same), will keep the school moving forward with change (Dweck, 2006).

Process of Initiating Change in Grading Expectations at the Research Site

In the 2013-2014 school year, I joined the middle school as an assistant principal. Some of my major roles and responsibilities revolved around curriculum and instruction. Taking my 11 years of experience as a classroom teacher and combining those experiences with my new role, grading procedures became an area of interest and I began doing more research on the topic. I began to read more on best practices in grading, specifically in middle school, and looking for ways to implement my new learning within our current systems and constraints. Though I was aware of inconsistencies within our building, I was still working my way into a new culture and learning about the current systems in place.

During the 2014-2015 school year, the school experienced a change in leadership. The administrative team then consisted of an assistant principal of two years, an assistant principal of one year, and a new principal. In three years' time, the administrative team had turned over. As the principal was meeting staff and talking about the systems in place through her transition period, one of the topics that was brought to her attention was the need to address grading discrepancies. Staff wanted some common practices (T. Perkins, personal communication, February 12, 2017). Based off feedback from multiple

sources, one of the systems the administrative team was wanting to shift in the building was that of grading expectations.

As an administrative team, we started by discussing some parameters we wanted teams and grade levels to stay within. We also wanted to have a clear vision of where we were headed without mandating how it would be done. Our team followed guidelines by AMLE (2012) such as starting with a vision that we could hold firm to in order for all students to reach their potential. Our vision also matched the district's middle school philosophy with AMLE (2012) by "being able to define and commit to norms that guide how the team operates" (p. 121). The goal was to have each team and grade level commit to common grading expectations, no matter the content taught by each teacher.

In preparation for creating the change, we anticipated some hesitation from some teachers. Not only did this new system represent change but it also challenged the status quo. Students come to the school high achieving in sixth grade. Student data shows high achievement, which some stakeholders may translate to evidence of no need to change what has always been done. However, this mindset does not consider the potential rate of student gains, nor acknowledge that students who are already proficient should be pushed to achieve at the next level (T. Perkins, personal communication, February 12, 2017). In addition to setting some administrative parameters, aligned to Board policy and building expectations, we created a document that provided teachers some talking points and questions to reflect upon (see appendix). These were the guiding tools to begin conversations with each grade level.

In the fall of 2015, grade levels gathered to have conversations about existing grading procedures and how to start coming together with common expectations. Teams

had multiple conversations and were asked to submit their expectations to implement for the school year. Some teams were ready before students arrived on the first day of school while others needed more time and were not ready on the first day. Teachers commented about being frustrated, feeling unprepared to start the year without a solidified syllabus in place. Eventually, through multiple meetings and conversations in the fall, each grade level came to a consensus on expectations they would follow.

Throughout the first year of implementation, the administrative team learned where teachers needed more support and guidance as well as what was working well. Reflecting on that first year, the administrators agreed on one mistake we made. We as leaders were not intentional with measuring along the way. Looking back, we should have scheduled specific check in times to monitor understanding and adjust as needed (T. Perkins, personal communication, February 12, 2017). The adjustment opportunity would have allowed for re-teaching of expectations with the adults, just as we would expect re-teaching of expectations with students.

This reflection from year one helped us to get off to a better start in year two of implementation. For example, conversations on teams' expectations started towards the end of the 2015-2016 school year to have them in place for fall of 2016. Teams agreed and submitted their expectations before leaving for the summer. This solved the problem of not having things in place when students arrived in the fall. It also freed up working time in the fall when teachers needed to be preparing for students instead of being in additional meetings to finalize grading expectations. In year 2, we also addressed the need to check in with staff along the way. In November of 2016, a google form survey was sent to all certificated staff asking a variety of questions to gain insights and collect

feedback on how the system is working. That data was used by administration to adjust staff development needs and plan for changes coming in the future. The survey sent in November 2016 is the instrument used for data collection for the purpose of this study.

Chapter 3: Methodology and Design

Study Design

Using the collected data, this study will contain mixed methods. An explanatory sequential design best suits this study since the starting point will be on-the-shelf quantitative data. Participants will be sorted by subgroups such as core teachers, exploratory teachers, other certificated staff, years of experience, grade level taught, content area, leadership experience, and experience teaching in more than one district. A Mann-Whitney Test or Kruskal-Wallis Test will be performed for each subgroup. Each group's median from Likert scale responses will be compared to the grand median of the whole. For example, in looking at the perceived level of fidelity of implementation, participants will be grouped according to the quantity of leadership experiences each teacher has had. The median of each of those sub groups will be compared to the median of all participants' rating of fidelity to determine if there is statistical significance. Then, qualitative data from the open-ended response items will be used to help explain some of the quantitative results. When analyzing free-response items from the survey, it will be determined if there are categorical justifications for the small population. Findings will indicate what categories emerge.

The results of the tests will be displayed according to subgroups in order to make comparisons amongst the groups. The qualitative data will then be used to help identify what characteristics and factors the group has in common that may be leading to the trend.

Also, the power of the data will lie in acknowledging the internal and external validity. With a small sample size (n=51) and all participants coming from the same building in one school district, the strength will come in comprehending the limitations of the study. Educational leaders will need to understand the context and culture in which this study takes place and understand the results cannot be extended to another building or district.

Population

The survey was sent to all (75) certificated staff members at the middle school research site. Of those 75 surveys, 50 were returned from a variety of content area and grade level teachers. There were 12 males and 39 females represented. Ages of participants ranged from 23-65 while the years of experience in education ranged from 2-37 years and the years at the specific research site ranged from less than one year to 25 years.

Instrumentation

The instrument used for this study was a survey used to provide building-level feedback to the administrative team at the research site. It was created with input from each administrator to dig deeper into data regarding the school's common grading expectations during year two of implementation. After participants identified their grade level and category of teaching assignment (core, exploratory, other certificated staff), the survey directed them to Likert scale questions. It was comprised of 10 Likert scale questions with responses of 1 (strongly disagree), 2 (disagree), 3 (agree), and 4 (strongly agree). Then participants answered 13 free response questions to record individualized

answers and expand on some of the Likert scale questions. The survey ends with basic demographic information.

After the survey was created, it was tested with six teachers who were not on staff at the research site. Three teachers each came from two different middle schools in the same district to take the survey and provide feedback. Based on their written feedback and responses to questions, the survey was altered to address some possible concerns. For example, the original survey asked a question about equity and another about equality. Of the six teachers, five said they did not know the difference between the two questions and felt it was redundant. Terminology and question structures were edited based on the feedback from the test teachers.

Data Collection

The source of the data collection for this study came from on-the-shelf data. The administrative team utilized a survey that went to all certificated staff asking questions ranging from Likert scale ratings to full, open response explanations about grading philosophies during the second year of implementing common grading expectations. Since the data was used at the building level to gain feedback and make administrative decisions, the data will be analyzed again through a different lens. All names of staff members were collected during the initial survey and will be deleted from the data for the purpose of this study. As the study is focusing on teacher perceptions, a survey is the best tool to gain insight of each individual's philosophies.

The survey was sent using a Google Form to certificated staff in mid-November, 2016. Staff had two weeks to complete it. Time usually dedicated to other meetings in

those two weeks were adjusted to provide time for all certificated staff to complete the survey. Multiple reminders were sent through email and the weekly communication document shared with all staff. The spreadsheet of submissions was shared only with the administrative team to view responses and sort data for building-level decision-making. At the building level, the data was viewed with individual names tied to responses. Names will be removed for the purpose of this study to analyze the big picture.

Research Questions

The central research question to be addressed is how do teachers formulate their own grading practices and what are their perceptions of implementing common grading expectations? Within the study, the following sub-questions will be answered:

1. What factors do teachers report impact their current grading philosophy and how do the findings vary among subgroups including grade level taught, core teachers, exploratory teachers, other certificated staff, years of experience, content area, leadership experience, and experience teaching in more than one district?
2. What are teachers' perceptions of implementation fidelity and how do the findings vary among subgroups including grade level taught, core teachers, exploratory teachers, other certificated staff, years of experience, content area, leadership experience, and experience teaching in more than one district?
3. What are teachers' perceptions of common grading expectations and how do the findings vary among subgroups including grade level taught, core teachers, exploratory teachers, other certificated staff, years of experience, content area, leadership experience, and experience teaching in more than one district?

Data Analysis

The survey results will be used to analyze groups of Likert scale responses and open-ended responses. To protect individual responses, the subgroups will include ranges to avoid unique identifying factors. For example, if a teacher has 30 years of experience, all teachers in the building would be able to identify that individual by one simple piece of information. Therefore, teachers will be grouped into ranges such as years of experience (0-3, 4-7, etc.) and those groups will be intentionally created to minimize the risk of identifying an individual. If the groups still allow for easy identification, suppression will be used to avoid identifying that individual.

The data analysis will be organized according to the three research questions posed. Below outlines specific Likert scale and qualitative open-ended questions from the survey that will be used to answer each research question.

Research Question #1: What factors do teachers report impact their current grading philosophy and how do the findings vary among subgroups including grade level taught, core teachers, exploratory teachers, other certificated staff, years of experience, content area, leadership experience, and experience teaching in more than one district?

Qualitative Questions

- Putting the common grading expectations aside, in 15 words or less, what is your personal grading philosophy?
- How did you arrive at your current philosophy of grading middle school students?
What has influenced you and your decisions?

- Was there a specific incident in your educational experience that influenced your personal beliefs and philosophy? Please describe the event.
- How has grading changed or not changed from when you were in school and you were graded as a student?

Research Question #2: What are teachers' perceptions of implementation fidelity and how do the findings vary among subgroups including grade level taught, core teachers, exploratory teachers, other certificated staff, years of experience, content area, leadership experience, and experience teaching in more than one district?

Likert Scale Statement

- I fully implement our grade level's grading procedures with fidelity.

Research Question #3: What are teachers' perceptions of common grading expectations and how do the findings vary among subgroups including grade level taught, core teachers, exploratory teachers, other certificated staff, years of experience, content area, leadership experience, and experience teaching in more than one district?

Likert Scale Statements

- "The current common grading expectations are what is best for students."
- "Our common grading expectations are fair and equitable for all students."
- "Our common grading expectations allow for students to take ownership for their own learning."
- "Our common grading expectations lead to grades that clearly reflect students' demonstration of knowledge."

- “The common grading expectations meet the needs of middle school students.”

Qualitative Questions

- How does your personal philosophy align with your grade level’s common grading expectations?
- How does your personal philosophy differ from your grade level’s common grading expectations?
- What do you find most effective about the common grading expectations?
- What do you find least effective about the common grading expectations?
- What outcomes have you observed that can be attributed to the common grade expectations (such as changes in effort, soft skills, student ownership of learning, relationships, etc.)?
- When you think of your grade level’s common grading expectations, what three adjectives come to mind?

For the research questions that have both Likert scale and open-ended, qualitative responses, the quantitative data will be analyzed using either a Mann-Whitney Test or a Kruskal-Wallis Test, depending on the sample size. Then the qualitative data from the open-ended response items will be used to help explain some of the quantitative results. When looking at open-ended response items from the survey, it will be determined if there are justifications for the test results for this specific small population.

Chapter 4: Results

For each research question listed below, a Mann-Whitney test was conducted when there were only two groups to compare. This test was used due to having no distribution assumptions and comparing only two groups. When there were three or more groups, a Kruskal-Wallis Test was conducted. This test was selected over a one-way Anova to avoid meeting the assumptions of a normal distribution with equal variance across groups. Though Kruskal-Wallis is not as powerful as an Anova, it meets the needs of this research. In viewing the results, the asymptotic significance (Asymp. Sig.) was used on the larger samples as the approximate p-value based on the chi-square statistic instead of an exact p-value.

Due to each test being a median test, all will have the same null and alternate hypothesis of:

$H_0 =$ *All medians are equal.*

$H_a =$ *At least one median is different.*

The common grading expectations created at this middle school resulted from a collaborative effort among teams and grade levels to create a system of grading that was equitable for all students across the grade level. After one year of implementation, certificated staff members were surveyed using Likert Scale questions on a scale of 1-4 as well as open-ended questions to gain further insight into their thoughts on grading philosophy, implementation, and perceptions of the practices in place.

Research Question #1: What factors do teachers report impact their current grading philosophy and how do the findings vary among subgroups including grade level taught, core teachers, exploratory teachers, other certificated staff, years of experience, content area, leadership experience, and experience teaching in more than one district?

Research question #1 was measured by four open-ended, qualitative questions. In reviewing the responses, there were no significant differences among any subgroups in how they arrived at their current grading philosophy except for the years of experience group. This group identified more with responses tied to their years of experience in the classroom. Otherwise, subgroups did not show significant differences but more overarching ideas of the population did evolve.

Qualitative Questions

Putting the common grading expectations aside, in 15 words or less, what is your personal grading philosophy?

There were some clear themes that came out in teachers' responses. Thirty-four teachers mentioned grades needing to reflect knowledge and learning. For example, one teacher said, "Grading is a tool for teachers/parents/students to understand student progress". Three teachers spoke specifically to grades being a snapshot in time and should tell the story of a student's learning journey. Nine teachers focused on grades being a form of accountability and students taking ownership for their learning (the word "ownership" was a focus point in the school's mission statement). Other responses that were recorded included grades representing effort, focused on assessment performance,

the need for redos at the middle level, and two teachers responded they would like to get rid of grades all together to focus more on learning and less on grades.

How did you arrive at your current philosophy of grading middle school students?

What has influenced you and your decisions?

Most teachers reported their grading philosophies were developed from their classroom experiences as a teacher. Some responded with simple statements about teaching successfully for a certain number of years while others combined years of experience with other reasons. One teacher responded with “influence from experience and building expectations” while others reported a combination of experience and understanding relationships with students and families. Of the fifty respondents, twelve referenced their philosophy coming from professional development opportunities or reading about best practice from researchers such as Rick Wormelli and Ken O’Connor. Two teachers referenced an experience they had as a student which shaped their philosophy while two other teachers discussed the importance of what they would want for their own children in school.

Was there a specific incident in your educational experience that influenced your personal beliefs and philosophy? Please describe the event.

Responses to this question seem to be in conflict with the previous question. Here, only 15 teachers said their experience as a teacher is what influenced personal beliefs and philosophies. Some referenced experiences with meeting individual student needs with differentiation while others commented on student responses to grading philosophies. One teacher cited, “Watching apathy grow when there are retakes” as an

outcome that influenced grading philosophy. Thirteen teachers acknowledged there was not one specific incident they could pinpoint. Five teachers discussed outside pressures from parents and administration. One teacher said, “There was no particular event; however, constant and consistent external pressures (parents, other teachers, administrators, district and building philosophy, our current culture, etc.) have shaped and reshaped my philosophy and practice”. Five teachers referenced professional readings or staff development, mostly citing *Grading For Learning* by Ken O’Connor, which was a district initiative over the years. Only two teachers talked about a college course that assisted in shaping grading philosophy.

How has grading changed or not changed from when you were in school and you were graded as a student?

By far, the number one theme on this question revolved around redoing work and retaking assessments. Thirty-three teachers referenced an “excessive” amount of retake opportunities that did not exist when they were students. Four teachers said there is less accountability for students now and the burden has shifted to teachers to ensure students earn higher grades. Two teachers said technology has increased accessibility to grades, making the grading process more transparent. Students and parents have easy access to check on grades and it is no longer a secret until the end of the grading period as it once used to be. Other changes cited by a few teachers were different expectations of parents at home in supporting academics, grades being less vague now with needing more evidence to support a grade, and teachers having less autonomy to grade individually for each classroom.

The open-ended questions allowed staff to elaborate on how they arrived at their own beliefs on grading and reflect on the common grade expectations. Though the common grading expectations came through a consensus, further investigations were done to seek feedback on the fidelity of implementation.

Research Question #2: What are teachers' perceptions of implementation fidelity and how do the findings vary among subgroups including grade level taught, core teachers, exploratory teachers, other certificated staff, years of experience, content area, leadership experience, and experience teaching in more than one district?

Research question #2 was answered by one Likert scale survey question. Staff responded to the statement "I fully implement our grade level's grading procedures with fidelity". Following is the analysis of the one Likert scale question by each subgroup. There were no significant differences among any subgroups in analyzing the perceptions of implementation fidelity.

Table 4.1 Overview of Median Frequencies by Teacher Type (Core, Exploratory, and Other) on Perceptions of Implementation Fidelity

		Teacher		
		Core	Exploratory	Other
I fully implement our grade level's procedures with fidelity. <i>Mdn</i> = 3, on 4-point scale	> Median	11	6	0
	<= Median	22	5	0

Table 4.2 Kruskal-Wallis by Teacher Type (Core, Exploratory, and Other) on Perceptions of Implementation Fidelity

	I fully implement our grade level's procedures with fidelity
<i>N</i>	44
Median	3.00
Chi-Square	1.566
<i>df</i>	1
Asymp. Sig.	0.211

Medians of the core, exploratory, and other certificated staff were not significant on perceptions of, "I fully implement our grade level's procedures with fidelity," $\chi^2(1, N=44) = 1.566, p = 0.211$.

Table 4.3 Overview of Median Frequencies by Grade Level on Perceptions of Implementation Fidelity

		Grade Level			
		6th	7th	8th	Multi
I fully implement our grade level's procedures with fidelity. <i>Mdn</i> = 3, on 4-point scale	> Median	3	7	1	6
	<= Median	7	5	10	5

Table 4.4 Kruskal-Wallis by Grade Level on Perceptions of Implementation Fidelity

	I fully implement our grade level's procedures with fidelity
<i>N</i>	44
Median	3.00
Chi-Square	7.503
<i>df</i>	3
Asymp. Sig.	0.057

Medians of the 6th grade, 7th grade, 8th grade, and multi-grade level staff were not significant on perceptions of, "I fully implement our grade level's procedures with fidelity," $\chi^2(3, N=44) = 7.509, p = 0.057$.

Table 4.5 Overview of Median Frequencies by Years of Experience on Perceptions of Implementation Fidelity

		Years		
		0-10	11-21	22+
I fully implement our grade level's procedures with fidelity. <i>Mdn</i> = 3, on 4-point scale	> Median	7	7	3
	<= Median	9	10	8

Table 4.6 Kruskal-Wallis by Years of Experience on Perceptions of Implementation Fidelity

	I fully implement our grade level's procedures with fidelity
<i>N</i>	44
Median	3.00
Chi-Square	.822
<i>df</i>	2
Asymp. Sig.	0.663

Medians of staff with 0-10, 11-21, and 22+ years of experience were not significant on perceptions of "I fully implement our grade level's procedures with fidelity," $\chi^2(2, N=44) = 0.822, p = 0.663$.

Table 4.7 Overview of Median Frequencies by Years of Content on Perceptions of Implementation Fidelity

		CONTENT				
		Math	ELA	Science	Social Studies	Other
I fully implement our grade level's procedures with fidelity. <i>Mdn</i> = 3, on 4-point scale	> Median	3	2	4	2	6
	<= Median	5	9	2	4	7

Table 4.8 Kruskal-Wallis by Content on Perceptions of Implementation Fidelity

	I fully implement our grade level's procedures with fidelity
<i>N</i>	44
Median	3.00
Chi-Square	4.315
<i>df</i>	4
Asymp. Sig.	0.365

Medians of Math, ELA, Science, Social Studies, and Other content areas were not significant on perceptions of "I fully implement our grade level's procedures with fidelity," $\chi^2(4, N=44) = 4.315, p = 0.365$.

Table 4.9 Mann-Whitney Test Comparing Teachers Who Have District Leadership Experience and Teachers Who Do Not Have District Leadership Experience on Perceptions of Implementation Fidelity

	I fully implement our grade level's procedures with fidelity
Mann-Whitney <i>U</i>	179.500
Wilcoxon <i>W</i>	245.500
<i>Z</i>	-0.063
Asymp. Sig. (2-tailed)	0.950
Exact Sig. [2*(1-tailed Sig.)]	.957

Perceptions of “I fully implement our grade level’s procedures with fidelity” were not significantly different for teachers with district leadership experience ($Mdn=3$) than teachers with no district leadership experience ($Mdn=3$), $U=179.500$, $p=0.957$.

Table 4.10 Mann-Whitney Test Comparing Teachers Who Have Worked in Other Districts and Teachers Who Have Not Worked in Other Districts on Perceptions of Implementation Fidelity

	I fully implement our grade level's procedures with fidelity
Mann-Whitney U	184.000
Wilcoxon W	355.000
Z	-1.391
Asymp. Sig. (2-tailed)	0.164

Perceptions of “I fully implement our grade level’s procedures with fidelity” were not significantly different for teachers who have worked in other districts ($Mdn=3$) than teachers who have not worked in other districts ($Mdn=3$), $U=184.000$, $p=0.164$.

The previous data supports the implementation of the common grading expectations. However, it is possible staff can implement a process with fidelity without complete buy-in. Digging deeper into overall perceptions of the common grading expectations may reveal differences that cannot be seen in a statistical analysis.

Research Question #3: What are teachers' perceptions of common grading expectations and how do the findings vary among subgroups including grade level taught, core teachers, exploratory teachers, other certificated staff, years of experience, content area, leadership experience, and experience teaching in more than one district?

Since this is the most robust research question, it is analyzed by both Likert scale survey questions and qualitative open-ended questions. As a reminder, the Likert scale statements used for the quantitative survey questions tied to research question #3 are:

- “The current common grading expectations are what is best for students.”
- “Our common grading expectations are fair and equitable for all students.”
- “Our common grading expectations allow for students to take ownership for their own learning.”
- “Our common grading expectations lead to grades that clearly reflect students' demonstration of knowledge.”
- “The common grading expectations meet the needs of middle school students.”

Using the Likert scale survey questions, there was no significant difference in subgroups when analyzing perceptions of common grading expectations.

Table 4.11 Overview of Median Frequencies by Teacher Type (Core, Exploratory, and Other) on Perceptions of Common Grading Expectations

		Teacher		
		Core	Exploratory	Other
The current common grading practices are what is best for students. <i>Mdn</i> = 3 on 4-point scale	> Median	3	2	0
	<= Median	31	10	4
Our common grading practices are fair/equitable for all students. <i>Mdn</i> = 3, on 4-point scale	> Median	6	4	0
	<= Median	28	8	8
Our common grading practices allow students to take ownership for their own learning. <i>Mdn</i> = 3, on 4-point scale	> Median	6	3	0
	<= Median	28	9	4
Our common grading practices lead to grades that clearly reflect students' demonstration of knowledge. <i>Mdn</i> = 3, on 4-point scale	> Median	3	2	0
	<= Median	31	10	4
The common grading practices meet the needs of middle school students. <i>Mdn</i> = 3, on 4-point scale	> Median	3	3	0
	<= Median	31	9	4

Table 4.12 Kruskal-Wallis by Teacher Type (Core, Exploratory, and Other) on Perceptions of Common Grading Expectations

	The current common grading practices are what is best for students	Our common grading practices are fair/equitable for all students	Our common grading practices allow students to take ownership for their own learning	Our common grading practices lead to grades that clearly reflect students' demonstration of knowledge	The common grading practices meet the needs of middle school students
<i>N</i>	50	50	50	50	50
Median	3.00	3.00	3.00	3.00	3.00
Chi-Square	1.089	2.451	1.279	1.089	2.791
<i>df</i>	2	2	2	2	2
Asymp. Sig.	0.580	0.294	0.527	0.580	0.248

There was no significant difference among core teachers, exploratory teachers, and other certified staff on perceptions of common grading expectations.

Medians of the core, exploratory, and other certificated staff were not significant on perceptions of, "The common grading practices are what is best for students," $\chi^2(2, N=50) = 1.089, p = 0.580$.

Medians of the core, exploratory, and other certificated staff were not significant on perceptions of, "Our common grading practices are fair/equitable for all students," $\chi^2(2, N=50) = 2.451, p = 0.294$.

Medians of the core, exploratory, and other certificated staff were not significant on perceptions of “Our common grading practices allow students to take ownership for their own learning,” $\chi^2(2, N=50) = 1.279, p = 0.527$.

Medians of the core, exploratory, and other certificated staff were not significant on perceptions of “Our common grading practices lead to grades that clearly reflect students’ demonstration of knowledge,” $\chi^2(2, N=50) = 1.089, p = 0.580$.

Medians of the core, exploratory, and other certificated staff were not significant on perceptions of “The common grading practices meet the needs of middle school students,” $\chi^2(2, N=50) = 2.791, p = 0.248$.

Table 4.13 Overview of Median Frequencies by Grade Level on Perceptions of Common

Grading Expectations

		Grade Level			
		6th	7th	8th	Multi
The current common grading practices are what is best for students. <i>Mdn</i> = 3, on 4-point scale	> Median	1	2	0	2
	<= Median	9	11	11	14
Our common grading practices are fair/equitable for all students. <i>Mdn</i> = 3, on 4-point scale	> Median	3	3	0	4
	<= Median	7	10	11	12
Our common grading practices allow students to take ownership for their own learning. <i>Mdn</i> = 3, on 4-point scale	> Median	2	3	1	3
	<= Median	8	10	10	13
Our common grading practices lead to grades that clearly reflect students' demonstration of knowledge. <i>Mdn</i> = 3, on 4-point scale	> Median	2	1	0	2
	<= Median	8	12	11	14
The common grading practices meet the needs of middle school students. <i>Mdn</i> = 3, on 4-point scale	> Median	2	1	0	3
	<= Median	8	12	11	13

Table 4.14 Kruskal-Wallis by Grade Level on Perceptions of Common Grading Expectations

	The current common grading practices are what is best for students	Our common grading practices are fair/equitable for all students	Our common grading practices allow students to take ownership for their own learning	Our common grading practices lead to grades that clearly reflect students' demonstration of knowledge	The common grading practices meet the needs of middle school students
<i>N</i>	50	50	50	50	50
Median	3.00	3.00	3.00	3.00	3.00
Chi-Square	1.752	3.702	.852	2.521	3.025
<i>df</i>	3	3	3	3	3
Asymp. Sig.	0.625	0.296	0.837	0.471	0.388

There was no significant difference among 6th grade, 7th grade, 8th grade, and multi-grade level staff on perceptions of common grading expectations.

Medians of the 6th grade, 7th grade, 8th grade, and multi-grade level staff were not significant on perceptions of “The common grading practices are what is best for students,” $\chi^2(3, N=50) = 1.752, p = 0.625$.

Medians of the 6th grade, 7th grade, 8th grade, and multi-grade level staff were not significant on perceptions of “Our common grading practices are fair/equitable for all students,” $\chi^2(3, N=50) = 3.702, p = 0.296$.

Medians of the 6th grade, 7th grade, 8th grade, and multi-grade level staff were not significant on perceptions of “Our common grading practices allow students to take ownership for their own learning,” $\chi^2(3, N=50) = 0.852, p = 0.837$.

Medians of the 6th grade, 7th grade, 8th grade, and multi-grade level staff were not significant on perceptions of “Our common grading practices lead to grades that clearly reflect students’ demonstration of knowledge,” $\chi^2(3, N=50) = 2.521, p = 0.471$.

Medians of the 6th grade, 7th grade, 8th grade, and multi-grade level staff were not significant on perceptions of “The common grading practices meet the needs of middle school students,” $\chi^2(3, N=50) = 3.025, p = 0.388$.

Table 4.15 Overview of Median Frequencies by Years of Experience on Perceptions of Common Grading Expectations

		Years		
		0-10	11-21	22+
The current common grading practices are what is best for students. <i>Mdn</i> = 3, on 4-point scale	> Median	2	3	0
	<= Median	15	19	11
Our common grading practices are fair/equitable for all students. <i>Mdn</i> = 3, on 4-point scale	> Median	3	7	0
	<= Median	14	15	11
Our common grading practices allow students to take ownership for their own learning. <i>Mdn</i> = 3, on 4-point scale	> Median	2	6	1
	<= Median	15	16	10
Our common grading practices lead to grades that clearly reflect students' demonstration of knowledge. <i>Mdn</i> = 3, on 4-point scale	> Median	2	3	0
	<= Median	15	19	11
The common grading practices meet the needs of middle school students. <i>Mdn</i> = 3, on 4-point scale	> Median	2	4	0
	<= Median	15	18	11

Table 4.16 Kruskal-Wallis by Years of Experience on Perceptions of Common Grading Expectations

	The current common grading practices are what is best for students	Our common grading practices are fair/equitable for all students	Our common grading practices allow students to take ownership for their own learning	Our common grading practices lead to grades that clearly reflect students' demonstration of knowledge	The common grading practices meet the needs of middle school students
<i>N</i>	50	50	50	50	50
Median	3.00	3.00	3.00	3.00	3.00
Chi-Square	1.604	4.729	2.321	1.604	2.297
<i>df</i>	2	2	2	2	2
Asymp. Sig.	0.448	0.094	0.313	0.448	0.317

There was no significant difference among teachers with varying years of experience on perceptions of common grading expectations.

Medians of staff with 0-10, 11-21, and 22+ years of experience were not significant on perceptions of “The common grading practices are what is best for students,” $\chi^2(2, N=50) = 1.604, p = 0.448$.

Medians of staff with 0-10, 11-21, and 22+ years of experience were not significant on perceptions of “Our common grading practices are fair/equitable for all students,” $\chi^2(2, N=50) = 4.729, p = 0.094$.

Medians of staff with 0-10, 11-21, and 22+ years of experience were not significant on perceptions of “Our common grading practices allow students to take ownership for their own learning,” $\chi^2(2, N=50) = 2.321, p = 0.313$.

Medians of staff with 0-10, 11-21, and 22+ years of experience were not significant on perceptions of “Our common grading practices lead to grades that clearly reflect students’ demonstration of knowledge,” $\chi^2(2, N=50) = 1.604, p = 0.448$.

Medians of staff with 0-10, 11-21, and 22+ years of experience were not significant on perceptions of “The common grading practices meet the needs of middle school students,” $\chi^2(2, N=50) = 2.297, p = 0.317$.

Table 4.17 Overview of Median Frequencies by Content on Perceptions of Common Grading Expectations

		Content				
		Math	ELA	Science	Social Studies	Other
The current common grading practices are what is best for students. <i>Mdn</i> = 3, on 4-point scale	> Median	1	0	1	1	2
	<= Median	7	11	5	5	17
Our common grading practices are fair/equitable for all students. <i>Mdn</i> = 3, on 4-point scale	> Median	1	1	2	2	4
	<= Median	7	10	4	4	15
Our common grading practices allow students to take ownership for their own learning. <i>Mdn</i> = 3, on 4-point scale	> Median	1	3	2	0	3
	<= Median	7	8	4	6	16
Our common grading practices lead to grades that clearly reflect students' demonstration of knowledge. <i>Mdn</i> = 3, on 4-point scale	> Median	1	1	1	0	2
	<= Median	7	10	5	6	17
The common grading practices meet the needs of middle school students. <i>Mdn</i> = 3, on 4-point scale	> Median	0	0	2	1	3
	<= Median	8	11	4	5	16

Table 4.18 Kruskal-Wallis by Content on Perceptions of Common Grading Expectations

	The current common grading practices are what is best for students	Our common grading practices are fair/equitable for all students	Our common grading practices allow students to take ownership for their own learning	Our common grading practices lead to grades that clearly reflect students' demonstration of knowledge	The common grading practices meet the needs of middle school students
<i>N</i>	50	50	50	50	50
Median	3.00	3.00	3.00	3.00	3.00
Chi-Square	1.876	2.446	3.140	1.034	5.559
<i>df</i>	4	4	4	4	4
Asymp. Sig.	0.759	0.654	0.535	0.905	0.235

There was no significant difference among different content teachers on perceptions of common grading expectations.

Medians of Math, ELA, Science, Social Studies, and Other content areas were not significant on perceptions of “The common grading practices are what is best for students,” $\chi^2(4, N=50) = 1.876, p = 0.759$.

Medians of Math, ELA, Science, Social Studies, and Other content areas were not significant on perceptions of “Our common grading practices are fair/equitable for all students,” $\chi^2(4, N=50) = 2.446, p = 0.654$.

Medians of Math, ELA, Science, Social Studies, and Other content areas were not significant on perceptions of “Our common grading practices allow students to take ownership for their own learning,” $\chi^2(4, N=50) = 3.140, p = 0.535$.

Medians of Math, ELA, Science, Social Studies, and Other content areas were not significant on perceptions of “Our common grading practices lead to grades that clearly reflect students’ demonstration of knowledge,” $\chi^2(4, N=50) = 1.034, p = 0.905$.

Medians of Math, ELA, Science, Social Studies, and Other content areas were not significant on perceptions of “The common grading practices meet the needs of middle school students,” $\chi^2(4, N=50) = 5.559, p = 0.235$.

Table 4.19 Mann-Whitney Test Comparing Teachers Who Have Held District Leadership Positions and Those Who Have Not on Perceptions of Common Grading Expectations

	The current common grading practices are what is best for students	Our common grading practices are fair/equitable for all students	Our common grading practices allow students to take ownership for their own learning	Our common grading practices lead to grades that clearly reflect students' demonstration of knowledge	The common grading practices meet the needs of middle school students
Mann-Whitney <i>U</i>	249.000	218.500	244.500	225.500	249.500
Wilcoxon <i>W</i>	915.000	323.500	349.500	330.500	354.500
<i>Z</i>	-0.082	-0.793	-0.176	-0.696	-0.067
Asymp. Sig. (2-tailed)	0.934	0.428	0.861	0.486	0.947

There was no significant difference among teachers with district level leadership and teachers with no district level leadership on perceptions of common grading expectations.

Perceptions of “The common grading practices are what is best for students” were not significantly different for teachers with district level leadership experience ($Mdn=3$) than teachers with no district level leadership experience ($Mdn=3$), $U=249.000$, $p=0.934$.

Perceptions of “Our common grading practices are fair/equitable for all students” were not significantly different for teachers with district level leadership experience

($Mdn=3$) than teachers with no district level leadership experience ($Mdn=3$), $U=218.500$, $p=0.428$.

Perceptions of “Our common grading practices allow students to take ownership for their own learning” were not significantly different for teachers with district level leadership experience ($Mdn=3$) than teachers with no district level leadership experience ($Mdn=3$), $U=244.500$, $p=0.861$.

Perceptions of “Our common grading practices lead to grades that clearly reflect students’ demonstration of knowledge” were not significantly different for teachers with district level leadership experience ($Mdn=3$) than teachers with no district level leadership experience ($Mdn=3$), $U=225.500$, $p=0.486$.

Perceptions of “The common grading practices meet the needs of middle school students” were not significantly different for teachers with district level leadership experience ($Mdn=3$) than teachers with no district level leadership experience ($Mdn=3$), $U=249.500$, $p=0.947$.

Table 4.20 Mann-Whitney Test Comparing Teachers Who Have Worked In Other Districts and Those Who Have Not on Perceptions of Common Grading Expectations

	The current common grading practices are what is best for students	Our common grading practices are fair/equitable for all students	Our common grading practices allow students to take ownership for their own learning	Our common grading practices lead to grades that clearly reflect students' demonstration of knowledge	The common grading practices meet the needs of middle school students
Mann-Whitney <i>U</i>	285.000	267.500	278.000	261.500	230.000
Wilcoxon <i>W</i>	750.000	732.500	743.000	471.500	695.000
<i>Z</i>	-0.377	-0.705	-0.472	-0.927	-1.717
Asymp. Sig. (2-tailed)	0.706	0.481	0.637	0.354	0.086

There was no significant difference among teachers who have worked in other districts and teachers who have not worked in other districts on perceptions of common grading expectations.

Perceptions of “The common grading practices are what is best for students” were not significantly different for teachers who have worked in other districts ($Mdn=3$) than teachers who have not worked in other districts ($Mdn=3$), $U=285.000$, $p=0.706$.

Perceptions of “Our common grading practices are fair/equitable for all students” were not significantly different for teachers who have worked in other districts ($Mdn=3$) than teachers who have not worked in other districts ($Mdn=3$), $U=267.5000$, $p=0.481$.

Perceptions of “Our common grading practices allow students to take ownership for their own learning” were not significantly different for teachers who have worked in other districts ($Mdn=3$) than teachers who have not worked in other districts ($Mdn=3$), $U=278.000$, $p=0.637$.

Perceptions of “Our common grading practices lead to grades that clearly reflect students’ demonstration of knowledge” were not significantly different for teachers who have worked in other districts ($Mdn=3$) than teachers who have not worked in other districts ($Mdn=3$), $U=261.500$, $p=0.354$.

Perceptions of “The common grading practices meet the needs of middle school students” were not significantly different for teachers who have worked in other districts ($Mdn=3$) than teachers who have not worked in other districts ($Mdn=3$), $U=230.000$, $p=0.086$.

Each statistical analysis of the Likert scale questions showed no significant difference in perceptions of the common grading expectations. To allow for individual voices of teachers to come through, research question #3 is also analyzed through six open-ended questions. When staff were prompted with additional questions to reflect upon regarding those expectations, some differences were discovered. The qualitative results of the six open-ended survey questions are to follow.

Qualitative Results

How does your personal philosophy align with your grade level's common grading expectations?

Instead of giving specific examples, most teachers referenced a level of alignment between personal philosophy and the grade level's common grading expectations.

Twenty-four teachers commented the two are aligned and referenced their belief in valuing re-teaching and having the grade reflect knowledge. Ten teachers said the two are closely aligned but had some conflicts such as wanting to account for effort in a grade when the grade level expectation was to leave out effort in the grade. Three teachers said there is no match between the two and have had to find ways to resolve that internal conflict. In doing so, one teacher said, "Our common grading practices align with the least of my personal philosophy and is perhaps more generous than I would be but I understand the need of it", acknowledging the need for a system to be in place. She had to adjust her personal philosophy to align with the grade level expectations in order to function day to day. Finally, one teacher said, "I think it aligns with what is on the sheet but not with what is practiced". This statement leads one to believe what was agreed upon is not what was being practiced.

How does your personal philosophy differ from your grade level's common grading expectations?

Though thirty-four teachers said they were aligned in the previous question, this question showed only ten teachers saying they do not differ. Again, the overarching theme centered on retakes and redos. The grade level common grading expectations generally stated students could retake and redo for 100% credit. Teachers commented

this conflicts with their own philosophy. Some teachers wanted flexibility in making that decision on a case-by-case basis. Others wanted the retake to focus on learning the material instead of changing the grade. Teachers commented on wanting some sort of consequence tied to the grade for late or missing assignments. Three teachers wanted to de-emphasize grades all together and get back to being focused on learning. Finally, two teachers wanted to include non-academic factors in grades such as effort and timeliness.

What do you find most effective about the common grading expectations?

There were three main ideas that emerged from this question. By far, the biggest idea was one of consistency and fidelity. Thirty-nine teachers felt consistency was the most effective outcome from common grading expectations. Two of the thirty-nine teachers went on to explain the benefit for parents and students. They noted it was no longer a popularity contest to get the “easy teacher” and parents could no longer request a certain team due to their grading expectations being different than the other teams. It also allowed for consistency between siblings. For example, if twins or triplets were on different teams, all students were afforded the same opportunities and were graded on the same criteria. There were also five teachers that noted they felt the common grading practices were getting back to grades meaning something since they were communicating the same levels of achievement for students across the board. Grades were no longer being inflated due to extra credit or different grading criteria.

What do you find least effective about the common grading expectations?

Again, the overarching theme returns to retakes. Sixteen teachers commented on the effect retakes have on student effort. Some referenced students who put forth little

effort the first time and constantly use the safety net of retakes. To quote one teacher, “Retakes/redoes/re-learning, while valuable to many students, also reflect and/or contribute to a culture in which these concepts are a daily expectation and regular occurrence, oftentimes regardless of the individual effort of many students. Students plan for retakes before taking (or during) the first assessment. Also, many students are caught in a seemingly never-ending cycle of re-learning in multiple classes, while the class has moved on to new concepts...This all becomes a source of frustration for all involved”. Nine teachers were concerned the grading practices were too standardized and did not account for different teacher strengths and team dynamics. They were hoping for more autonomy to meet student needs. Three teachers felt the practices were not conducive to be implemented across all content levels and two teachers wanted to go back and include, or reward, effort.

What outcomes have you observed that can be attributed to the common grade expectations (such as changes in effort, soft skills, student ownership of learning, relationships, etc.)?

Multiple outcomes were identified through this question. Teachers felt it was a benefit that students knew the expectations and therefore, put forth more effort. However, it was only ten teachers that acknowledged this. Yet again, teachers expressed frustration with student effort the first time around due to the option to retake. Teachers used explanations such as, “Over time, I have seen effort and student ownership of learning diminish. Students feel that they can retake and redo all assignments, so many students do not put effort in the first time” and “I’ve seen a lack in effort, study skills and responsibility. I’ve also seen a sense of entitlement develop in both students and parents”.

Other outcomes mentioned were improved relationships with students, better communication amongst teachers, and one teacher said there is no observed outcomes yet because she feels they are not on the same page.

When you think of your grade level's common grading expectations, what three adjectives come to mind?

Though this was an open-ended question, a few words came up multiple times. "Consistent" and "common" were expressed eighteen times. Next in line was "fair" with twelve occurrences and "challenging" with six. For all the adjectives, words were sorted into categories based on references to teachers versus students. Words like "compromise", "frustrating", "idealistic", and "best practice" were in the teacher category as they referred to descriptions of how the grading practices impact teachers. In the student category, words such as "reflective", "enabling", and "student-centered" were included as they describe how the practices impact students. In the end, there were twenty different responses teachers used to describe common grading practices that would indicate an impact (positive or negative) on teachers. On the other hand, there were sixty different responses used to describe common grading practices to indicate an impact on students, thus focusing more on students.

The story from quantitative Likert scale survey questions tells a different story than that of the qualitative data collected from open-ended survey questions. The narratives and ability to elaborate on perceptions proved to be the valuable part of this research in regards to grading philosophy, implementation, and overall perceptions on the common grading expectations.

Chapter 5: Conclusions

Determining the impact of change can be difficult. This study aimed to compare teachers' perceptions of common grading expectations among staff from one middle school. While all quantitative results showed no significant difference in medians from a variety of survey questions among subgroups, the qualitative responses provided additional insight into perceptions of certified staff.

Through the use of both qualitative and quantitative results, three main findings emerged from staff responses. First, while there was no statistical significance in differences among subgroups, there were differences found among individuals. Second, regardless of the question asked, the topic of retakes came up multiple times through subgroups. Finally, because of the individual differences, it is important to build trust among staff and administration in order to progress together through change. It is the daily conversations and interactions that will prove to be the most valuable when judging the impact of change.

Differences in individuals, not subgroups

Differing views among staff posed conflicts in responses among different teachers. There was not a clear variable identified to sort the staff into categories. For example, it was not one content area or grade level that felt strongly about fidelity or any of the survey topics. In each subgroup, the median remained constant but the responses to open-ended survey questions varied regarding the common grading expectations. Some staff felt overwhelmed by the expectations while others felt a sense of relief that a consistent message was being given to students and parents regarding grading practices.

No matter the grade level, years of experience, content area, leadership experience, or experience in other districts, staff responses varied on all open-ended responses within the qualitative data.

This conflict amongst individuals provides an opportunity to further discuss the procedures and what works well for some and not others. The staff can capitalize on the fact that no themes were found within subgroups. This means the common grading expectations are working for some and not others within each subgroup. Therefore, teachers can work within their content areas, grade levels, and other subgroups to reflect on what is working for some and why it is hindering others. It is an opportunity to learn strategies from each other. Often times there are groups that believe something does not work for a specific grade level, content area, or classroom. Yet with the variety of responses that flow throughout the subgroups, processes are working for some in each subgroup and that success can be shared with others to offer suggestions of how to overcome barriers that were identified through the open-ended responses.

Retakes

Whether investigating pedagogy, implementation, or individual history of teachers' grading pedagogy, retaking assessments was the one constant piece of common grading expectations that consistently showed in the qualitative data. As mentioned in the literature review, there are two sides to the argument on students retaking assessments. Those in favor reference the benefits to student learning. Those not in favor identify factors, such as time, that impact the ability to do retakes successfully as intended.

In this study, the majority of staff surveyed displayed a sense of frustration towards retake procedures. These frustrations seem to stem from three main points. First, teachers acknowledged the lack of time to meet the requirements set forth for retakes across grade levels. Students were constantly scheduling times before, during, and after school to retake, putting their time into the retake process while falling farther behind on current content. As Kamenetz (2014) learned, this process takes away instructional time from those students who need it most since they are demonstrating gaps in learning. Second, reteaching was seen a hoop to jump through for the sake of retaking an assessment to raise a grade instead of focusing on the learning aspect of reteaching. Lastly, some called retakes a “crutch” that lessened student preparation for the first attempt of the assessment. Often times students were counting on a retake before putting their best effort into the first opportunity. Teachers perceived retakes as an endless cycle of focusing on raising grades that lessened the value of current learning in the classroom.

On the other hand, some teachers expressed support for the processes and a willingness to follow through on the expectations for retakes. Staff in support of retakes commented on how they have adjusted instruction and routines within the classroom to offer retakes that are valuable and provide a clear picture of how students could demonstrate learning. The conversations with students focused less on changing a grade and more on demonstration of knowledge. They found ways to manage the time commitment by altering the length of the retake or having students demonstrate the knowledge through a different mode such as a verbal explanation from student to teacher. Teachers also sought to learn more about different assessment options or ways to grade

the assessments. For example, Marzano (2017) suggests shifting “from an assessment perspective to a measurement perspective” (p. 26). Using this model, teachers use a proficiency scale to measure explicit knowledge instead of combining individual scores in some way to produce an overall score. Whether it is adjusting instructional time, mode of assessment, or learning new ways to assess, all staff will need to have a voice and learn more about implementation before bringing it to the classroom level. Finding ways to make school reflect the real world by providing additional chances will benefit students in the long run (O’Connor, 2002). As new research is published and schools try new ideas, being able and willing to investigate ways to improve processes and allow others to have a voice in those changes will create buy-in and trust to continuously improve grading practices.

Building Trust

Though the quantitative data showed no significant differences among subgroups, this study has demonstrated the importance of building trust with stakeholders and focusing on the conversations that build relationships. If one would only consider the responses on the Likert scale questions and what the statistical analysis revealed, the true learning from the survey would be lost. It was when staff was able to elaborate and express in their own words their thoughts about perceptions of the common grading expectations that differences emerged and unique descriptors were identified.

For example, asking staff to rate on a scale how they feel about the common grading expectations that were created collaboratively produced a numeric value and forced participants into a category of strongly disagree, disagree, agree, or strongly agree. Two participants rated a specific question a 3 but then used their own words as

descriptors. Words such as “frustrating”, “idealistic”, and “reflective” would not have emerged without the ability to elaborate and extend their feedback. Rating how strongly someone agrees or disagrees does not allow for distinguishing remarks to further understand their perceptions. When leaders trust their staff to provide honest and authentic feedback, it is easier to understand the larger impact of change.

Returning to Fullan’s (2004) components of leadership, building relationships is one major component. While he acknowledges the positive power of relationships, there is also the opposite side where relationships can be powerfully wrong. Being able to use powerful words as descriptors such as “enabling” and “compromise” can be signs of trust or they can be signs where relationships stand between leaders and staff. With the qualitative data presented, one could argue relationships are strong and there is a sense of trust to speak up and share the truth behind their perceptions as a means of helping the grading expectations evolve and improve. However, one could also argue that the use of the powerful words is a sign of powerfully wrong relationships where their responses are more a sign of unhappiness and push back. While this study will not make the determination on the current strength of relationships, it still provides an example of how relationships will impact the success of any change initiative.

As the researcher and stakeholder at the research site, this process has shown the importance of seeking more details when soliciting feedback. It is about the conversation and asking questions to understand. While the study started as a means to possibly identify subgroups to target in order to make progress with next steps at the middle school building, this research has shown that support, frustration, knowledge, willingness, or fidelity of implementation (among other descriptors) could not be used to

label a group. Instead, inconsistencies emerged at the individual level. There was no conclusion that more experienced teachers were satisfied or Science teachers were more frustrated. It was individuals across all subgroups that expressed a variety of feelings and responses. The implication for this finding shows the need to meet people where they are. Professional development and support should be individualized for teachers and individual conversations should reflect a willingness to listen and provide support to bring the groups together as a whole.

Recommendations for Further Research

With so many responses from staff revolving around retake procedures and policies, this study could be extended to further investigate that one piece of the common grading expectations. This whole study could be replicated with a focus on retakes. A researcher could ask about pedagogy, practices, implementation, and arrival at retake philosophies. Qualitative responses could be geared at learning more about the pros and cons of retakes at this research site. From there, the findings could assist the administrative team in determining how to frame professional development opportunities to learn more about research surround retakes at the middle school level. Perhaps using opportunities to learn more about the “why” and “how” would provide the staff with a deeper understanding of middle school philosophy with retakes. Using Weber’s (2016) recommendation of giving time to study the why and what, provide opportunities to voice opinions, plan for professional development, and measure success could help to assist individuals understand the purpose of the change in grading expectations, specifically retakes. This time devoted to learning supports the role of the change agent to provide relevancy, meaning, and authenticity (Sheninger, 2016).

This study could also be replicated over time to further investigate the long-term effects of implementing common grading expectations. The current survey was a one shot look at feedback from staff after one year of implementation. The data currently presented could be used as a pre-test and follow it up with a post-test later in time to look at growth patterns and changes (or lack thereof) in practice and pedagogy. As the staff turns over and as common grading expectations becomes the norm and part of the culture of the building, teachers' perceptions may also change over time. A pre-test and post-test may also reveal specific changes teachers are seeking due to roadblocks they have encountered over time.

Finally, this study could be used to further extend research on implementing change at the research site. The administrative team could go back and ask further clarifying questions framed around the open-ended responses already provided. Gaining additional feedback on this change initiative may provide valuable information about how staff feel in regards to change in general, not simply about change surrounding grading procedures. It may also bring to the surface specific feedback about the level of trust and building of relationships among all staff. Schools are constantly changing to meet the needs of students, to meet academic standards, and to stay current with educational research. Understanding the impact of change on the staff and how the outcomes affect the classrooms on a daily basis may provide insight into ways to improve change processes.

Summary

While there were no significant differences in medians across subgroups, qualitative data proved to be the most informational feedback for the administrative team.

If only looking at the Likert scale responses, it would have been deducted that the majority of staff felt at ease with the change in grading expectations. Yet, the open-ended qualitative responses added to the understanding and helped differences in perceptions emerge. While the common grading expectations seem to be acceptable overall, there is still room for improvement.

It will be important for the leaders of the building to consider the feedback as a whole and decide how to further support teachers in this grading transition. In order to sustain change overtime, the leaders will also need to continuously evaluate the processes in place and the level of fidelity of implementation. Remembering Fullan (2006) says change happens most effectively when we speak to people's feelings, it may be beneficial to continue the conversation with staff about where they are and where they hope to be with grading practices. Building a culture with a strong foundation of trust, respect, positivity, and listening to understand rather than reply will support a lasting, successful implementation of common grading expectations (Peeples, 2016).

References

- Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 218-223.
- Alm, F., & Colnerud, G. (2015). Teachers' experiences of unfair grading. *Educational Assessment*, 20(2), 132-150. doi:10.1080/10627197.2015.1028620.
- Association for Middle Level Education. (2012). *This we believe in action*. Westerville, OH: AMLE.
- Bailey, J., & McTighe, J. (1996). Reporting achievement at the secondary level: What and how? *Yearbook (Association for Supervision & Curriculum Development)*, 119-140.
- Boaler, J. (2016). *Mathematical mindsets*. San Francisco, CA: Jossey-Bass.
- Bonesronning, H. (2010). Do the teachers' grading practices affect student achievement? *Education Economics*, 12(2), 151-167. doi:10.1080/0964529042000239168.
- Campbell, E. (2005). Challenges in fostering ethical knowledge as professionals within schools as teaching communities. *Journal of Educational Change*, 6(3), 207-226.
- Carifio, J., & Carey, T. (2013). The arguments and data in favor of minimum grading. *Mid-Western Educational Researcher*, 25(4), 19-30.

- Cherkowski, S., Walker, K. D., & Kutsyuruba, B. (2015). Principals' moral agency and ethical decision-making: Toward a transformational ethics. *International Journal of Education Policy and Leadership*, 10(5), 1-17. Retrieved from <http://journals.sfu.ca/ijepl/index.php/ijepl/article/view/149>.
- Cole, L. B. (2014). The teaching green school building: a framework for linking architecture and environmental education. *Environmental Education Research*, 20(6), 836-857. doi:10.1080/13504622.2013.833586.
- Covey, S. M. (2006). *The Speed of Trust*. New York: Free Press.
- Cross, L. H., & Frary, R. B. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied Measurement in Education*, 12(1), 53-72. doi:10.1207/s15324818ame1201_4.
- Dianis, L. (2003, September). *By the numbers: A data bank on education trends for district leaders*. Retrieved from Apex Learning: www.apexlearning.com.
- Duncan, C., & Noonan, B. (2007). Factors affecting teachers' grading and assessment practices. *The Alberta Journal of Educational Research*, 53(1), 1-21.
- Dweck, C. S. (2006). *Mindset*. New York: The Random House Publishing Group.
- Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin.
- Earthman, G. I. (2011). Decisions and decision-makers in planning educational facilities.

The ACEF Journal, 2(1), 5-15.

Elder, L., & Paul, R. (2011, November 19). *Ethical reasoning essential to education*.

Retrieved from The Critical Thinking Community:

<http://www.criticalthinking.org/pages/ethical-reasoning-essential-to-education/1036>.

Feldman, J. (2014, August 27). Students deserve discussion of teachers' grading practices. *Education Week*, 34(2), 24-24.

Fiarman, S. E. (2016, November). Unconscious bias. *Educational Leadership*, 74(3), 10-15.

Friess, S. (2008, May 19). At some schools, failure goes from zero to 50. *USAToday*.

Fullan, M. (2004). *Leading in a culture of change*. San Francisco, CA: Jossey-Bass.

Fullan, M. (2006). *Turnaround leadership*. San Francisco, CA: Jossey-Bass.

Fullan, M., & Quinn, J. (2016). *Coherence: The right drivers in action for schools, districts, and systems*. Thousand Oaks, CA: Corwin Press.

Gardner, H. (2008). *5 minds for the future*. Boston: Harvard Business Press.

Gordon, M. E., & Fay, C. H. (2010). The effects of grading and teaching practices on students' perceptions of grading fairness. *College Teaching*, 58, 93-98.

doi:10.1080/87567550903418586.

Griffiths, D., & Portelli, J. P. (2015). *Key questions for educational leaders*. Burlington,

Ontario, Canada: Word & Deed Publishing Incorporated & Edphil Books.

Hargreaves, A. (2003). *Teaching and the knowledge society*. New York: College Press.

Hargreaves, A., & Fink, D. (2006). *Sustainable leadership*. San Francisco: Jossey-Bass.

Harland, T., McLean, A., Wass, R., Miller, E., & Nui Sim, K. (2015). An assessment arms race and its fallout: High-stakes grading and the case for slow scholarship. *Assessment & Evaluation in Higher Education*, 40(4), 528-541.
doi:10.1080/02602938.2014.931927.

Hoff, R. (2011, March/April). Schools get smart about energy. *Leadership*, 18-20.

Kagan, S. (1995, May). Group grades miss the mark. *Educational Leadership*, 68-71.

Kallick, B., & Zmuda, A. (2017). *Students at the center*. Alexandria, VA: ASCD.

Kalthoff, H. (2013). Practices of grading: An ethnographic study of educational assessment. *Ethnography and Education*, 8(1), 89-104.
doi:10.1080/17457823.2013.766436.

Kamenetz, A. (2014, November 17). *Testing: How much is too much?* Retrieved from nprEd: <http://www.npr.org/sections/ed/2014/11/17/362339421/testing-how-much-is-too-much>.

Kanter, R. M. (2004). *Confidence: How winning and losing streaks begin and end*. New York: Crown Business.

Kurtzleben, D. (2014, February 27). 1/3 of states will have fewer high school graduates

by 2022. *U.S. News Digital Weekly*, 6(9), 9-9.

Lackney, J. A. (2007). Thirty-three educational design principles for schools and community learning centers.

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

Marzano, R. J. (2006). *Classroom assessment & grading that works*. Alexandria, VA: ASCD.

Marzano, R.J. (2017). *The new art and science of teaching*. Bloomington, IN: Solution Tree Press.

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

Miller, J. V. (2016, September 13). *Want a top-performing team? Make sure speaking up is encouraged*. Retrieved September 17, 2016, from SmartBrief: <http://smartbrief.com/original/2016/09/want-top-performing-team-make-sure-speaking-encouraged>.

Neason, A. (2017, January). Does homework help? *Education Update*, 59(1), 1-4.

O'Connor, K. (2002). *How to grade for learning*. Thousand Oaks, CA: Corwin Press.

O'Connor, K. (2016, July). A repair kit for grading: 15 fixes for broken grades.

Olson, L. (1995, June 14). Cards on the table. *Education Week*, 23-28.

- Peeples, S. (2016). Dare to go first. *Educational Leadership*, 10-14.
- Randall, J., & Engelhard, G. (2008, January/February). Differences between teachers' grading practices in elementary and middle schools. *The Journal of Educational Research*, 102(3), 175-185.
- Richmond, E. (2008, February 4). A floor for failing grades: Parents, educators debate whether kids should get 50 points for doing nothing. *Las Vegas Sun*. Retrieved from www.lasvegassun.com.
- Schneider, J., Feldman, J., & French, D. (2016, November). The best of both worlds. *Phi Delta Kappan*, 98(3), 60-67.
- Sheninger, E. (2016). Leaders can remove barriers for change agents. *ASCD Express*, 11(22).
- Sieh, M. (2016). The top ten! *NCSA Today*, 9-11.
- Smith, D. E., & Fleisher, S. (2011, March). The implications of grade inflation: Faculty integrity versus the pressure to succeed. *Journal of Research in Innovative Teaching*, 4(1), 32-38.
- Sparks, Sarah, D. (2015, December). Teachers feel declining classroom independence, fed data shows. *Education Week*. Retrieved February 11, 2017, from http://blogs.edweek.org/edweek/inside-school-research/2015/12/teachers_feel_declining_classr.html?cmp=enl-enl-eu-news2.

- Stanley, G., & Baines, L. (2010). No more shopping for grades at B-Mart: Re establishing grades as indicators of academic performance. *The Clearing House: A Journal of Educational Strategies, Issues, and Ideas*, 77(3), 101-104. doi:10.1080/00098650409601237.
- Starratt, R. J. (2007). Leading a community of learners: Learning to be moral by engaging the morality of learning. *Educational Management Administration & Leadership*, 35(2), 165-183. doi:10.1177/1741143207075387.
- Starratt, R. J. (Fall 2009). Administering an ethical school system. *Education Canada*, 49(4), 23.
- Surowiecki, J. (2004). *The wisdom of crowds*. New York: Doubleday.
- Tomlinson, C. A. (2005). Grading and differentiation: Paradox or good practice? *Theory into practice*, 44(3), 262-269. doi:10.1207/s15430421tip4403_11.
- Vatterott, C. (2009). *Rethinking homework: Best practices that support diverse needs*. Alexandria, VA: ASCD.
- Weber, C. (2016). Change pain: If it's predictable, it's preventable. *ASCD Express*, 11(22).
- Wiggins, A. (2016, December). At home and away. *Educational Leadership*, 74(4), 76-79.
- William, D. (2016, July). Formative assessment: Minute by minute and day by day.

Papillion, NE.

Wolter, D. (2016, November). The opportunity gap in literacy. *Educational Leadership*,

74(3), 30-33.

Zoeckler, L. G. (2007, Spring). Moral aspects of grading: A study of high school English

teachers' perceptions. *American Secondary Education*, 35(2), 83-102.

Appendix

Each member of the Team will have the same approach/procedure on the 7 following topics. Teams may add to these 7.

***Items in the boxes are examples that could live in each domain; this is not an inclusive list.**

Topic:	Team Procedure-What does it look like and sound like for students and teacher?	How/when will this procedure be communicated to students/guardians?
1. Tardy to Class:	Enter into IC, every class At 3 tardies teacher (TEAM) will contact home to establish a plan.	-Teachers will introduce expectations 1 st week and revisit throughout year. -Info added to syllabus for students and parents to reference
2. Guided Study:	-Planners to be carried with students when not in GSP (i.e. to see another teacher or use the restroom) -Homework will be displayed on the IWB daily	-Teachers will introduce expectations 1 st week of Team Study and revisit periodically,
3. Homework Procedures, including assigning, incomplete, grading of HW	-All teachers will guide students in recording the day's assignment or topic covered. - Homework must be recorded in the planner each day. - If late/missing work becomes a consistent issue the teacher/team will address the issue with the student and guardian to develop a plan.	Teachers will communicate via IC messenger and email and homework page on website.
4. Grading Procedures, for late/missing, re-teaching/re-dos with grades. link-> (See pg 22 in the Handbook)	-Reteaching will occur with students who score less than 77% on formative assessments and/or homework, as needed. -Students' grades are reflections of knowledge for redo work.	-Students be informed on an individual basis as needed.

	-Grade books will be updated on a consistent basis.	
5. Passes, this includes within team area and main halls.	-Planner will be used for all passes outside of the team area	-Teachers will introduce expectations 1 st week of school and revisit as needed.
6. Communication with Parents	-Respond to calls/emails within 24-hour work day. -If an email is sent more than twice regarding the same topic, it warrants a phone call to discuss and clarify the issue. -Document communication on IC	
7. Before School, After School, 9th hour procedures	-Teachers will provide the office passes as needed. -Students need to clear hallways by 3:10 unless supervised by teacher	-Teachers will introduce expectations 1 st week of school and revisit as needed.

Each member of the Team will share individual approach/procedure on the following 11 topics; *these topics are necessary discussion points, not necessarily items to be consistent across the team.*

1. **How are students expected to enter your classroom? How will they know this expectation?**
2. **Entry/beginning of class; during the first 3-5 mins of class, what are students expected to do in your classroom?**
3. **Where will your daily 4-part objective be posted? How will it be incorporated in your instruction?**

4. **What is the teacher's/student's role when a student is absent?**
5. **When are students able to socialize in your classroom?**
6. **When can students be excused to use the restroom; what is the procedure in your class?**
7. **How/when are students dismissed from your class?** *Students should not be allowed to line-up at the door, prior to the bell ringing.
8. **When can students be excused to use the locker; what is the procedure in your class?**
9. **How do you ensure students understand the rules and expectations for your classroom/team?**
10. **How will students know if they are supposed to see you during GSP?**