Illinois State University

ISU ReD: Research and eData

5-13-2020

Utilizing a Learner Centered Design Model to Support a Change in Grading and Assessment Practices

Carmen Margot Bergmann
Illinois State University, jcbrgmnn@icloud.com

UTILIZING A LEARNER CENTERED DESIGN MODEL TO SUPPORT A CHANGE IN GRADING AND ASSESSMENT PRACTICES

CARMEN MARGOT BERGMANN

124 Pages

Practices in grading and assessment have been studied extensively. Guskey and Brookhart (2019) found many studies dating back to the early 1900s, and one as early as 1888, that looked at the reliability and consistency of grades. Despite more than 100 years of research on grading and assessment practices and many efforts to change these practices, most practices today are remarkably similar to those utilized in the early 1900s. The purpose of this qualitative study was to explore the effects of using a Learner Centered Design Model to support a school staff in altering their beliefs and practices regarding grading and assessment. The Learner Centered Design Model, developed by Cale Birk and Charity Allen (2017), is comprised of five phases designed to support a change or initiative by involving members of the organization most impacted by the change. The study was conducted in an elementary school and included teachers who were members of the Learner Centered Design team as well as teachers who were informed about the process by team members during staff and grade level meetings. Information was collected through interviews, observations, and document analysis. All data were analyzed to determine whether there was evidence of change in beliefs and practices related to grading and assessment. Although the time frame for this study was not sufficient to establish lasting change in assessment and grading practices in this school, there was evidence of changes in practice regarding the schoolwide commitment to separate academics and behavior and the identification

of focus standards at each grade level, and there was evidence of changes in beliefs regarding

parent voice and the purpose of grades.

UTILIZING A LEARNER CENTERED DESIGN MODEL TO SUPPORT A CHANGE IN

GRADING AND ASSESSMENT PRACTICES

CARMEN MARGOT BERGMANN

A Dissertation Submitted in Partial
Fulfillment of the Requirements
for the Degree of

DOCTOR OF EDUCATION

Department of Educational Administration and Foundations

ILLINOIS STATE UNIVERSITY

2020

UTILIZING A LEARNER CENTERED DESIGN MODEL TO SUPPORT A CHANGE IN

GRADING AND ASSESSMENT PRACTICES

CARMEN MARGOT BERGMANN

COMMITTEE MEMBERS:

Neil Sappington, Co-chair

Lydia Kyei-Blankson, Co-chair

Guy Banicki

Garth Larson

ACKNOWLEDGMENTS

I have been blessed in my life, and especially in my academic life, by an extraordinary

support system, which includes educators, family, and friends. An attempt to mention all of

them would fail because I would certainly leave some people out, but I will highlight a few of

the most notable, beginning with perhaps an educator who has been a major influence in my life

from elementary school to the present day, John Pye.   John was an assistant principal at my

elementary school many years ago. He was the exemplar of a school leader who was caring and

supportive but who challenged students in his charge to meet his high expectations. He has been

integral to my development from the very beginning. He taught me how to be a leader by his

example, by mentoring me, and by hiring me for my first three principal positions. Many other

friends, colleagues, and leaders in education have nurtured me along the way, and I am thankful

for each and every one of them.

For 14 years, Illinois State University has been at the center of my educational journey,

first as a student-athlete during my undergraduate years, and continuing through two graduate

degrees. My experience as a Redbird has been immensely rewarding, and I have many people to

thank for that. I am grateful first and foremost for Melinda Fischer, the head softball coach who

recruited me to ISU as a student and an athlete. In truth, I had too many wonderful professors

over the years to name them all. I want to thank everyone in the EAF Department at ISU, but

especially Dr. Neil Sappington and Lydia Kyei-Blankson, my co-chairs. I was blessed to be a

student in a number of Dr. Sappington's classes during my doctoral work, and he has been an

amazing mentor along the way. Dr. Kyei-Blankson challenged me to recognize my own

potential and never failed to encourage me along the way. I want to thank Dr. Guy Banicki, who

has been my professor for all classes related to finance and helped me secure a Chief School

Business Officer license. As a valued member of my committee, he provided the perfect balance of encouragement and constructive feedback.

Garth Larson, the fourth member of my committee, was one of the first people to suggest I could write a book someday, which was more encouraging than he will ever know. He introduced me to consultancy, which I hope will prompt me to push for the kind of educational change I envision for all children.

Many friends have encouraged me along the way. I love and appreciate every expression of caring support they have given me. I want to specifically mention Laura O'Donnell, a valued friend, who talked me into pursuing this degree and who has shared the journey with me from day one. She inspires me with her ability to lead with grace and integrity. I have had a whole co-hort of people I consider my friends who have been on the same path for the last three years.

I saved my family for last because they are the solid rock foundation of my support system. When I am unlovable, when I am anxious and stressed, they are my unfailing stalwarts. My parents, Lloyd and Pamela Hulit, have been and continue to be all I could ever ask for. They breathed fire into my spirit while providing me with the structure to channel it in positive ways. They have supported every single endeavor I have embraced, and they have never failed to believe in me. My sister, Yvonne Nyman, and her family, Scot, Christopher, and Benjamin Nyman have endured months of my being a missing person when I was obsessed with my work and with school, but they always welcomed me back to the fold when I was present. My husband, John, and my children, Lance, Peyton, and Brianna deserve medals of honor for all they sacrificed when my academic work and my day and night job dominated my time and energy. They are strong, loving, and fiercely loyal to me and our entire family. I love them more than I can express, and I appreciate their unwavering love and patience.

iii

CONTENTS

TABLES

FIGURES

CHAPTER I: INTRODUCTION

Among the oldest practices in education are grading and assessment, closely related processes that have remained largely unchanged over time. Teachers learn about grading and assessment through their own schooling experience and continue to build that paradigm through their teaching experiences. Their teachers, however, learned about grading and assessing at the start of their schooling experiences as well, and the cycle has continued relatively unabated generation of teachers after generation of teachers.

The persistent status quo of grading and assessment practices is contrary to the fact that access to school and the purpose of public school have changed significantly.   Schools today serve a very different purpose than they did a hundred years ago when they were not available to all students. In the 19th century, most students did not progress past elementary school (Gusky, 2009). Even when more students had educational access, schools operated primarily on an industrial model, preparing students to work in industry except for the select few who received higher levels of education in preparation to lead (Guskey, 2015). Educators did not believe they could possibly teach every student the skills of every vocation, so measures of ability were needed to predict the future roles of students and the appropriate pathways of their learning (Shepard, 2000). In 1892, the Committee on Secondary School Studies, also known as the Committee of Ten, was formed in an effort to create educational standardization across the country. This committee made the decision that all curriculum would be based on a college preparatory format (Armstrong, 2006). Original grading and assessment practices, therefore, were designed to sort and rank in order to determine students who have the ability to complete college work apart from those who might be better served in vocational training or skilled labor.

1

Today, we know more about how students learn in general, and we know that not all students learn in the same way or at the same rate. As Abbott (2015) points out when discussing the need for educational transformation, "Most of the schools that today's children attend were designed when prevailing cultures assumed that children were born to be taught rather than to learn" (p. 95). The most recent federal education act, Every Student Succeeds Act, (ESSA) includes provisions that indicate a shift in focus to the learner and the learning process. In suggesting a movement to a competency-based education framework, Marzano (2017) highlights the following ESSA expectations that support personalized learning:

- Redesigning assessments for student-centered learning
- Piloting new assessment systems that align with competency-based approaches
- Implementing personalized, blended, and online learning approaches

Brooks and Brooks (1999) also emphasize the importance of focusing on student learning. They assert that school reform efforts have missed the mark because they have not reached the nucleus of education, which is how students learn, or construct understanding. A shift in the focus of policies, funding, and reform efforts to the actual learning process would likely include a focus on not only instruction, but also the assessment and grading that are associated with that instruction. In order to provide an education that prepares all students to handle and manipulate information, no matter their future paths, educators must develop grading and assessment practices that honor all students and all learning styles. Unfortunately, this focus on all students is not inherent in many of the practices currently employed in our schools.

The Problem of Practice

With the passage of the compulsory attendance laws in the early 20th century and the reauthorization of the Individuals with Disabilities Act in 1997, schools are more diverse, and

students are staying in school longer (Bradley & Calvin, 1998; Guskey, 2015). We now need all students to learn to the fullest extents of their individual potentials, and we need to develop each student's strengths instead of assuming that all students have all abilities and all interests in common. It is our responsibility to teach all students and to develop systems of instruction, assessment, and grading that address the issues of equity and accuracy.

O'Connor (2007) makes the point that, in education, the term *fairness* has tended to mean *uniformity*. He notes the dissonance produced when educators construe *fair* to mean *same*. That is, all students take the same tests at the same time and grades are calculated the same way, even though educators would never ask a student who wears glasses to take them off for a test so that his/her physical condition is the same as that of the student's non-glasses-wearing peers. If, instead of fairness meaning *uniformity*, we want fairness to mean *equity*, this disparity in thinking needs to be addressed in all areas of instruction and assessment.

We have moved towards equity in instruction as we have developed standards that include content and skills to equip all students with the skills and knowledge that will help them to be successful in a variety of careers or in higher education. All states now have common standards that provide guidance and coherence in terms of what is taught and expected of students, yet the decisions about how and what to grade are left to individual teachers (Guskey, Jung & Swan, 2011; Keni, 2011). This infusion of standards into the curriculum has led to a need for standards-based instruction, which is a new and different approach for many educators. Unfortunately, despite this change in instruction, grading today looks much like it did over a hundred years ago, and there are significant differences in how teachers assign and report grades, with no research to support the practices employed (Guskey, 2011; Guskey, Jung, & Swan, 2011; Peters, Kruse, Buckmiller, & Townsley, 2017).

3

Teachers in the present era are guided by specific standards that explicitly describe what students need to know and be able to do, yet grades continue to represent widely varying factors, depending on what individual teachers value. In addition, there continue to be differences among teachers in terms of which criteria they use to assign grades (Brookhart et al, 2016, Guskey, Swan & Jung, 2011). The academic knowledge and skills described by the standards could lead to grades determined by any combination of factors such as effort, behavior, length of time it takes to learn, or ability to follow directions. In studies that have considered teacher perceptions about grading, teachers consistently indicated that they factored equity, consistency, accuracy, and fairness when making judgments about grading rather than using the algorithm, which creates even more variance within and between classrooms (Brookhart et al, 2016). A system like this does not provide any external measurement accountability (Schneider, Feldman & French, 2016), nor does it provide specific feedback about the learning that has occurred or not occurred.

It is important to emphasize that the traditional grading system has continued, not because it is the best we can create, but because it is all we have ever known (Schneider, Feldman & French, 2016). In order to change something as deeply ingrained in our society as grading, we need to address long-standing teacher beliefs and ideas about grading and how they should be applied to learning, effort, and discipline. Making a fundamental change to this system is made even more challenging by the fact that ideas about grading are based on the personal experiences of all those involved (Brookhart, 2011). Lortie and Clement (1975) identify this phenomenon the "apprenticeship of observation" and makes the point that the teaching profession, unlike any other profession, has been learned by educators beginning in their own childhoods. We observed our teachers and constructed our own ideas about why they

4

do what they do, what effect they have on students, and what makes some teachers better than others. This kind of paradigm makes any change in education extremely difficult because educators' lifelong personal experiences are subsumed in their professional development. Many, perhaps most, teachers had positive experiences with education that began when they were very young, and they repeat what they personally experienced, especially the aspects of their own educations they perceived positively. This creates a never-ending cycle of similar positive experiences for students, who, for a variety of reasons, actually do "fit" the mold of *school*, and specifically its paradigms of grading and assessment. Guskey and Brookhart (2019) point out that this tendency is exacerbated by the fact that teachers have little to no specific or intentional training in grading and are left to make grading decisions using their own professional judgements, shaped by their own personal educational experiences. These authors assert that school leaders are not providing guidance to teachers because they also have received little to no training in effective grading practices, and this is precisely why the cycle of *same old, same old* continues.

In summary, the problem of practice is this: It is the responsibility of educators to produce grades that are accurate and equitable, but traditions and old paradigms have made it difficult to shift grading practices to fit these expectations.

Purpose of the Study

As schools and districts have attempted to make shifts in their grading and assessment practices, as is true with other changes in education, the change process does not have authentic teacher involvement. The process has been imposed on teachers rather than facilitated by teachers with their input throughout the design and implementation of the changes. The general process for any change in education has been to identify a deficit or problem and then provide

5

teachers with professional development and resources to support them in addressing the deficit or problem. This method seems to follow more of what Hargreaves and Fullan (2012) call the business capital view of educators than the professional capital view. In the business capital view, the focus is more on a quick return on the investment and developing a system that will allow for temporary or short-term educators. The professional capital view focuses more on valuing and developing the talents of the individuals, while building capacity for educators over the long term. Therefore, a method or process that effectively supports change and aligns with the professional capital view would include the educators in the process and would place value on their knowledge and abilities. These qualities are embedded in the Learner Center Design Model developed by Cale Birk and Charity Allen (2017).

Cale Birk and Charity Allen (2017) detail the Learner Center Design Model in the book, *Changing Change Using Learner-Centered Design: From Failed Initiatives to a Change Process that Connects, Empowers, and Actually Works*. The model is a 5-phase participatory action research process that involves the learners who are experiencing the change in the change process. A representative team collaboratively and intentionally works through each of the five phases first to co-create ideas, then test and improve the ideas in the process of coming up with measurable solutions. It is still a new intervention in the field, but it has been used in the Salmon Arm School District in the British Columbia and at the King George V private School in Hong Kong, and it is similar to the process suggested by Mintrop (2016) as a method for educators to respond to problems of practice at the local level.

The first phase, *appreciate*, involves taking the time to empathize with, learn about, and truly understand the situation and what it is that is happening. From there, the team moves to the next phase, *illustrate*, where they will set clear goals and targets for a solution. They are not yet

naming a solution, but rather setting parameters for what it is they want the solution to encompass. It is not until the third phase, *ideate*, when the team will envision possible solutions. During this phase, the team essentially brainstorms ideas and then narrows their list based on the goals and targets that were set during the *illustrate* phase. The ideas that remain are tested and tried at the next phase, *iterate*. During this phase, possible solutions are tested and improved until they are ready to be measured and shared in the final phase, *proliferate*.

The purpose of the present study is to understand how engaging teachers in a Learner Center Design Model, which is inherently aligned to a professional capital viewpoint, can lead teachers in developing their thinking and practices related to grading and assessment. I am specifically looking to see whether teachers' involvement in this structured, albeit collaborative model in which they have the space and time to engage in constructivist learning will result in teachers shifting from traditional mindsets around grading and assessment. Furthermore, I will determine if there are observable changes in their practices as a result of their engagement in the Learner Center Design Model.

I will work with an elementary staff that has already begun the process of shifting their grading and assessment practices. They have participated in professional learning experiences related to standards-based grading and have engaged in dialogue about the process for developing a standards-based report card. This population is well suited for this study because they have already entered into the change process but have not discerned shifts in thinking or practices. I will, therefore, be able to see whether or not the Learner Center Design Model, in fact, helps them to experience some of the changes they have sought, but have not realized yet.

The elementary school includes grades kindergarten through third grade. This range is well suited for the study because the teachers at these grade levels teach all subjects and are more

7

focused on skill development than content even though both skill development and content are targeted. It would be more difficult to conduct a study like this on a wider grade band that includes teachers with a varied focus.

Through interviews, focus groups, and document analyses before and after their participation in the Learner Center Design Model, I will be able to capture teacher perceptions about grading and assessment in order to see evidence of changes or development in their practices.

Research Questions

Looking at this problem of practice through this framework and Learner Center Design Model can help answer these questions:

- In what ways does a learner design system approach work to support teachers in changing their perceptions about grading and assessment practices?
- What is the evidence of changes in grading and assessment practices as a result of a teacher's participation in a learner design system?

Organization of the Book

In an effort to address the challenges associated with changing historical and well-established grading and assessment practices, this dissertation will explore the potential effects of engaging educators in a collaborative Learner Center Design process. The first chapter establishes the problem of practice and outlines the scope and purpose of the study. The first chapter also outlines the questions that will be addressed by the study and the basic plan for the study.

The second chapter includes the conceptual framework and a synthesis of the educational theories that will guide the study and the analysis of the data. The conceptual framework is

followed by a review of the literature related to historical grading and assessment practices as well as a review of current practices. The review considers the history of grading and assessment within the context of school reform. The chapter closes with a brief review of the literature related to collective teacher efficacy and collaboration.

The methodology and a description of the study design is presented in chapter three. The chapter includes a description of the context of the study, the district, and the participants. The design of the study is detailed and the ethical considerations are outlined and explained. My positionality and the effects it may have on the study are disclosed along with any potential limitations of the results.

Chapter four will include the results of the study and will outline the data gathered from the case study. The data will be analyzed and interpreted in relation to the research questions. A summary of the results and implications for future work or future studies will be shared in chapter five.

CHAPTER II: CONCEPTUAL FRAMEWORK AND REVIEW OF THE LITERATURE

A key concept to consider in looking for a solution to this grading and assessment problem comes from Heifetz and his work around *technical* versus *adaptive* challenges. A problem that can be solved through the knowledge of an expert is a *technical* challenge, whereas a problem that requires leadership with solutions that reside within the people themselves is an *adaptive* challenge. According to Heifetz and Linsky (2004), technical challenges reside in the head, but adaptive challenges lie in the stomach and heart. They go on to say that solving adaptive challenges requires changing values, beliefs, habits, ways of working, or ways of life, and these changes do not come easily or quickly.

Perhaps we have been looking at this change in grading and assessment as a technical problem in which we just need to set forth new grading protocols based on current knowledge and current research to solve the problem. In reality, a new assessment paradigm is an adaptive problem that requires a much more complex and collaborative solution. We are blessed to have so many experts offering ideas and solutions, but grading and assessment are very personal practices that educators must truly own, and their own grading paradigms have deep roots extending back into their own childhoods. While a technical problem can be solved by applying existing know-how, adaptive problems require individuals within the organization to change their thinking because the solutions lies within them (Heifetz & Linksy, 2004). Heifetz and Laurie (2001) point out the importance of shifting responsibility to the people involved so they can work collectively and collaboratively to find the solution.

Fullan's work with Educational Change and his work with Hargreaves on Professional Capital, combined with the work of Linda Lambert on the Constructivist Leader, synthesize well to build a conceptual framework for how educators might collaborate to find a solution to the

adaptive challenge of changing grading practices. The journey to change begins with the establishment of a driver for the change. Most of the work, however, involves collaboration and the development of meaning that occurs as a result of the collaboration. The role of the leader is more facilitative than directive, or as Heifetz and Linsky (2002) say, sometimes you have to use the skill of "getting off the dance floor and going to the balcony" (p. 66).

The conceptual framework (see figure 1) consists of a driver or goal as the seed at the center of a circle. The driver is surrounded by educators working in a space of social capital constructing new knowledge and meaning about what it means to assess students and assign grades in ways that are equitable and accurate. Through this constructivist learning and collaborative process, decisional capital is produced, which forms the outer rim of the circle. The decisional capital is the expertise of the educators applied to making accurate and equitable decisions about student learning, thereby creating new grading and assessment practices.

Figure 1- Visual representation of conceptual framework

11

Establishing a Driver

According to Lambert (2002), constructivists work well when they have a shared vision and shared values with a common goal, but the objectives and strategies cannot be predetermined or too tightly drawn. Hargreaves and Fullan (2012) list the following as essential to large scale reform: shared experiences, trusting relationships, personal and social responsibility, and transparency. Fullan (2016) talks about two components of effective change: capacity and ownership. He contends that "effective change processes shape and reshape good ideas, as they build capacity and ownership among participants" (p 41). He outlines four criteria he believes can be used to judge an effective driver of change: "(1) fostering intrinsic motivation of teachers and students, (2) supporting continuous improvement of instruction and learning, (3) encouraging collective endeavors or teamwork, (4) affecting all teachers and students- 100%" (p 41-42).

In the present era, when all students, regardless of race, ability, background, or social status, are judged to be entitled to an education based on high academic standards, it is important to have a grading and assessment system that accurately delineates performance while also supporting learning.   A universal system of grading that can be used to effectively measure performance while providing feedback to students about how to continue their learning is long overdue. A driver for this problem of practice should focus on developing grading practices that are equitable and accurate.

Using Fullan's four criteria for an effective driver, the second criterion, supporting continuous improvement of instruction and learning, is addressed by the aforementioned driver that focuses on developing grading practices that are equitable and accurate. When teachers have grading practices that honor all students and produce accurate feedback, there will

12

inherently be improvements in instruction and learning because teachers and students will know exactly what each student knows and where each student needs to go next (Nottingham & Nottingham, 2017). The collaborative process and the fact that the driver addresses grading practices used by teachers and students address the last two criteria – the collective endeavor and affecting all teachers and students. The first criterion, fostering intrinsic motivation of teachers and students, is not as easily connected to the driver. However, if grading practices are more equitable and accurate, it can certainly be argued that learning becomes more intrinsically motivating for students than traditional grades that simply sort and rank students.

Collaboration, Constructivist Learning, and Social Capital

Fullan (2016) asserts that "Significant educational change consists of changes in beliefs, teaching style, and materials, which can come about only through a process of personal development in a social context" (p. 107). When educators are collaborating in this way and producing quality changes, Hargreaves and Fullan (2012) refer to it as *social capital*. In order for educators to truly change beliefs or produce change while engaged in collaboration or as a result of their *social capital*, the constructivist learning theory would say they must also be engaged in shared learning, thinking, or the development of shared meaning. In fact, Lambert (2002) contends that the reciprocal process, or the exchange of ideas and concerns that leads to meaning-making, actually occurs within the context of relationships we have within our community.

When changes in education are initiated and implemented by policymakers or administrators without significant involvement from teachers, the opportunity for collaboration, or the development of shared meaning and *social capital* is diminished. Hargreaves and Fullan (2012), assert that "the movement from power over to power with is still a struggle" (p. 9). There

13

is great power in the collaboration of teachers, especially when they are expected to make changes. In looking at successful initiation and implementation of change, Fullan (2016) found that the level of interaction and collegiality between teachers had a direct correlation to the degree of change and the success of implementation. It makes sense, then, to engage teachers in a collaborative design system when approaching potential changes.

The potential benefits of a fusion between constructivist learning and collaboration of educators are limitless. When educators are engaged in shared meaning-making, they are really listening to one another for the ideas, experiences and insights that are being shared, and as a result, are developing relationships grounded in trust. Once these trusting relationships are established, people are more willing to change their own beliefs about their work and their purpose during their collaborative conversations (Lambert, 2002). The collaboration, or *social capital*, also provides educators with access to more knowledge. As Hargreaves and Fullan (2012) point out, "it gives you access to other people's human capital" (p. 90). *Social capital* can improve each individual's *human capital* simply by virtue of increased access to knowledge and a larger capacity for meaning-making.

Ultimately, what we want from our educators is for them to believe they can make a difference for each and every child they are charged to nurture. They need to know what to do to help each child to succeed. If we truly want this from our educators, we must entrust them with the power to make discretionary judgements in all aspects of their jobs. When it comes to grading and assessing, we want educators to feel empowered to make decisions about how best to provide feedback about what a student knows and can do. Teachers, based on the new knowledge and meaning they construct through collaboration, will use what Hargreaves and Fullan (2012) call *decisional capital* to make these decisions and establish new practices.

14

Hargreaves and Fullan believe that *decisional capital* is enhanced through practice and through reflection and dialogue with colleagues, which can be facilitated by continued collaboration.

A Review of the Literature

From one-room schoolhouses to virtual classrooms, education has seen massive and sometimes rapid changes. In the 19th century, students of all ages were taught by one teacher in one room and most students did not progress past elementary school (Guskey, 2015). With the passage of the compulsory attendance laws in the early 20th century and the reauthorization of the Individuals with Disabilities Act in 1997, schools are more diverse and students are staying in school longer (Guskey, 2015; Bradley & Calvin, 1998). We were once an industrial society, using schools to either prepare students to work in industry or select the few who would receive higher levels of education to prepare them for leadership roles (Guskey, 2015). Schools recognized that they could not possibly teach every student the skills of every vocation, so measures of ability were needed to predict the future roles of students and thus the paths of their learning (Shepard, 2000). We now need all students to learn well and we need to develop all students' strengths. More recently, the influence of technology has created a situation in which educators are preparing students for jobs that do not yet exist and will require knowledge and skills not yet known. All of these changes have brought with them a need for extensive educational reform.

We have developed standards that include the content and skills that will provide students the skills and knowledge they will need to be successful in higher education and in a variety of careers. All states now have common standards that provide guidance and coherence on what is taught and expected of students, yet the decisions about how and what to grade are left to individual teachers (Guskey, Jung & Swan, 2011; Keni, 2011). Unfortunately, grading

15

looks much like it did over a hundred years ago and there are huge differences in how teachers assign and report grades with no research to support the practices used (Guskey, 2011; Guskey, Jung, & Swan, 2011; Peters, Kruse, Buckmiller, & Townsley, 2017). A change in grading practices is the next logical step in the current movement toward standards-based instruction. This need for change is not an indication that what educators have done in the past is wrong, but rather a reflection of the evolution of society and of schooling. Any educator would profess to protecting students from physical harm and fight for their safety, yet, according to Reeves, Jung and O'Connor (2017), "some commonly used grading policies, we believe, rise to this level of urgency in that they threaten the emotional well-being and academic outcomes of children" (p. 42). A shift in grading and assessment practices that aligns with the standards and focuses on children and on their learning is a necessary change, but to date, that shift has been remarkably elusive.

A review of the literature in grading and assessment and student learning will help provide clarity about the problems with traditional grading and assessment practices and the benefits of shifting them, which, in turn, will help to inform the intervention for the problem of practice. I began my search of the literature by using the ERIC (EBSCO) database to search the leaders in the field of grading research including Guskey, Jung, Reeves and Wormeli. I then expanded the search to additional resources cited by these researchers in their writing, especially when the authors or researchers referenced learning or the impact of grading and assessment on learning. The resources related to grading and assessment are generally found within the past decade, although some of the resources on assessment preceded the full standards-based grading movement and are found in the older literature. A review of the literature in teacher efficacy and change will provide insights into the ways in which educators may be supported in sustainable

16

change. For this part of the literature review, I referenced current texts on the topic of teacher efficacy and changes that have been cited in peer-reviewed research journals or articles.

The review of the literature will begin with the alignment of assessment and traditional grading practices, followed by ways in which these traditional practices have become problematic in our current educational setting, establishing a need for reform. This will be followed by the history of grading and a brief overview of past school reform efforts and how they have been aligned with learning in order to see how this problem of practice has evolved. A review of the literature on teacher efficacy and change will then help to connect the problem of practice to the suggested intervention.

<div align="center">**Assessment and Grading Practices**</div>

Assessment and grading are so interconnected that it is difficult to discuss grading without also addressing assessment because assessment often drives grading practices. Assessing student achievement is always complex. Humans think and process information in unique ways. They communicate in different ways, and they learn at varying rates. This inherent complexity creates challenges in the administration of assessments and in the evaluation of the validity of the results. This multifarious concept of assessment is even more complicated when you consider the various types and uses of assessments and how they are associated with grading practices. This section of the review will expand on the purposes and types of assessments and then explore how current assessment and grading practices are related.

<div align="center">**Purpose of Assessment**</div>

Peters, Kruse, Buckmiller & Townsley (2017) pose this question about the purpose of assessment in education: Is assessment used to select or to develop talent? According to Hattie and Timperley (2007), the answer is that it is intended to develop talent. This, of course, does

not align with the many years of academic achievement discourse that included intelligence testing and mass standardized testing (Armstrong, 2006). Hattie and Timperley indicate that assessment should provide information about the discrepancy between where students are in their learning and where they should be. Wiliam and Black (1996) suggest that in order for an assessment to provide feedback, it must generate information about how to close the gap between actual and desired performance. However, what is often found in practice is that assessment simply provides a snapshot of learning rather than providing feedback that can be used by students and teachers to advance learning (Hattie & Timperley, 2007). Assessment should measure learning, but it should also promote learning by identifying a student's level of proficiency and determining what comes next in the learning progression (Erkens, Shimmer & Vagle, 2017, Black & Wiliam, 2003).

Assessment feedback is often a simple reporting of right and wrong with no student understanding of the criteria used to determine correct answers (Shepard, 2000). In fact, with the introduction of computer scanning devices for multiple choice assessments, students sometimes receive, as feedback, a list of bubbled in circles with no access to the questions that were asked. The problem with this approach is that students are left to make guesses about how best to prepare for success on an assessment. If they do not answer correctly, they have no information to help them improve and if they do answer correctly, they may or may not be able to replicate their strong performance on a subsequent assessment. When assessments are used in a way that does not benefit the student, they simply become a tool that rewards or punishes students, often due to factors that are out of their control. Shepard (2000) rejects this approach to assessment in stating that "our aim should be to change our cultural practices so that students

18

and teachers look to assessments as a source of insight and help instead of an occasion for meting out rewards and punishments" (p. 10).

Assessment Types and Uses

Assessment is often categorized into two general types: formative assessment, which is designed to keep students moving toward proficiency and focuses on specific and discrete skills, and summative assessment, which is focused on the verification of a learner's overall achievement level (Erkens, Shimmer & Vagle 2017). Summative assessments, in the form of end-of-learning tests and standardized assessments, have been a part of the education process almost as long as public schools have existed. The earliest existence of summative assessments in the form of standardized testing was in the 1850s in Boston to measure quality of education across the city (Schneider, Feldman & French, 2016). Even with the infusion of formative assessments, summative assessments intended to measure student learning after practice and instruction have occurred will still be a part of classroom assessment plans. Standardized assessments are quite prevalent in the educational culture of the United States and will likely remain for some time, although there are waivers included in ESSA that will allow school districts to use locally created assessments for the purpose of determining student competency and school progress (Schneider, Feldman & French, 2016)

In 1967, Michael Scriven coined the term *formative* in discussing educational evaluation, but Benjamin Bloom is credited with developing the concept of using formative assessment as a way to use feedback from assessment to determine corrective actions to help students close the gap between intended learning and actual learning (Wiliam & Black, 2003; Guskey, 2015, Popham, 2008). While Bloom's work originated in the 1960s, the use of formative assessments

19

became a more prevalent tool used by teachers in the mid 1990s, when learning standards were established in all states.

In 2006, the Council of Chief State School Officers (CCSSO), as part of an initiative to focus on a balanced use of educational assessments, created consortium groups focused on formative assessment. These groups adopted the following definition of formative assessment:

Formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes (Popham, 2011, p. 5).

A key point in this definition is that it is a process rather than a specific test. In order for there to be feedback that informs the next steps in learning for both students and teachers, formative assessment cannot be a one-time event (Nottingham & Nottingham, 2017). Wormeli (2006) suggests that assessment should not be arbitrarily scheduled for a particular day at a particular time. Rather, it should occur when it makes sense in the learning process to check for mastery. When discussing formative assessment, Black and Wiliam (1998) use the term *assessment* to include any activity in which teachers or students are assessing themselves in a way that provides information that can be used as feedback to modify teaching and/or learning. These definitions clearly delineate formative assessments from the traditional view of assessments as "tests" or test-like events.

Black and Wiliam (1998) conducted an extensive review of the literature, 1989-1998, which included 250 articles or book chapters in an effort to determine whether there was evidence of formative assessments raising achievement and where there is room for improvement. They found that the typical effect sizes were larger than most educational interventions, falling between .4 and .7 effect size. They equated the .7 effect size with raising

20

the score of a nation falling in the middle of the pack on an international assessment to the top five. They also found that an improvement in formative assessment had a greater effect on low achievers, effectively closing the achievement gap while increasing overall achievement.

If the purpose of assessment is to empower learners, it is important to provide them with timely and specific feedback during the learning process (Wormeli, 2006; Nottingham & Nottingham, 2017). Traditional summative assessments are not always useful in guiding learning because students do not learn while taking them (Wormeli, 2006; Black & Wiliam, 2003). On the other hand, formative assessment has been shown to be effective in raising student achievement and engaging students in their own learning and improvement. A balanced assessment plan that is intended to bolster learning opportunities will include summative assessments to measure student learning after instruction and practice, but it will also incorporate numerous formative assessment processes to guide and support learning for students and instruction for teachers.

Grading Practices

For the purposes of this research, assessment will be viewed in terms of its ability to impact the learner or the learning process. The type of assessment used and how it is used ultimately impacts grades and also drives grading practices. It is these practices that can have an impact on the learning process, especially when viewing learning through a constructivist lens where learning is seen as student discovery enhanced by constructive feedback. When grading practices produce grades that are inaccurate and/or incomplete, many students see school as being about compliance and points accumulation rather than learning (Reeves, Jung, O'Connor, 2017). Many of the researchers in the area of grading have a common list of grading practices that should be examined when considering alternative grading practices.

21

O'Connor (2007) calls grading practices that do not accurately communicate achievement "broken," and offers "fixes". Guskey (2009) refers to grading policies "that imposed procedural barriers to the implementation of standards-based reforms in grading and reporting" (p. 10). Schimmer (2016) suggests that teachers establish a grading "true north," which is essentially a grading philosophy or standard based on accuracy of reporting and the building of student confidence. Teachers should ensure that their grading practices meet this standard and then check for alignment between their *true north* and their grading practices. There is common agreement that many of the current practices result in grades that do not provide an accurate picture of knowledge or proficiency, and therefore, do not provide information to support learning. Some grading practices are focused on how the grade is determined, such as the averaging of grades, whereas other practices involve what is being graded: homework grades, grades based on first assessments during the learning process, and grades that include multiple sources of evidence. There are some practices that are simply used for punishment like zeroes for incomplete work, or grade reduction for late work.

Averaging Grades

The current system of averaging grades is burdened by two huge problems: equity and accuracy. It favors only students who start out with strong skills and maintain strong skills. It does not value the learning that occurs over time because it gives the same weight to assignments and assessments completed early in the learning process and the assessment given at the end of the learning process (Westerberg, 2016). This system does not allow educators to track student progress and provide specific feedback, and it is possible that two students who reach the exact same level of proficiency will receive different grades because the student whose early results were higher will receive the higher grade (Schimmer, 2016). Not only does this approach

22

produce grades that are inaccurate for some students, it discriminates against students for not having background experiences that some students of privilege may have. Additionally, calculating an average inherently punishes a student for taking risks while engaging in the learning process because a low grade at any point of the process will negatively impact the overall grade. Reeves, Jung, & O'Connor (2017) assert that "a grading system that persistently punishes mistakes instead of rewarding eventual progress and mastery guarantees the stagnation of learning" (p. 43). Student engagement will likely decline if students feel there is personal risk in responding incorrectly or failing to complete a challenge (Hattie & Timperley, 2007). In the real world, people who ultimately achieve in a task or performance are not typically penalized for starting out with a lower level of proficiency.

Grading Homework

The practice of assigning points for homework discriminates against students who do not have the support, time, or safe learning spaces at home to do schoolwork. Homework often serves three purposes: (1) to create more time for the completion of work by extending the work or projects initiated in class into the evenings or weekends; (2) to prepare by introducing knowledge and skills that will be needed for class; and (3) to have students engage in practice opportunities outside of the classroom (O'Connor, 2007). In all three cases, it is likely that students receive very different levels of support in very different home environments, which makes it an invalid source of measurement of what a student knows or can do.

Another concern with grading homework rests on the potentially damaging effects of grading students on their performance on work that is intended to be practice. Schimmer (2016) uses the athletics analogy of practice and games to point out the misalignment of grading homework. A parent who picks a child up from an athletic practice would never ask their child

23

what score was received for practice that day. It is clear that practice is an opportunity to improve established skills and try new skills, while playing the game provides an opportunity to verify that practices have been successful in developing the skills necessary to properly play the game. It is important for students to know that they can use practice opportunities to take risks and that it is acceptable to learn from their own mistakes. Unfortunately, this is not the message sent when homework is graded. Research indicates that motivation is undermined when students are evaluated while they are in the midst of confronting challenging problems subsumed in the process of learning (Shepard, 2000).

No Redos or Retakes

The idea that students only get one chance to show what they know does not honor the student who may not have it yet, but who is determined to keep trying. The goal of educators should be for all students to learn, not just those who can learn on a specific timeline. Wormeli (2011) observes that, "lawyers who finally pass the bar exam on their second or third attempt are not limited to practicing law only on Tuesdays or only under the watchful eye of a seasoned partner for the duration of their careers" (p. 25). When students do not have the opportunity to redo work, we are essentially waiving their responsibility to not learn the content we have deemed important to learn (Wormeli, 2006). If there are skills or content worth teaching and assessing, they are also worth re-teaching until students are successful (Guskey, 2015, Westerberg, 2016). In order to teach students the importance of learning, educators must be persistent in giving them the grace of time they need to learn even if it is not on the expected timeline.

Grades as Punishment

The use of grades as punishment, specifically, the use of zeros, is a practice that can be particularly damaging to children. Grades are not intended to teach students responsibility. Rather, grades should be accurate indicators of student mastery (Wormeli, 2006). In fact, Wormeli (2006) calls "letting the low grade do our teaching an abdication of our responsibilities as educators" (p. 16). Educators will be more effective in meeting their objectives if they create incentives for work students complete on time with accuracy. An appropriate response for a student failing to complete an assignment on time is to hold them accountable for getting it done. In many cases, students are less motivated by a threat of failure than by being given the opportunity to complete their assigned work and move on (Reeves, 2004). There are many fundamental problems associated with issuing grades of zero on a one-hundred-point scale. Besides the fact that a score of a zero does not provide any information about what a student knows or can do, there are mathematical concerns with the assignment of zeros. When looking at a typical one-hundred-point scale, the range of scores for each grade A-D is ten points, whereas the interval from a D to a zero is 60 points. In essence, a zero out of one hundred is the same as a negative six on a four-point scale (Reeves, 2004). On a one-hundred-point scale, it would take two perfect scores in a row just to move from an F to a D and nine perfect scores in a row to move to an A. If a student were to receive a couple of zeros in a class, he/she would be facing a situation in which no amount of hard work would allow him/her to escape the fate of failure, and the only option would be to give up and withdraw (Peters, Kruse, Buckmiller, & Townsley, 2017).

Hodgepodge Grading

Brookhart (1991) refers to the combining of academic and nonacademic factors a "hodgepodge grade of attitude, effort and achievement" (p. 36). A single letter or symbol which includes pieces of achievement, behavior, work habits, responsibility, and attendance is often assigned to a student in a given subject as a grade that is supposed to communicate a student's performance in that subject. Guskey (2015) equates this type of grading with combining measures of height, weight, diet, and exercise into a single number or symbol to represent a physical condition. Despite this incongruity, this is a grading practice with deep roots in the education world, and it continues to be a commonly and widely used application of grades.

When all of these factors are included in a grade, it is possible to have students who are clearly learning, but who are missing some work, earning lower grades than students who have learned to "play school," but have actually learned less than the students who missed some work (Scriffiny, 2008). This makes grades very difficult to interpret and reduces their validity. The basic purpose of grades is to measure performance, yet when grades are not specific, they are interpreted differently by parents, employers, and postsecondary officials who are trying to make decisions based on students' grades (Bradley & Calvin, 1998). In effect, a student who earns an A could be a student who knew the material before the instruction and did the minimum amount of work to maintain knowledge without enhancing it, or it could be a student who never learned the material, but tried hard, completed extra credit, and participated in class (Jung & Guskey, 2009).

A Need for Reform

Many of the grading practices utilized by schools and teachers in the United States were born in an industrialized society, a time during which the purpose of grading was to sort students

26

and determine whether students were prepared for leading or working in industry. Additionally, many grading practices were developed at a time when schooling was not available to all students. Report cards today look very much like they did when the Committee of Ten convened in 1892 to consider high school reform (Cox, 2011). It should be emphasized that these practices are not inherently bad, but they were created for an education system that had a different purpose and structure than the system in which we currently operate. The traditional grading systems have continued, not because it is the best we can do, but because it is all we have known (Schneider, Feldman & French, 2016). Perhaps the standards movement can be the catalyst for change in grading practices throughout the United States and can provide teachers the support they need to construct new meaning around what grading and assessment can, and should, look like.

The standards movement began in the 1990s when states were required to establish curriculum standards and assessments, and it continued into the early 2000s with the 2002 reauthorization of the Elementary and Secondary Education Act, which made the argument that higher standards and measurable goals would increase achievement and reduce the achievement gap. The most recent reauthorization, the Every Child Succeeds Act, has continued the movement with an added provision of more autonomy for states and school districts (Erkens, Schimmer & Vagle, 2017). Teachers now have specific standards explicitly describing what students need to know and be able to do. Unfortunately, grades continue to represent multiple factors that vary depending on what teachers value and there continue to be differences among teachers in terms of which criteria they use to assign grades (Brookhart et al, 2016, Guskey, Swan & Jung, 2011). In addition to the academic knowledge and skills described by the standards, in a case of *deja vu* from an earlier era, a grade could include any combination of

27

factors such as effort, behavior, length of time that it takes to learn, or ability to follow directions. In studies that looked at teacher perceptions about grading, teachers consistently indicated that they considered equity, consistency, accuracy, and fairness when making judgements about grading rather than using the algorithm, which creates even more variance within and between classrooms (Brookhart et al, 2016). A system like this does not provide any external measurement accountability (Schneider, Feldman & French, 2016), nor does it provide specific feedback about the learning that has occurred or not occurred. Students deserve accurate, specific, and timely feedback that can inform their learning (Nottingham & Nottingham, 2017).

Additionally, any system of grading or assessing students that is related to the normal curve is structurally inconsistent with an instructional framework that is intended to support all students in their learning. Using the normal curve in grading or assessment produces information about student performance relative to the student's peers, but it does not provide information about what the student has learned or achieved (Guskey, 2015). Another concern with using the normal curve is that it is designed to be a distribution of randomly occurring events. Assigning grades to students should not be random. Rather, grades should guide teachers to interventions for students who need them in order to learn (Guskey, 2011).

The basic purpose of grades has traditionally been to measure performance, but when grades are not specific and well-defined, they are viewed differently by students, parents, employers and postsecondary school officials based on their own prior experiences or interpretations of grades (Bradley& Calvin 1998). In the present era, when all students, regardless of race, ability, background, or social status, are entitled to an education based on high academic standards, it is important to have a grading and assessment system that accurately

communicates performance while also supporting learning. A universal system of grading that can be used to effectively measure performance while providing feedback to students about how to continue their learning is long overdue.

The use of grades and grade reporting has changed throughout history, but not significantly. The strong desire for standardization and accountability has made change difficult, and over time, practices have simply become institutionalized, or more pointedly, entrenched.

Grading History

In the very early years of formal education in the United States, when students were educated in one-room schoolhouses, progress was determined by which page the students were on in their books or which list of facts they were working to memorize. Grading and reporting were virtually unknown in American schools prior to 1850 (Guskey, 2015). It was not until after the Civil War, when students were organized into graded classes, that teachers began to report progress to parents (Brodinsky, 1972). These early reports of progress were in the form of oral and written reports from teachers to parents, and they varied greatly from teacher to teacher in form and content (Brodinsky, 1972; Brookhart et al., 2016).

As student populations became more diverse and the content being taught became more complex, the need for more specific grades evolved. During the early 20th century, the standard model of A-F appeared, but when concerns were raised about the lack of uniformity in grades, the standardized assessment movement was born. The very first standardized tests were created in Boston in the 1850s, but they became more prevalent across the country in the early 1900s (Schneider, Feldman, French, 2016). In May, 1917, mass intelligence test were given to millions of military recruits serving in World War 1 for the purpose of sorting them into potential leadership roles (Armstrong, 2006). This desire for more standardized grades brought

29

percentages and ratings that were used to report student standing in each subject (Brodinsky, 1972; Brookhart et al., 2016). The 100-point percentage scale that teachers used to create the A-F model was based on an average grade of 50, with grades above 75 and below 25 representing only a small portion of the total grades (Guskey, 2015). The admittedly good intention of using the normal curve was to provide standardization of grades, but problems quickly became apparent.

Around this time, a study completed by Starch and Elliott (1912) debunked the idea that the use of percentages to rate student work really provided standardization. In their study, the authors presented an English paper that had received an 80% to 140 teachers to grade. They discovered that there were different standards for what percent was considered passing, and there was startling variability in the final grades. Of the 140 teachers who graded the paper, fifteen percent of them gave the paper a failing score, and twelve percent of them gave the paper a score over 90%.

The Progressive Education Association joined the grading discussion in the 1920s with demands that grades should consist of more than just a number or a letter. They suggested that grades include information about a child's strengths and weaknesses. Their demands and suggestions came from concerns about how competition and grading might impact students' mental health. The Progressive Education Association also suggested that grading practices should be aligned with the philosophy and objectives of each school. Around this time, child psychologists were also focused on "the whole child" and the importance of considering more than just a summative view of a student's achievement (Brodinsky, 1972). However, by the 1940s, more than 80% of schools in the United States still utilized an A-F grading scale, and as

late as the 1960s, the normal curve was still installed as a "silent partner" making sure there were not too many As or Fs (Brookhart et al., 2016).

The Russian launch of Sputnik in the late 1950s triggered a renewed focus on grading methods. There was an increase in demands to work children harder and to establish the elite learners who would be capable of propelling America to the fore in its quest to be the smartest and most progressive country in the world. Grades were again used as rewards and punishment, building on the assumption that grades served as a motivational tool for students. Perhaps the predominant purpose of grades was to sort students and identify the top tier, which in turn, cast many students as failures (Brodinsky, 1972). In 1971, in an effort to purge the stigma of failure from the experiences of children, the National Council of Teachers of English adopted a grading and reporting policy that emphasized the importance of assessing students in relation to their own work rather than for the purpose of ranking within a group (Brodinsky, 1972).

At the end of the 20th century, there was concern about grading generated by an obvious lack of agreement between teacher practices and measurement theory. There was evidence of great variability among teachers and even within teachers' own practices in terms of what and how they were assessing students and reporting progress (Tierney, Simon, Charland, 2011). In the mid 1990s all states were required to adopt or create academic standards and assessments for English and mathematics. The No Child Left Behind Act of 2002 continued the movement of standards and associated benchmarks (Marzano, 2017). In response to this focus on standards and the assessment of the standards, standards-based grading was introduced in the late 20th century, precipitating the next large-scale shift in thinking about grading (Brookhart et al., 2016).

31

Change in Education

Change in education has historically been difficult and painfully slow. A variety of factors has driven reform efforts over time. During the late 20th century, political leaders and the general public turned a critical eye on the nation's education system, raising questions about teaching methods with particular attention to a study that exposed gaps in student achievement. In 1955, Rudolph Flesch published *Why Johnny Can't Read,* a book in which he questioned the teaching strategies used to teach reading. He pointed out that American children were slower to advance in their learning of reading than European children. He implied that this discrepancy in learning was a threat to American democracy (Armstrong, 2006). Almost 30 years later, in *A Nation At Risk*, a report by The National Commission on Excellence in Education, differences in achievement were affixed to specific groups of students and the gaps that existed in their achievement levels. In that report, an indicator of risk was the functional illiteracy of 17-year-olds. It was noted that 13 percent of American 17-year-olds were functionally illiterate, but for minorities the rate may have been as high as 40 percent (Gardner et al, 1983). Because educators cannot possibly know everything students need to know, they must transition from serving in the role of deliverer of content to the role of facilitator of learning. In discussing the need for educational transformation, Abbott (2015) notes that "Most of the schools that today's children attend were designed when prevailing cultures assumed that children were born to be taught rather than to learn" (p. 95). It is educational shifts like this that have made school reform not only inevitable, but imperative.

It must be acknowledged that educational reform is not a new phenomenon in the United States. Over the years, there have been a number of federal reform efforts as well as countless state reforms. While the desired outcome of education is commonly believed to be student

learning, it appears that much of the reform is directed at factors that are not specific to the learners themselves. In the history of education in the United States, there have been multiple reforms or policy events that have focused on buildings and on the basic infrastructures of education. For example, the Northwest Ordinance required every township to set aside land for education. Money from the Public Works Administration was used to build schools in 1934. Some historical policies and events have stimulated changes in curriculum. For instance, the Smith-Hughes Act of 1917 introduced vocational education to public schools, and the Goals 2000 introduced the idea of creating national education goals and led to a provision in the 1994 reauthorization of the Elementary and Secondary Education Act for states to adopt academic standards. These efforts were focused on content being delivered through instruction in classrooms (Cross, 2015).

In the late 1950s into the 1960s, during a time that Fullan (2016) identified as the "adoption era," the U.S. federal government launched massive curriculum reform. The federal government set aside large sums of money to fund curriculum initiatives in content areas such as science and social studies and in the study of innovations such as open-plan schools, flexible scheduling and team teaching. By the early 1970s, however, it was clear that these initiatives were yielding little to no scalable changes in the classroom.

Fullan (2016) explains that "large-scale reform failed in the 1960s because it focused primarily on the development of innovations and paid scant attention to the culture of the schools and districts in which innovations would reside" (p. 9). Brooks and Brooks (1999) also contend that many of the reforms have not gone deep enough. According to these authors, "Educational reform must start with *how* students learn and *how* teachers teach, not with legislated outcomes" (p. 3). They see successful reform as mirroring the complex construction of knowledge that is

involved in learning. They discuss the discord between Franklin Bobbitt's early 20th century view of education as preparation for adult life and John Dewey's view of education as a process of living and not a preparation for living. When the learning process is honored, education can be a student-centered process that prepares them for their futures.

Many events, policies, and reforms have focused on student achievement and access to learning rather than on the learning process itself. The Elementary and Secondary Education Act of 1965 was aimed at providing poor students access to educational opportunities. The National Assessment of Education Progress (NAEP), in 1969, created "The Nation's Report Card" for the purpose of publishing regional achievement data. The Individuals with Disabilities Education Act (IDEA) granted rights to education for children with disabilities in 1975, and the publication of *A Nation at Risk* in 1983 led to the 1988 reauthorization of the Elementary and Secondary Education, which addressed the lack of improvement in the education of African-American students (Cross, 2015). In 2002, passage of the No Child Left Behind Act (NCLB) marked the largest intervention of the federal government into education in U.S. history.   President Bush promoted NCLB in an attempt to replicate state models that had been used in Texas and New York. NCLB's end goal was that by 2014, 100% of students would meet prescribed standards, regardless of ability, proficiency, or if they were recently immigrated or were English Learners (Hursh, 2007).

While these recent reform efforts were concerned with access to learning and the achievement of learners, they lacked a direct connection to the learning process and failed to have a significant impact on student achievement. These reforms, consistent with most reforms in education, have produced little substantive change to the major institution of education. Academically disadvantaged students still do not have equal access to an equal education, and

34

there continues to be a significant gap between the achievement of African-American students and white students. On the National Assessment of Education Progress (www.nationsreportcard.gov), the gap in the average 4th grade reading score between white students and black students is still 26 points and between students eligible for free or reduced lunch and not eligible is 28 points.

The heavy focus on the delivery of curriculum standards and the assessment of them without a meaningful focus on learning has made victims of low-income students of color. Between 1993 and 2002, at the height of high-stakes testing in the United States, 48% of the nation's largest 100 districts where students of color are concentrated graduated fewer than half of their 9th grade class in 4 years. This was happening in America at the same time that other countries were using their assessments to guide curriculum decisions and inform higher education decisions. As opposed to the too common practice in America of denying high school diplomas, these countries were increasing their graduation rates (Darling-Hammond, 2010).

Changes in American education have been difficult to achieve because historically used educational strategies, including grading and assessment practices, are frustratingly entrenched. Despite local and national efforts from multiple vantage points to create meaningful reform, especially as it relates to student learning, the needle has moved very little. This is important to keep in mind when considering a change in practice, especially a practice that has deep and intractable historical roots across multiple generations.

<p align="center">**Teacher Efficacy and Change**</p>

Michael Fullan (2016) has been writing about the *Meaning of Educational Change* since 1980 when the idea for a book of that title came to him on a trip from Toronto to San Francisco. Since then, he has written five editions of the book in his quest to find the solution to

accomplishing a whole system change in education. In the 5th edition of his book, he focuses on shared meaning and the power of learning in context. When discussing research on professional learning communities and their impact on professional capacity, he states that "…professional learning 'in context' is the only learning that ultimately counts for changing classrooms" (p. 121).

The placement of Collective Teacher Efficacy at the top of John Hattie's Visible Learning ranking of factors related to student achievement with an effect size of 1.39 has resulted in an abundance of research and discussion around collective teacher efficacy and how effective collaboration can impact student achievement and change in education. In his ranking, an effect size of .40 is the hinge point, or the average effect, which makes collective teacher efficacy at 1.39 almost 4 times more effective than the average factor. In recent years, there has been an increased focus on practices such as Professional Learning Communities (PLCs), collaborative inquiry, and the development of shared meaning.

Collective teacher efficacy is essentially a combination of Albert Bandura's 1997 definition of self-efficacy and collective efficacy placed in a school setting (Dewitt, 2016). It is the belief of a group of educators that their efforts will have an impact on student learning. "Some staffs believe that through their collaborative efforts they can help students achieve in measurable ways, while others feel they can do very little to impact student results" (Donohoo, 2016, p. 27).

The potential for the escalation of teacher efficacy when teachers are collaboratively involved in the work is documented in Darling-Hammond's 2010 book, *The flat world and education: How America's commitment to equity will determine our future.* When discussing standards-based reform and state testing, she shares that some states have used performance-

based assessments integrated into the work of the classrooms and that the teachers were trained to score them collaboratively. In these states, the teachers assigned more writing, more complex mathematics problem solving, and student achievement on higher-order skills improved. The "teachers who were involved in scoring these assessments with other colleagues felt they became smarter about student learning and about how to teach effectively as they examined student work and thinking directly and talked with their colleagues about what constitutes good work and how to produce it" (p. 60). Because of their collaboration and involvement in the assessment process, these teachers believed they could influence student learning, and they did. On the other hand, in other states in which teachers were not collaboratively involved in scoring the assessments and did not have as much knowledge about how the results aligned with their classroom work, teachers did not believe they could impact student learning. In fact, 40% of the teachers reported they could raise test scores without improving learning. (Darling-Hammond, 2010).

Hargreaves and Fullan (2012) reference a study conducted by Carrie Leana on the relationship between human and social capital to emphasize the power of the group. Leana found that teachers with higher social capital increased student math scores more than teachers with lower social capital did, and that those teachers with lower human capital but high social capital could perform just as well as teachers with average human capital. The authors noted that social capital actually increases your knowledge because it provides you access to other people's human capital. DeWitt (2016) refers to high-quality professional learning design that takes advantage of the collective teachers' voices and expertise, or human capital, as collaborative inquiry. He sees it as an alternative to top-down, short-term professional learning approaches because it is, instead, a self-sustaining process that puts teachers in a leadership role.

37

Donohoo (2016) provides a four-stage model for collaborative inquiry in which teams identify learning needs, implement changes based on intentional evidence collection and knowledge sharing, collect data on their change in practice, and then reflect on their work. She states that this collaborative inquiry process has been effective in increasing efficacy. She points to factors that contribute to increased efficacy. One is that teachers have more influence and are empowered, and they are able learn about one another's work. In addition, the process creates more cohesion within the staff and results in a more supportive environment for successful collaboration.

Fullan (2016) argues that effective learning organizations are a requirement for educational change because "when enough people start doing the right thing in the settings in which they work, they end up changing their very context" (p. 265). As teachers work collaboratively, they begin to realize their personal and collective efficacy through the change they are able to make. When collective teacher efficacy exists in an educational organization, they are poised to address the four obstacles that, according to DeWitt (2016), must be considered when planning a significant change: time, understanding, resources, and people. An organization that has nurtured a collaborative culture has already addressed the time obstacle and will continue to make time a priority as long as they continue to value the collaboration. The time educators spend in collaboration and in establishing shared meaning ensures that they understand what it is they are doing and why they are doing it. If educators truly have collective efficacy, they maintain a strong belief in their ability to make a difference in student learning and they will be clear about what resources they need. In a collaborative setting with collective teacher efficacy, getting people on board to implement change is not difficult as long as the change aligns with the vision of the organization.

38

Summary and Conclusion

When public schools were created in the United States, almost 200 years ago, they were created for a different purpose, population, and society in comparison to 21st century America. School structures and protocols were developed to serve the purposes and populations of that time period and to serve that society. As changes in the population have led to more diverse schools and changes in technology and the work force have led to a different purpose for schooling, tradition has kept many of the core structures and practices in place. Many of those practices are inequitable and inaccurate and some are actually damaging to student motivation and learning in today's world.

There have been changes in education over time, of course, but those changes have been primarily the result of population growth or technological advances. Actual education reform initiated in an effort to change old paradigms or practices that no longer align to the current population of student or purpose of education have not been well-established. These reforms have not always been focused on actual learning, and they have not effectively engaged the involvement of the people who are closest to the instruction – teachers. Just like other reform efforts in education at large, changes in grading and assessment have been elusive. Although there has been considerable research shared and many efforts made to advance grading and assessment practices consonant with what we know is in the best interests of instruction and learning, traditional core practices remain in most schools today.

With current research indicating the high effect size of teachers believing in their ability to impact student learning, it makes sense that teacher involvement in a change process is crucial. Engaging teachers in a design process in which they are constructing their own meaning about what equitable and accurate grading looks like is essential and the best hope for moving

39

these changes forward. If teachers are involved in this effort, we may, at long last, be able to

effect the changes needed in assessment and grading practices.

CHAPTER III: METHODOLOGY

This study addressed a problem of practice involving the experiences of educators as they encounter and engage in changing grading and assessment practices with the support of a Learner Centered Design model. For this kind of study, an experimental design of any type is not appropriate because an experimental design focuses primarily on causality. In addition, in a scenario in which educators are engaged in change designed to benefit children, the manipulation of an independent variable to create comparisons to a control group is inappropriate and probably not practically feasible.

If this study were simply looking at how teachers respond to a change in grading and assessment, and the intention was to understand why the change continues to be difficult, I would have considered a phenomenological study of the lived experiences of the educators. This type of study would have allowed me to really understand the challenge and how it impacts educators. However, I took the process a step further and applied a change method I believed would work, and then analyzed how well it worked and made inferences regarding ways to improve the process. Yin (2003) contends that a case study allows the researcher to gather evidence from multiple sources to answer a current question or phenomenon. A case study approach is a better fit for this particular study.

Stake (1995) suggests three design options for case study research depending on the function and characteristics of the study: intrinsic, instrumental, and collective. An *intrinsic* case study is used to learn about one particular case and is not used to create general theories or to generalize any findings. An *instrumental* case study can be used to answer a research question or create a general understanding, so the study is directed at gaining a broader insight into a

41

question by studying a particular case. A *collective* case study is the collection of multiple instrumental case studies and is used in order to enhance the ability to generalize the results.

Research Design/Questions

I conducted a qualitative case study in a local elementary school during the 2019-2020 school year while engaging staff members in a learner design system focused on addressing the targeted problem of practice, which was to shift traditional grading practices to make them equitable and accurate. In order to study this problem of practice using a case study design, I engaged in an *instrumental* case study, which is what Stake (1995) suggests works well to gain insight and understanding of a particular situation or phenomenon. The focus of the research was on one school engaged in the change process, but I gained broader insight into the effect of a design system on change. Yin (2003) supports the use of multiple data sources to increase credibility. This study included interviews, observations, and document analysis prior to the intervention, during the intervention, and following the intervention. The information gathered and the conclusions drawn are specific to the context of the study, but the longer-term goal is to share the findings to be used as a resource with other schools that plan to make changes in their grading practices.

I used the intervention and case study to answer the following questions:

● In what ways does a learner design system approach work to support teachers in changing their perceptions about grading and assessment practices?

● What is the evidence of changes in grading and assessment practices as a result of a teacher's participation in a learner design system?

Context of the Study

The study was conducted in the Ashland School district, which is a district that sits just outside of an Illinois city in a small town named Mifflin. While the setting is somewhat rural, and there are a number of families who live on farms and work in agriculture, the area serves as more of a suburban area for many of the employees in the adjacent city. A very large business in the area employs a large number of the Ashland parents.

Ashland was formed in 1962 as the result of a consolidation between two rural school districts. It covers approximately 100 square miles and has a total enrollment of approximately 1050, with a steady growth of 3.5% per year over the last decade. The district includes one elementary building, a middle school building, and a high school building, and employs 85 certified and 55 non-certified staff members across the three buildings.

The participants of the study were selected from the elementary building, which houses early childhood through third grade, although the study will not include early childhood. The elementary school has approximately 397 students with an average class sizes around 21 and each grade level has a team of three teachers. The student population is fairly homogeneous, with a racial breakdown of 92% White, 2% Black, 2% Hispanic, 1% Asian, and 3% two or more races, and only 12% qualifying for free or reduced lunch. The student attendance rate is high at 96%, the chronic absenteeism rate is low at 1%, and the student mobility rate is low at 2%. The school has a rich history of academic success, having earned the Academic Excellence award nine years in a row for having 90% or more of the students meeting or exceeding state standards.

Ashland Elementary School began looking at their grading practices in the summer of 2018. They had a contingency of teachers who attended a three-day training that provided the participants with justification and support to clarify the need for a change in grading and

43

assessment as well as structures and potential next steps for initiating some of the shifts away from traditional practices. Throughout the 2018-2019 school year, there were small groups of teachers throughout the district, including a few elementary teachers, who participated in a book study using Ken O'Conor's book, *A Repair Kit for Grading: 15 Fixes for Broken Grades.* At the onset of this study, however, there had been no school or district level changes to grading and assessment policies or procedures.

Participants

The participants in the study were teachers from Ashland Elementary School. Some of the participants were members of the district school improvement team, which became Learner Centered Design (LCD) team and received support as they progressed through the process of using the LCD model.   They used the model to collaboratively engage in a change process related to the grading and assessment problem of practice. The other participants in the study were teachers who were not on the LCD team, but received information from the LCD team members

No one was coerced to participate in the study. When asking for volunteers, the principal clearly articulated not only the responsibilities of the team members as participants in the LCD model, but also the research expectations for the participants in the case study. Based on a review of qualitative research studies, Creswell (2014) found that a good sample size for a case study is about four to five cases, so I used a sample size of four to five teachers. There was an open invitation for teachers to participate after learning about the study. There were initially six teachers who volunteered to participate in the study, so I reduced the number of volunteers to the targeted sample size after the initial interview by selecting participants from a range of grades with varying perspectives about shifting from traditional grading and assessment practices.

44

There were more than five teachers who volunteered to participate on the LCD team, so I reduced the number of volunteers to the targeted sample size after the initial interview by selecting participants from a range of grades with varying perspectives about shifting from traditional grading and assessment practices. I was able to select not only teachers who were extremely supportive of change, but also teachers who did not support a change. I also included participants with varying years of experience in education.

Description of Participants

After the initial meeting, there were six teachers who volunteered to participate in the study. All six participants were interviewed initially, using the interview protocol. During the interview with one of the teachers, a kindergarten teacher, it was clear that this teacher's instruction and assessment structures were significantly different from the teachers in the other three grade levels because of the introductory nature of the grade level. In order to collect as much data as possible from the teachers who were engaged in this change around grading and assessment, I decided to include the five volunteers who teach in grades first through third.

The final group of participants then was made up of five teachers: Jill, Julie, Lacy, Jenna, and Tim. They represented each grade, first through third, and special education. Their years of experience ranged from three years to thirty years in total teaching experience and from three years to twenty years as teachers in Ashland. At the time this study was conducted, they had taught a range of grades from pre-school to college level students.

Jill was a veteran teacher with 21 years of experience, all but one year in the Ashland school district. She began her career in a private pre-school, then taught one year of 5_{th} grade at the Ashland middle school before moving to 1_{st} grade where she has been for 16 years. She served on the LCD team.

Lacy had a wide range of teaching experiences spanning 30 years. She taught moderate to severe 4th and 5th graders for seven years in a community that is about an hour from Ashland School District. She then spent seven years at a local four-year university supervising student teachers and teaching methods courses. At the time of this study, she had been at Ashland for 16 years, beginning with 11 years in the middle school teaching moderate to severe students before moving to the elementary school working with learning disabled students as a co-teacher. At the time of this study, she was filling that role for the 2nd grade. She served on the LCD team.

Julie is a young teacher, who is new to the profession. She was in her third year of teaching and has spent all of them at Ashland Elementary in the 2nd grade. She student taught in a large district in the city near Ashland School District. She also served on the LCD team.

Tim was a nine-year veteran in education, including seven years in the Ashland School District. He began his career in another rural community near the Ashland School District and taught for one and a half years in the 4th grade. To date, he has been teaching 3rd grade at Ashland Elementary for seven years. Tim recently completed his master's degree in administration and hopes to find a principal position in the near future. He was not on the LCD team, so he was informed about the change work via grade level and faculty meetings.

Jenna had eight years of experience teaching, all in Ashland School District, but spread out over 20 years. She taught middle school for two years beginning in 1999, then she took time off from teaching to raise her children. She returned to teaching in 2014 at the third-grade level at Ashland Elementary. She was not on the LCD team, so she was informed about the change work via grade level and faculty meetings

Learner Center Design Model Implementation

The Learner Center Design Model is a change process model developed and detailed by Cale Birk and Charity Allen (2017) in their book, *Changing Change Using Learner-Centered Design: From Failed Initiatives to a Change Process that Connects, Empowers, and Actually Works.* This model places learners themselves at the center of the change process to create and shape solutions to the problems in schools. It is an extension of participatory action research, which includes co-creation of ideas and is democratic in its design. Team members generate their own ideas about solutions based on their personal experiences and based on input from their peers as they work through the first three phases of the model. The first three phases, *appreciate, illustrate, and ideate*, include reaching a deep understanding of the problem, generating a bank of ideas, and developing goals and a vision for the solution. In addition to working through the initial three phases of the model, team members, as the developers and end users of their recommended solutions, are responsible for testing and making suggestions for improving their delineated ideas as they navigate the last two phases of the model. The final two phases, *iterate and proliferate*, involve the actual design, development and testing of solutions and then the measurement of success and sharing of the solutions. Throughout the entire process, the teachers met with grade level teams between meetings in order to engage in two-way communication; sharing what was being discussed at the LCD meetings and then collecting feedback to take back to the LCD meeting.

While the teachers who participated in this study had had experience with collaborative teams or action research, this specific model was new to them, and the overall structure was a shift from typical change processes. The team received support from Cale Birk through virtual training sessions and discussions. I scheduled five sessions in November, December, January,

47

February, and April for the team, one session dedicated to each phase of the model. During two of the training sessions, Cale provided support by asking exploratory questions about the team's thinking and reasoning for the decisions they were making. He also explained the purpose of each of the phases, along with what their work should include during that phase and how they would know they were ready to move on. He was also available to answer questions the team had about their work or their next steps. During the sessions when Cale was available, I was a participant as I supported the team, but I also observed and collected data.

The team continued to meet regularly, as an LCD team and also with their individual grade level teams as they engaged in each phase of the LCD model. There was a longer period of time between the fourth and fifth meetings, during the iterate phase, when they were actually testing their solutions in their classrooms and making adaptations. During this phase, I asked teachers to write reflections to capture their thinking as they experienced the LCD model.

Data Collection Process/Tools

The purpose of my study was to determine whether the use of an LCD model to support a change in grading and assessment practices would produce changes in perceptions and practice. I designed the study to address two specific questions:

(1) Does an approach utilizing a learner design system support teachers in changing their perceptions about grading and assessment practices?

(2) Is there evidence of changes in grading and assessment practices as a result of a teacher's participation in a learner design system?

I needed to ensure, therefore, that I collected data that would provide answers to these questions.

Interviews. Interviews are a popular data collection method in qualitative research and especially in case studies because they provide insights into participants' thinking, beliefs, and

48

perceptions, insights that might not be gained by simply asking participants direct questions about these cognitive/emotional components (Glesne, 2016, Hancock & Algozzine, 2016). A strength of interviews is that the researcher can utilize probes, or prompts to clarify responses or to ask for more information, in order to obtain the most useful in-depth data (Johnson & Christensen, 2017). Interviews allow the researcher to learn how participants view the world, what participants are thinking or feeling, or obtain information about past events that cannot be replicated (Merriam, 2015).

Researchers must consider a number of factors when planning interviews for data collection. The researcher must decide whether the interview will be face-to-face or on the telephone and whether it will be conducted one person at a time or in a focus group format. The researcher must also take the structure of the interview into account. There is a range of interview structures. An interview can be highly structured, in that the wording and the order of the questions are predetermined, or it can be essentially an oral form of a written survey. At the other end of the continuum, an interview might be unstructured. The questions are open-ended, flexible and exploratory, and the interview resembles a give-and-take conversation (Merriam, 2015). The researcher also needs to consider the content of the questions, the type of questions, and how they are worded. The content of the questions is guided by what the researcher wants to understand. Types of questions include questions about experiences or behaviors, opinions or values, feelings, knowledge, sensory, or background or demographics. Typically, a researcher will begin with questions about experiences or behaviors because they are usually the easiest for participants to answer. On the other hand, knowledge questions, which feel more like a test, should be reserved for later in the interview when the participant is more comfortable. When

49

considering the wording of the questions, it is important to avoid *yes* or *no* questions and multi-query questions (Glesne, 2014).

In order to understand and detect changes in perceptions, I wanted to hear from the teachers themselves. I did this through face-to-face interviews, observations and written reflections. The interviews were semi-structured, which means I had predetermined questions, but they were flexibly worded, and I included follow-up questions to probe more deeply into the thoughts and ideas expressed by each interviewee. Hancock and Algozzine (2016) suggest that semi-structured interviews are well suited for case study research because interviewees can express themselves openly and can describe the targeted phenomenon from their own perspectives without intentional or unintentional insertion of the researcher's perspective. I designed the question protocol so that I began with questions about experiences and behaviors or feelings before asking questions about knowledge.

I interviewed each teacher once at the beginning of the study, before the teachers engaged in the LCD process, and once after the teachers completed the LCD cycle. The questions were focused on the practices teachers used for grading and assessment, their thinking about how their practices addressed equity and accuracy, and their thoughts about the possibility of making changes to their practices. The pre- and post-interviews contained similar questions, although there may have been questions added to the post-interviews based on themes that emerged from the data collected during the study. I interviewed each teacher individually, face-to-face, at a location of their choice, which was their classroom in each case. The interviews took 45 minutes to an hour, and I audio-recorded the interviews with the voice memo app on my phone in addition to taking notes, then I transcribed the interviews into a Word document. I saved the audio recordings and the transcripts on my password protected computer until I completed the

50

data collection and analysis of the data. Once I have defended my dissertation, I will erase all files from my computer.

Observations

Observations are a primary source of data in qualitative research, much like interviews, but they differ in that they take place in the setting of the phenomenon, and they are a firsthand account rather than a secondhand account of what happened (Merriam, 2015). Observations can be beneficial in qualitative research about people because people do not always do what they say they do (Johnson & Christensen, 2017).   In order to consider observation a research tool, it must answer a specific research question and be subject to the checks and balances necessary to produce trustworthy results (Merriam, 2015). Skilled observers are attentive, able to take descriptive notes, and they can separate relevant details from trivia (Merriam, 2015).

In order to be effective as an observer, social skills are important because trust and rapport with the participants are essential for obtaining accurate data. Regardless of the role the observer takes, the rapport established and the level of trust that the participants have with the observer may impact the accuracy of the data collected (Johnson & Christensen, 2017, Merriam, 2015). There is a continuum interaction between the observer and participants depending on the role the observer takes on. Other roles, which include complete participant, observer-as-participant, and complete observer, have varying levels of interaction between the researcher and the participant. The highest level of observer participation is *complete participant*, which means that the observer is a part of the group, and the participants are not aware that the observer is conducting research. In the next level of observer participation, *participant-as-observer*, the observer spends a significant amount of time with the group and is focused on being an active participant, and the participants are aware that they are being studied. The role of *observer-as-*

participant is similar to *participant-as-observer*. The difference is that the participation role is secondary to the observation role. The least participatory role is *complete observer*. The researcher is an outside observer and observes the group without participating (Creswell, 2014, Johnson & Christensen, 2017).

I used the *observer-as-participant* type of observation, which means I participated in the trainings and meetings, but my observation role was primary to the participation role. This allowed me to be seen as a part of the group while also focusing on collecting data. An advantage to using the observer-as-participant method is that the observer can receive timely feedback about the observations and tentative conclusions from the participants as the observation is occuring, which is also helpful in building trust (Johnson & Christensen, 2017).

I observed LCD training sessions and grade level team meetings as opportunities to collect data related to the group's discussions and questions. My goal was to look for evidence of changes in the participants' thinking about grading and assessment and for evidence of intended changes in actual practice related to grading and assessment. I used an open-ended approach, which means I was a participant and asked questions, taking on the role of observer-as-participant, one of the most common roles for researchers (Creswell, 2014, Johnson & Christensen, 2017). I observed and participated for the duration of each meeting, which lasted two to three hours each, and I took descriptive notes of all comments and behaviors that provided insight toward answering my research questions.

Document Analysis

According to Merriam (2015), documents include anything in existence prior to the research, with the exception of researcher-generated documents which are prepared by the researcher or by participants after the study has begun. Johnson & Christensen (2017)

differentiate between secondary or existing data, which are data that already existed prior to the study, and constructed data, which are data provided by the participants during the study. Documents can be classified into two major categories: public or official records, which are records that document any information about an organization, and personal documents, which include diaries, letters, scrapbooks or any other narratives of human experiences (Merriam, 2015). Some of the advantages to using documents include the convenience of accessing them, the opportunity to easily capture the language and words of participants, and saving time because there is no need to transcribe (Creswell, 2014).

When using documents, it is important to determine the authenticity of each document as well as how and why each document was produced as part of the research process. This will help to determine how relevant the documents will be to the research as well as how valuable they will be in providing accurate information. A system of coding or cataloging data is another part of the process and that system may be influenced by the nature of the documents. The aim is to use coding and analysis to record what is contained in each of the documents and to understand the communication of meaning (Merriam, 2015).

For this study, I utilized a combination of public records and personal documents to capture the thinking and beliefs of the participants. I accessed previously constructed public documents such as handbooks, grading policies, sample report cards, and teacher gradebooks to understand the historic beliefs and the current beliefs and ideas held by personnel in the targeted school district regarding grading and assessment practices. I then analyzed similar documents at the end of the study, which may have been re-crafted as a result of the LCD process. I also utilized personal documents generated by the researcher or constructed documents in the form of reflective journals written by participants throughout the research process.

The participants were asked to write reflections, which were open-ended, in which teachers, after LCD meetings or grade level team meetings, recorded their thoughts about any changes they made, or were considering making. I asked the participants to respond to specific questions after I had observed them if I believed something in the observation might have needed elaboration, clarification, or more detailed responses. While these journals may not always have been reliable accounts of exactly what happened, they had the potential to provide considerable value to the researcher's ability to understand participants' potentially evolving attitudes and beliefs and to gauge their perspectives about what they were experiencing (Merriam, 2015).

Some of the data collected through the interviews, observations, and reflections provided evidence of a change in practice, but I also utilized document analysis to collect evidence. The interviews helped guide some of this data collection because I collected and reviewed all documents referenced by the teachers. I examined documents such as teacher gradebooks, classroom handbooks, assessments, and report cards. I looked for statements or policies about grading and assessment and evidence of grading and assessment procedures followed by the teachers. Creswell (2014) notes that documents represent data to which participants have given attention. Therefore, written policies and the procedures teachers follow when assigning grades help me understand the priorities of the educators and the priorities reflected in the documents themselves. The policies and procedures were given attention and were intentional, so they provided a different perspective than what I heard from the teachers during interviews and observations, during which their responses and statements may have been impacted by their knowledge that the study was focused on change in grading and assessment practices.

Interview Protocol/Reflection Questions

The interviews were conducted in person, were recorded, and took place in a location and at a time frame selected by the participant, which was their classroom in each case. Each interview lasted approximately 45-60 minutes and was guided by the following questions, with some variance in follow-up questions depending on the participants' answers.

- What do you remember most about grades and assessments from when you were in school?

- What do you see as the purpose for assessing students?

- What do you see as the purpose for assigning grades to student work?

- How do you think grades are impacted by a student's background experiences or home life?

  - Do you think that students whose parents are more involved in their schooling are more motivated by grades?

- What is it that contributes to a student's grade in your class? Tell me about the typical entries you would record in the gradebook.

  - If there is homework- ask about support at home, readiness,

- In thinking about a grade that a student receives on a report card, tell me what the grade communicates about that student?

  - Can a student who has a lot of content knowledge receive a low grade?

  - Can a student who struggles in a content area receive a high grade?

- How is the grading in your classroom different from the grading when you were in elementary school? How is it similar?

  - How have your grading practices changed as your career has progressed?

- What are your thoughts about students having the opportunity to re-assess?

  o Do you think students should be accountable to do something to improve their learning before re-assessing?

  o How often should they be able to re-assess?

  o What kind of "credit" should they get for subsequent attempts?

- How do you think you should assign grades to those students who struggle with the skills or content at first but are able to demonstrate mastery by the end of the grading period?

  o Should their grades be different than the students who show mastery earlier?

- How do your assessments help students learn?

  o How do you know if the grades students earn accurately communicates what they know and can do?

The reflections were completed after some of the LCD meetings or grade level meetings and were completed if the teacher was considering making an adjustment or trying a new grading or assessment strategy. The reflection document included the following questions:

- Based on the discussions you have had with your colleagues, are you considering an adjustments or new strategies?

- Describe the change you plan to make or the strategy you would like to try.

- What caused you to consider the change? What do you hope to accomplish?

A final reflection was offered to all participants after the final LCD meeting and grade level meeting with the following open-ended question:

- Describe any shifts you have made to your grading and assessment practices and what the impact has been for you and your students.

56

Timeline

I conducted the study during the 2019-2020 school year with the intention of engaging the staff in a study and intervention throughout the school year when they were working and interacting with students. This timeline allowed for results and analysis to be shared with staff prior to the end of the school year so they could utilize the information as they plan for the next year.

November 2019- I conducted interviews with participants and analyze documents such as parent letters, handbooks, gradebooks, and report cards.

November 2019 – April 2020- The LCD team received training on the components of the Learner Centered Design Model with author Cale Birk via two virtual collaboration sessions. I participated and observed in these sessions and led and observed in the other sessions.

November 2019 – April 2020 -The LCD team met regularly as they worked through the phases of the LCD model and addressed the problem of practice. I observeed the LCD sessions and asked them to write reflections regarding any changes they were implementing throughout the process.

April 2020- I conducted interviews with participants and analyzed documents such as parent letters, gradebooks, graded work, report cards, and reflections.

Data Analysis

When discussing qualitative data analysis, Merriam (2015) points out that the collection and analysis of data is simultaneous in qualitative research. According to Saldaña (2016), "The majority of qualitative researchers code their data both during and after collection as an analytic tactic, for coding is analysis" (p. 9). The ongoing data analysis guides subsequent data collection.

57

However, once the data are all in, the analysis does become more intensive and comprehensive. I utilized several methods for organizing and analyzing the data as I collected them and prepared for a summative analysis.

During the study, I utilized pattern coding to look for phrases and words that appeared often in pre- and post-interviews, in the observations, in the documents, and in the reflections. Pattern coding allows for the organization of data into sets or themes that can be searched for rules, causes or explanations (Saldaña, 2016). At the end of the study, I entered the data into a matrix with the early data on one side and the later data on the other and utilized longitudinal coding. This type of coding involves looking for similarities and differences in data from one time period to another in order to generate inferences of change (Saldaña, 2016).

I generated themes from the pattern coding and the longitudinal coding and continued to analyze them and categorize them into themes and sub-themes in an effort to develop a storyline or explanation for the change process. As I generated themes throughout the study, any contradictions or convergences I observed helped guide my questioning for the post-interviews. The theory and explanation generated from the themes helped shape my recommendations for next steps and alternative applications of the intervention.

Trustworthiness/Credibility/Authenticity

According to Yin (2003), case study research is viewed by many research investigators as a less desirable form of inquiry than experiments or surveys because it lacks rigor, is more biased and is not generalizable. Therefore, research validity, or trustworthiness, is a factor to be considered in case study research. Attention to trustworthiness while planning a case study can help to allay these concerns. Trustworthiness refers to the quality and rigor of the study, and the degree of confidence others can have in the interpretation of the data (Glesne, 2016).

58

Glesne (2016) and Johnson and Christensen (2017) provide tables which outline a number of strategies that can promote trustworthiness in a qualitative study. Both texts include triangulation, member checking, and peer review as strategies that can be used to promote qualitative research validity. Given that I am vulnerable to a degree of researcher bias due to my extensive involvement with this topic and my desire to see change, it was important to engage a range of strategies to ensure the highest possible level of validity. The strategies selected also enlisted additional people to ensure multiple perspectives and viewpoints.

Triangulation

Triangulation involves the cross-checking of multiple sources of information and the use of multiple procedures. When the information or data agree, or when they triangulate, they are considered *convergent*. However, when the information or data do not triangulate, they are *divergent* (Johnson & Christensen, 2017). In order to triangulate what the teachers were saying in their interviews, I included the reflections they wrote throughout their change experience. In addition, I had the data from the observations which will were used to document what they were saying and doing as they engaged in the change process with their peers.

Member Checking. One way to ensure that I did not interject my own bias and personal interpretation to the information generated via the interviews and observations was to utilize *member checking*, the process of verifying with the participating community that what was written and summarized was an accurate representation of their thoughts and ideas (Glesne 2016, Johnson & Christensen, 2017). Glesne (2016) points out the challenge of finding the right balance when it comes to member checking. Asking participants to review all transcripts and all observation notes would be overwhelming and would place unreasonable demand on the participants' time. Instead, I shared with participants the content of the summary and the

59

analysis and asked them to judge if I was consistent with their thinking and what they intended to share.

Peer Review

To ensure that my conclusions and explanations were credible, I used *peer review*, which involves a discussion and review of the actions, coding, interpretations and conclusions of the researcher with other knowledgeable professionals (Johnson & Christensen, 2017). I used the feedback from one peer reviewer during the process of coding interview and observational data throughout the study and when I was coding, organizing data into themes, and analyzing data at the end of the study. I then asked a panel of peer reviewers to question my thinking, push back on my ideas and to make good faith efforts to deconstruct my conclusions.

Positionality

I was only eight years old when I decided I wanted to be a teacher. From that moment forward, my desire and commitment to be a teacher remained firm despite the fact that my understanding of education, over time, changed drastically. My views of teaching and learning have been shaped by my experiences as a student, as a parent, as a teacher, and as an administrator. Each experience has contributed substantially to my interpretation of effective teaching practices and to my understanding of how children learn. In full disclosure, my experiences in education have been inconsistent and, at times, have provided conflicting feedback to my ever-developing understanding of teaching and learning. My personal educational philosophy has been shaped by my experiences – good, bad, and indifferent – in all of my educational encounters, as a student, parent, teacher, and administrator.

One aspect of education that has been of particular interest to me and has contributed to my educational philosophy has been grading and assessment. The integral connections between

60

instruction, assessment, and learning have always been very clear for me. I realized early in my career that assessment is like a map that creates a route for learning, and without it we would never know where we are going. Even when I was young, I sought feedback from the adults in my life during the learning process to ensure that I was going in the right direction. My views and opinions have changed and evolved over time as a result of experiences, discussions, and research. My motivation to seek change in grading and assessment practices has, however, been most influenced by my experiences.

As a student, my experience with grading and assessment was predominately positive, but definitely changed over time. I came from a well-educated family with parents who valued education and sent me to school well prepared and continued to support my education in many ways. I had help when I needed it, access to resources that supported my learning, and I always had a quiet place to study and work on my academic assignments. Learning the content was not difficult for me and I rarely struggled to learn new concepts or skills. When I was younger, however, my grades were sometimes negatively impacted by my propensity to socialize, my lack of focus, and my desire to be the boss of others. By the time I reached high school, I continued to find learning to be fairly easy, but I also developed significantly improved functional skills. While still social, I learned when it was and was not appropriate to socialize. I became much more focused, and I learned the importance of following directions and adjusting my work based on what teachers expected. Through high school and college, I discovered that different classes required different levels of effort and that grades were based on a variety of factors. Given that I was willing to ask for help when needed and that I learned from my mistakes, I rarely struggled to get the highest grade. I became very proficient at the "game of grading". Grading and

assessment, in my mind, were fair and sensible, and it never occurred to me to question the grading practices in place.

My experiences with grading and assessment as a parent varied greatly due to the extreme differences in my children. My oldest child, a recent college graduate who began a teaching career in the year I conducted this study, is very intelligent and was identified as "gifted" at a young age. Learning new concepts and skills was never hard for him. His challenges in school were always the result of a strong will, a lack of attention to detail, and a lack of motivation to put time and effort into work that did not interest him. If he did not see value in the work he was asked to do, or he was not interested in the topic or assignment, he would sometimes choose to not do it. This resulted in his receiving zeros on assignments, which then resulted in low grades despite his high levels of mastery of the content. His strong will and lack of attention to detail often resulted in work products that failed to conform to the instructions provided by his teacher. This also had a deleterious impact on grades in subjects in which his mastery of the content was high. He was not interested in playing the "game of grading" and, therefore, received grades that did not accurately reflect his high ability.

My second child, a college student, has had a different experience in school. He is also very intelligent, but he has a different way of thinking about things. He is self-motivated, thrives on challenges, and wants to please. When he was in grades K-12, these characteristics worked very well in the traditional format of grading and assessment most of the time. His challenges were related to his unique perspectives and thinking. At times he was penalized for approaching assignments in ways that differed from his teachers' expectations. There were other times when, due to his unique processing of ideas, he needed more time to learn a concept. If he performed poorly the first time, the skill was assessed accordingly, but when he demonstrated complete

mastery at a later time, his summative grade suffered as a consequence of his earlier struggles. He was, however, very organized and was a pleaser, so he could usually figure out how to bring his grade up, even when he faced these challenges, although he was not always able to achieve the highest grades.

The experiences of my third child, now in high school, provide yet another perspective on grading and assessment. She has a high level of intelligence, is a pleaser, and a perfectionist. She has never struggled to learn new concepts, and she has easily demonstrated mastery on assessments. Her challenges have been grounded in her perfectionism and an unhealthy fear of failure. When taking tests, she sometimes runs out of time because of her compulsive need to check and recheck her answers. When this happens, she sometimes loses points for the questions she does not have time to address, or she is forced to rush through parts of the test in order to finish on time. When she knows that a homework assignment is going to be graded, she spends an inordinate amount of time trying to make her work perfect even after she has clearly reached proficiency on the intended target of the assignment.

I have developed many insights, opinions and ideas about grading and assessment as I have navigated all levels of schooling with my children. Over the eighteen years I have been a parent to a student in school, I have questioned grading practices such as giving students zeroes, averaging grades, and grading homework. Ken O'Conor calls these practices "broken" and offers fixes for them (2007). I have realized the importance of student choice in grading and assessment, and I have experienced frustration with the lack of flexibility in many of the current assessment models. I have seen, through the prism of my own experiences as a parent to three uniquely different children, how grading and assessment practices can impact student motivation

and self-concept and can contribute to levels of anxiety that are detrimental, not just to academic progress, but to emotional and social health.

My experiences as a teacher have also shaped my insights and opinions about grading and assessment. In my eleven years of teaching, I had students with special needs in my class every year, and I had gifted students in my class ten of the eleven years. I often struggled with the challenge of using a vaguely described grading scale to produce a single grade in each subject which adequately described the performances of my students whose readiness skills, strengths, and needs varied greatly. I questioned the fairness of two students receiving the same grade on a test that was heavily modified for one of the students. I wondered if it was acceptable for a student to receive an A in math despite the fact that she had not yet mastered one of the concepts covered in the quarter. I noticed that I gave partial credit on math problems when students demonstrated proficiency in solving a problem but had the wrong answer due to a small addition error, even though other teachers gave no credit in similar situations on the same test. I eventually came to question the integrity of an entire grading system that subsumed these inconsistencies.

As an administrator, I have been able to view grading and assessment through a much wider lens. I have had the opportunity to observe the practices of many different teachers in widely varying school environments. As a principal, I sat in on special education meetings and problem-solving team meetings. On many occasions, I witnessed situations in which grading practices had become barriers to student learning because students either did not know how, or who were unwilling, to play the "game of grading." In these cases, students who scored poorly on an assessment, or who received a low grade on homework because of a lack of support at home, allowed those grades to define them and accepted their assigned grades as factual

descriptions of their abilities. Many of these same students did not take advantage of extra credit or opportunities for "easy points." Conversely, I encountered situations in which students were so good at playing the "game of grading" that they appeared to be doing well in their classes even when there were large gaps in their knowledge and understanding. Patricia Scriffiny (2008) calls this scenario "playing school" and rightly points out that some students who score poorly on assessments, but who receive high grades on homework and take advantage of extra credit, may well struggle in future classes.

As a researcher in the area of grading and assessment, I recognized that it was important for me to accept and account for my personal biases. There is no doubt that my experiences as a parent, teacher, and administrator have created personal biases against traditional grading practices, although I would like to think of these biases as reasonable conclusions based on countless observations over the course of my life and my professional career. There are two biases of which I am acutely aware: (1) I believe that traditional grading practices produce inequity, and (2) I believe that traditional grades do not engage students in their own learning. In the words of Alfie Kohn (2011), "the more students are led to focus on *how well* they're doing, the less engaged they tend to be with *what* they are doing" (p. 29). While many experts in the field hold similar beliefs and make statements that align with my thinking, it was important to search out and cite research that supported the point of view with which I am comfortable before making specific claims in my own research. As a qualitative researcher, I needed to be cognizant of how my biases might evidence themselves in my questions or in the manner in which I interacted with research participants. Furthermore, I needed to be committed to the proposition that facts and truth are more important than advancing any personal agenda, and I remained committed to holding tight to that proposition.

Another consideration for me, as a qualitative researcher, was my conceptualization, as well as those of the educators in the school I studied, of my place within the institution. Banks (1998) contextualizes the idea that individuals who are socialized within a community have ways of knowing that differ from those who are socialized in other communities with what he calls a *typology of crosscultural researchers.* According to his typology, I fell into two categories, depending on the scope of community. In one category, I was an educator in the same state and region as the school in which I worked, and therefore, had a good understanding of the state expectations, and of the historical context of school challenges and progress in the area. In this sense I was an *indigenous-insider*, which according to Banks, is someone who endorses the unique perspectives, beliefs and knowledge of the community and is perceived to be able to speak with authority about it. On the other hand, I have never been a parent or an employee in this very unique, tight-knit school district. This fact made me more of an *external-outsider* according to Banks. An *external-outsider* is someone who is socialized outside of the community and has little to no appreciation for the values and perspectives of the community, a point of view that might lead to misunderstandings.

My personal experiences with, and my interest in, grading and assessment as well as what I trust will be the justified biases I have developed have driven my work and will continue to contribute to my future learning. There are many opportunities to add to the body of work in the area of grading and assessment. While there has been an increase in the volume of research that has focused on classrooms grading and assessment, much of it has been in the form of surveys intended to gather teacher views on grading practices (Hochbein & Pollio, 2016). My ability to produce valuable research will be enhanced if I continue to grow and nurture my understanding of how teachers approach the process of making changes in grading and assessment. Throughout

the research process for this study, I needed to be keenly aware of my biases and how they might impact my work, and as I continue my research efforts in this area, I need to be always cognizant that my bottom line concern, and what should be the bottom line concern of all educators, should be to continually refine what is best for students and what will best equip teachers to teach effectively, and to employ assessment and grading practices that are fair and equitable for all students.

Ethical Considerations

I recognize the important role the participants played in this research study. In fact, there would not have been a study without their willing participation. Creswell (2014) emphasizes the importance of ethical considerations in designing a research study for the protection of the participants, for the development of trust, and for the promotion of integrity within the research. He points out that there are ethical considerations at all stages of research and presents ethical issues and possible ways to address them for each stage: prior to conducting the study, beginning the study, collecting the data, analyzing the data, and reporting, sharing and storing the data.

Prior to conducing the study, I obtained approval through the Institutional Review Board (IRB). I completed the Collaborative Institutional Training Initiative (CITI) Human Subjects Research Training course in August 2018, so I have fulfilled the IRB training requirements. The IRB includes a process for informing participants of their role and includes a form for participants to give their consent as well as a process for them to withdraw their consent at any time. I gained permission from the principal at Ashland Elementary, and she communicated general information with her staff to ensure that there would be sufficient interest in participation.

As I began the study, I established complete transparency with the staff at Ashland in sharing the purpose and process of the study as we sought volunteers. The option to participate was voluntary, and the teachers knew exactly what they are volunteering for prior to giving their consent. They also had the expressed right to opt out at any time. I made the research proposal available to anyone who was interested in learning more about the study prior to their committing to participate.

While collecting, analyzing, and reporting the data, I developed an interview protocol with general wording for the questions so that I was consistent in the initial line of questioning. I allowed the participants to select the time and place of the interviews to ensure that they were comfortable and to make sure I minimized disruptions to their daily routines as much as possible. I was conscious of ensuring that I reported all perspectives even if they were contradictory. As detailed earlier in reference to trustworthiness, I utilized triangulation, member checking, and peer review of the coding to ensure accuracy of the data. I explained to the participants prior to their interviews that they would have an opportunity to member check my data and that they would have an opportunity to withdraw any information with which they were not comfortable from the study after reading my interpretations of their perspectives.

Reciprocity

Curry (2012) defines reciprocity as the "cooperative exchange of help in which two parties arrive at an arrangement where everyone benefits" (p. 92). When addressing reciprocity in qualitative research, the researcher considers the honor and care of the participants as well as any possible benefits the research might accrue to the participant. Some conventional practice has viewed reciprocity in terms of monetary compensation for participants in order to acknowledge their contribution (Trainor & Bouchard, 2013). Curry (2012), however, suggests a

more collaborative relationship focused on mutual benefit when it comes to research in education. That is, the researcher will obtain the data needed for the study, but the participant will also benefit from the lessons learned during the conduct of the study.

The primary benefit to the participants in this study is the support and experience they received as they learned to utilize a Learner Centered Design Model to support change in assessment and grading practices. The participants benefited from taking part in the five or more virtual training sessions, two of which were led by Cale Birk, the creator of the LCD Model that guided their experiences, augmented by the opportunity to ask Cale questions, making the training sessions truly interactive. In addition, the team members had the opportunity to engage with their peers in a collaborative way as they learned a new process that may be helpful to them in the future. In terms of long-term benefit, it must be emphasized that the LCD Model can be used to support any type of change as long as one shares the effort with a group of like-minded colleagues willing to invest the necessary time and energy to craft meaningful solutions to whatever problem of practice they mutually agree to target. The entirety of the Ashland Elementary School will benefit by having a team of staff members on site who can provide leadership in supporting future LCD work.

Finally, again considering the long-term, the participants in the study and the school at large will accrue the benefits derived from the team's experiences in a genuinely collaborative effort that will, if the results are deemed useful, result in assessment and grading practice changes that will improve teaching and learning, both of which will benefit the primary recipients of quality education — the young people whose minds teachers and administrators are charged to nurture.

Limitations

This study was conducted in a small school district with a fairly homogeneous population. The students in this district perform well academically, and parent support is high. The need and the sense of urgency to address the issues of equity associated with some of the current grading and assessment practices are not as high within this district as they would be in many other settings. Despite the fact that change in schools takes a lot of time, this study will be conducted during one school year, limiting the potential extent of change. The researcher acknowledges that these factors may impact the motivation of the teachers to remain engaged in change.

The fact that this study was conducted in one small school also means that it is not necessarily scalable, and there will be no effort to generalize the findings to other schools and districts whether or not they are similar in size and demographics. The researcher does hope, however, that the ideas and insights drawn from this study can be used by other organizations for guidance as they engage in similar work in districts that will be larger and have more diverse student populations.

CHAPTER IV: RESEARCH FINDINGS

Leading Change With a Learner Centered Design Model

Change in education is difficult, especially when the change involves practices that are deeply embedded in the historical context of education. The practices associated with the assessment of student performance and the assignment of grades have remained essentially unchanged for over 100 years despite the fact that they are not appropriate or responsive to the changes that have occurred in the purpose and scope of public education (Guskey and Brookhart 2019). Grading and assessment practices have passed from generation to generation as students grow up to be teachers and parents who continually expect and apply and the same practices they experienced. The teaching profession, unlike other professions, has been learned by educators beginning in their own childhoods. Lortie and Clement (1975) identify this phenomenon as the "apprenticeship of observation." Change in education is made extremely difficult because the traditions of the profession are rooted in the lifelong educational experiences of the very people responsible for guiding and shaping the profession.

Effective change in education needs to be focused on student learning and how teachers can support that learning. Many reform efforts in the past have focused on aspects of education that are too far removed from learning. It is important for change efforts to be systemic and connected so that there is coherence and direction which is clear to all stakeholders. Michael Fullan (2001) warns that "the main problem is not the absence of innovations but the presence of too many disconnected, episodic, piecemeal, superficially adorned projects. Rather than contributing to substantial improvements, adopting improvement programs may also add to the endless cycle of initiatives that seem to sap the strength and spirit of schools and their communities" (p. 109). It is important to situate change efforts within the context of the

71

organization's existing priorities while building on strengths that already exist within the organization.

The purpose of this study was to determine if the use of a Learner Centered Design Model to drive changes in grading and assessment will impact changes in thinking and practice within a school community. The Learner Centered Design Model is a series of 5 phases designed to support change or initiative by involving the members of the organization who are most impacted by the change. It was developed by Cale Birk and Charity Allen (2017) and is described in their book, *Changing Change Using Learner-Centered Design: From Failed Initiatives to a Change Process that Connects, Empowers, and Actually Works*.

Looking at this problem of practice through this framework and the Learner Center Design Model can help answer these questions:

● In what ways does a learner design system approach work to support teachers in changing their perceptions about grading and assessment practices?

● What is the evidence of changes in grading and assessment practices as a result of a teacher's participation in a learner design system?

Pre-Intervention Interviews and Document Analysis

In order to assess the teachers' current perceptions and practices about grading and assessment, I conducted one-on-one interviews with each of the five teachers involved in the study on November 18, 19, and 20, 2019. I asked them questions about their personal experiences with grading and assessment as well as their thoughts about how they utilize assessment and grades in their current practices. Table 1 provides basic information about the teachers and their teaching experience along with the pseudonyms by which I referred to them while summarizing and analyzing data.

Table 1

Participating Teachers

Pseudonym	# of years' experience	Job assignment(s)
Jill*	1 year at private Pre-School, and 20 years at Ashland	1 year in pre-school, 4 years in 5th grade, 16 years in 1st grade (current assignment)
Lacy*	7 years at a school district an hour south of Ashland, 7 years at a local four-year university, and 16 years at Ashland	7 years in 4th and 5th grade moderate to severe disability classrooms, 7 years supervising student teachers and teaching methods courses, 11 years moderate to severe middle school students, and 5 years elementary learning disability (current assignment)
Julie*	3 years at Ashland	3 years in 2nd grade

(Table continues)

Pseudonym	# of years' experience	Job assignment(s)
Tim	1.5 years in a nearby rural community, and 7 years at Ashland	1.5 years in 4th grade, and 7 years in 3rd grade (current assignment)
Jenna	8 years in Ashland	2 years in middle school, and 6 years in 3rd grade

* Member of LCD Team

The interviews were conducted in the classroom of each teacher so they had access to any resources they wanted to reference or share. I recorded the interviews using an application on my phone and then transcribed each recording into a Word document. The teachers appeared to be comfortable and were very willing to share their thoughts. Each interview was framed around the same set of questions, and each interview lasted approximately 45 minutes to an hour.

I identified two big ideas to begin organizing the teacher responses: purpose and meaning of grades and assessment and grading practices. For each interview transcript, I highlighted teacher comments to be coded, categorized, and then to generate main themes. The two main ideas were categorized into the following themes: personal recollections regarding grades and assessment, purpose of assessment and grades, re-assessment, homework, and perceived accuracy of assessments and grades. As part of the analysis of the data, I triangulated the interview data with my document analysis of the assessments and grading documents provided by the teachers as well as the data from my observations of the early LCD meetings.

Personal Recollections Regarding Grades and Assessment

While none of the teachers had well developed memories of how grading and assessment were implemented when they were in elementary school, there were some commonalities within their responses, including how grades made them feel. For example, Jill and Lacy remembered reading groups and mentioned the fact that students knew exactly what it meant to be in each group. Lacy pointed out that "no one wanted to be in the turtle group because they knew what that meant." Jill shared frustration about not being in the group she thought she should be in. Lacy, Julie, and Jenna had memories of grades they associated with a content area that was a struggle for them. Lacy said that math was horrible because she was not good at it and got low grades, and Julie stated that when she started getting C's in math in third grade, she concluded that she was bad at math. Jenna recalled not being a good speller and getting lower grades in spelling.

The teachers were able recall some details about how they were graded and assessed when they were in school. Jill and Tim mentioned that grades were always based on an average. Lacy talked about the emphasis placed on getting good grades because "scholarships were tied to them and they could get you in college." Jenna did not remember getting ABCDF grades and thought that a different set of letter grades was used for all of elementary, whereas Julie and Tim specifically mentioned ABCDF grades being used. Jill and Lacy specified that grades were always based on pencil-paper activities. Jenna remembered the emphasis being different because "grades in elementary used to be mostly about character building and learning behaviors."

Purpose of Assessment and Grades

When asked about the purpose of assessment, there was agreement among the teachers interviewed. All five of the teachers indicated that assessments provide them information about

75

what students know and can do or whether the student has mastered the content or skill. Jill said that "assessments help us know what the students know." Lacy said she "uses assessment to see if they know the skill. It is important to understand where the kids are and what they understand from what you have taught them. Julie indicated that "an assessment is gathering information on students so I can know whether they know the content." Jenna shared that she uses assessments to "see if they're understanding the material that's covered." In addition to knowing what students can do and what they have mastered, Tim and Jill added that assessments reflect the teaching. Tim said assessments "are a reflection of what I am teaching them. If the majority of the class is missing a skill, maybe we need to re-teach it," and Jill said "if students don't do well, I may not have done my job properly."

The teachers also referenced assessments as being formative and ongoing. Jill pointed out that assessments "should not be a one check, one look; they should be continual, and the results should be able to change." Jenna talked about the use of assessments to determine what students know before teaching the unit. Tim talked about using assessments to determine the pacing of instruction or to change the strategy used for instruction.

There was some disparity in the answers when the teachers were asked if the assessments were only for their purpose or whether they served a purpose for students as well. Julie and Tim saw a definite purpose for students, and Julie included parents as consumers of assessment information by saying "I use assessments to report out to parents, to communicate to them if their child knows the content." Jenna pointed out that although assessments are usually a confidence builder for students, sometimes assessments serve more of a negative purpose for students because they make them self-conscious. In contrast, two of the teachers were definitive in saying that assessments were for their information only. Jill said, "I do not know that our

www.manaraa.com

students see or understand the meaning" and Lacy said "Assessments are for me. The kids don't get caught up in point grubbing." Julie was more undecided about whether the assessments served a purpose for students, saying "assessments do not always serve the right purpose for students. I am not sure they are understanding them the way that we are intending them to."

The teachers' views on the purpose of grades were not as consistent as their views about the purpose of assessment. Jill, Julie, Jenna, and Tim specifically said that grades offer communication to parents. There was concern raised, however, that this communication may not be accurate. For example, Lacy said "they don't necessarily tell you what a kid knows and understands." Similarly, Tim said "the grade on the report card may communicate to a parent that their student is doing well, when really they are only doing well in one part," and Lacy said "You can't tell by looking at a grade how much support has been given."

While there was an overarching theme of grades being used as a vehicle for communicating how a student is doing in a general sense, there was some deviation in thought about what it is that grades communicate. Jill said "grades help parents know where their students are and what needs to be worked on, what they've mastered and what still needs work," butt Jenna and Lacy saw grades as a way to communicate labels or comparisons for parents to understand where their children rate in comparison to other students. Lacy said "Grades are a label. We have to label somehow." Similarly, Jenna said "grades communicate what parents are looking for, kind of how their child compares to other third graders. They want them to be at the top."

The sample assessment documents provided by the teachers show some alignment with the teacher comments about the purpose of assessing and assigning grades. The assessments did, in fact, provide students with multiple opportunities to demonstrate mastery of skills or

www.manaraa.com

concepts, although I did note some inconsistency in their stated purpose for assessments and how the results of assessment were reported. For example, most of the assessments covered multiple skills or concepts, yet only one score was generated and recorded in the gradebook. This would make it difficult to use assessment to determine specifically what students know and can do. Julie and Lacy did have a separate document they used to record student mastery of skills or knowledge they observed in the classroom, but they did not use these documents to break down the written assessments into skills.

Re-Assessment

When asked about re-assessment opportunities, there was complete agreement among the teachers interviewed that re-assessment should be an option and that they offer opportunities for students to re-assess in their classrooms. This is consistent with their responses about the purpose of assessment in the sense that there was agreement among the teachers that assessments help teachers track what students know and can do. Their responses reflected a genuine desire to ensure that students learn the content and skills. Julie said that reassessment is "a really good thing, especially for students who really want to have the opportunity to learn the content and show that they know it." Lacy said, "I agree with reassessment – the opportunity to go back and relearn some of the material." Tim pointed out that retakes are important for students "because a lot of times they make silly mistakes or they misread the problem, or maybe they were just exhausted when they took it."

There was some disagreement about the conditions or policies associated with re-assessment opportunities. Specifically, they did not show complete alignment in their responses to a question about how many times a student should be able to re-assess and how the re-assessment score should be calculated. Lacy and Jenna said they believe students should be able

78

to re-assess once, whereas Jill, Julie, and Tim indicated a desire to give students unlimited opportunities. Jill, Julie and Tim did preface the desire to provide unlimited opportunities by pointing out that providing unlimited opportunities may not be realistic for teachers. Julie suggested a limit of three to four opportunities per quarter, and Tim said "ideally students would retake as many times as they need to in order to show they can do the work at a high level, but the reality is that we have to move on." In response to a question about how the grade of one or more re-assessment(s) should be calculated, Jill, Julie and Lacy said they replace the grade and Jenna and Tim said they give students half credit. The teachers did seem to be aware of and sensitive to these inconsistencies, because during the LCD work, two of the current concerns or challenges they identified were related to grading policies or retakes. One challenge they noted was *grade reporting consistency between grade levels*, and another was *developing consistent policies for re-takes*.

Homework for a Grade

The use of homework in general, let alone the use of homework for a grade, was not very prevalent among the staff I interviewed. Four of the teachers indicated that they give some homework, but the homework does not count toward the grade. Lacy, who said that homework does count toward a grade, indicated that it is a completion grade and that homework that is sent home is not work requiring extra support because it is intended to reinforce what they have already done in class. Jenna indicated that homework is not taken for a grade, but that students can be fined within their class currency structure for homework that is not completed. Julie stated that while she does not issue grades for homework, the rest of her team does. Jill and Tim pointed out that, since their homework is practice and reinforcement, student performance is likely to be impacted and can affect their overall grades.

The fact that the team identified *grade reporting consistency within and between grade levels* during the phase of the LCD model during which they articulated why they believe a change is necessary, indicates a desire and a belief that it is important to have consistency in how teachers approach grading. The use of homework as a contribution to grades is one consideration for the team when attempting to achieve consistency.

Perceived Accuracy

In order to understand their perceptions about the accuracy of grades, I asked the teachers questions about whether they believed a student who struggles with the content who is assessed and graded might end up with a high grade or that a student who is proficient with the content might end up with a low grade. I also asked them what contributes to the student grade and what the grade communicates about a student.

Their responses were somewhat contradictory, although the overall sentiment was that their grades were accurate. For example, Jill shared that they had stopped giving grades for time spent reading in order to be more accurate because they realized that they were really giving a score based on whether parents were writing down reading times and not based on student performance. Lacy shared that their grades are based on performance and that points are not given for participation, which makes them more accurate.

When asked if students might receive a grade that does not match mastery of the content, some of their responses uncovered possible inaccuracies. Jill, Julie, and Jenna indicated students will receive a grade that matches their level of skill and knowledge in the content area. Jenna, however, also said that if students participate and pay attention, they will get an A or B and that some students get a D or F because they do not pay attention. Tim said "a student with low content knowledge in math could get a high grade because that student might do well on the

fluency quizzes which count five times as much as the skill checks. Lacy shared that student grades are often inflated because of the level of support they receive and are, therefore, not consonant with their content knowledge.

In looking at the documents shared with me, there was evidence of variance in what contributes to grades and how grades are calculated. For example, each grade level uses different symbols to communicate student grades. In first grade, students receive the following marks: E for excellent or above grade level, S for satisfactory or at grade level, N for needs improvement or below grade level, and U for unsatisfactory or well below grade level. In second grade, the same marks are used for grammar, social studies, and science, but A B C D F is used for reading, math writing, physical education, and music. In third grade, the marks used are A B C D F, and are based on percentages.

When analyzing gradebooks and what they contribute to grades, I focused on math and reading because the assessment documents shared with me were from these two content areas. In both subjects, in all grade levels, there was one grade calculated from a variety of gradebook entries. Each grade level had a different combination of contributing assessments and assignments, but there were some similarities. For example, at all grade levels, math grades are a combination of skill tests or quizzes graded on points, and fluency checks or timed tests graded on points. In third grade, they also utilized grades from weekly cognitively guided instruction problems that graded on rubrics. In all grade levels, the reading grade was a combination of vocabulary quizzes or sight words graded on points and skill tests based on specific learning standards graded using points. In third grade, there were also Accelerated Reading grades based on whether students accumulated enough points according to goals pre-set for the grade level.

81

Learner Centered Design Model Observations

I was an observer-participant during four three-hour meetings designed to support the team, which was made up of Jill, Lacy, and Julie, in working through the Learner Centered Design Model. The meetings took place on November 13, 2019, December 11, 2019, January 21, 2020, February 18, 2020 at Ashland Elementary. Cale Birk, author of *Changing Change Using Learner-Centered Design: From Failed Initiatives to a Change Process that Connects, Empowers, and Actually Works* (2017) consulted with the team through a Zoom call as they worked through the first two phases of the model.

Phase 1- Appreciate

During the first phase, *appreciate*, which involves taking the time to empathize with, learn about, and truly understand the situation and what it is that is happening, the team reflected on questions posed by Birk. He started with the basic question of "Why change?" The teachers discussed a desire to have grades that reflect student learning and to stop doing things simply because it is how they have always done them. Birk followed up by asking, "Who is this a problem for?" to which Julie immediately responded by saying that the superintendent was hearing reports of grading practices that were ineffective and unfair to students. When Birk responded with questions about how they decided it was an issue, who the biggest resisters are, what the biggest challenges are, and who is impacted, the teachers began to stop and think more about their answers and engaged in exploratory dialogue about who the users of the grading and assessment are and what the challenges are for each entity. The group talked a lot about teachers as users, but Birk pushed them to also think about the role of parents as users. There was a realization on the part of the group that while they had thought about how parents might react to a change, and had planned to include parents in communication about the change, and had

www.manaraa.com

planned to provide support for parents to understand the change, they had not thought about soliciting information about parents' current views, their understanding of the inherent problems involved in grading and assessment, and their needs for information flowing from these processes.

Julie explained that the team had already taken steps to explore and understand teachers as users prior to the start of the study by engaging the teachers in exploratory discussions during institute days and staff meetings. They completed this work as part of their responsibilities as members of the district leadership team. Jill also shared that many of the teachers in the school had engaged in a book study on Ken O'Connor's book, *A Repair Kit for Grading: 15 Fixes for Broken Grades.* From that book study, the district leadership team learned from the teachers that they had a strong desire to ensure that grades accurately reflect student learning, but they did not have a clear picture of what that should look. In addition, the teachers indicated that they would need support in making any kind of change. There were concerns from teachers about increased workloads and whether the grading program would work with any changes recommended and implemented.

When discussing the parents as users, Birk asked the teachers what concerns they think parents might have about changes to grading, assessment, and reporting. Lacy and Jill shared that they have heard parents express concern about college entrance and college scholarships and their perception that colleges will only consider traditional grade point averages in making admissions decisions. Lacy also reported that some parents worry that allowing students to re-take assessments will cause students to develop poor study habits because they will choose not to study before the assessment with a plan to re-take it if they do not succeed. Julie pointed out that

parents do not realize that grades are not accurate and would likely not understand the need to make changes.

Next, Birk asked the teachers how they might collect information about the parents' views, understanding, and needs in relation to grading and assessment to see how well they align with the perceptions of the teachers. Julie suggested an online survey, and Jill and Lacy agreed that a survey was a good way to collect the information. Jill suggested they send home the current report card and a mockup of what a new one might look like, but the discussion led the team to decide that there were a couple of concerns with that approach. First of all, it might invite comments and suggestions from parents that are unrelated to the subject at hand and fail to focus on the changes they are considering. Secondly, it could send the message that the new way of reporting has already been developed. Birk directed the group to decide what it is they want to know about parents' user experiences. The teachers talked about effective communication and what that would look like and how they could learn from parents about their vision of successful learning. They decided they wanted to know what information about their own children's performance parents most valued and how often they would seek that information. They developed a survey (Appendix A) and decided to send it to parents using Google Forms. The team used the results of the survey to inform their thinking about what a solution would look like.

Phase 2- Illustrate

The next phase, *illustrate*, during which the team sets clear goals and targets for a solution was accomplished through questioning about how they would use information gathered from teachers and parents, along with brainstorming from the team in response to the following prompt: *The ideal grading system will....* Birk asked them what the "pain points" in their

grading system are, meaning what aspects of the grading system would create sufficient conflict for the Ashland community that they would seriously consider a change. Jill talked about the inconsistencies in grading symbols and scales between grade levels and the challenge with setting up Skyward to meet all of the different iterations. Jill and Julie discussed the need to hold students accountable and the importance of communicating student efforts and behaviors. Lacy expressed concern about how the supports that some students receive are factored into grades. She said "How much is too much when it comes to accommodations? I read assessments to my students when I realize the reading is getting in the way. Should their grade then be adjusted?" All three of them discussed the desire from parents and staff to have grades that are clearly defined and consistent. Julie pointed out that it wouldn't be too difficult to grade specific skills while the students are learning. She said, "Why can't I just ask a student about the main idea of a story we are reading and enter that as a grade?"

From their discussion, they began to develop a framework to guide their journey to a solution. On large post-it notes, they wrote the functions of grades and assessments from the perspective of the teacher, the student, and the parent. The concepts that appeared in all three lists, included clear communication, supporting improvement. Jill pointed out that if communication is important, they need to know that what they are communicating is accurate. Julie added that to make the communication accurate, it is important to separate academics and behaviors. Julie and Lacy both shared that if grades are going support improvement, one-time assessments will not suffice. Jill brought up the fact that part of accurate communication should also include grades that are specific to skills. Julie shared that, in order to support improvement, it would be helpful to have methods for showing growth over time. They did not work on naming a solution yet, but rather they established parameters and descriptors for what it is they

www.manaraa.com

want the solution to encompass. The team decided that the descriptors for the ideal grading and assessment plan are that it will:

- Be clearly articulated and communicated

- Be based on Ashland Essentials that are expressed in understandable language

- Include the use of multiple probes to assess student learning and progress

- Accurately describe student performance at a given point in time

- Identify specific areas of need for individual students

- Demonstrate student growth and progress over time

During this discussion, the teachers' final product was a PowerPoint presentation that identified and fleshed out ideas about what they believed an ideal grading and assessment system would look like (Appendix A). They shared the presentation with their full staff during an institute day in January and asked for feedback from the entire staff.

Phase 3- Ideate

In the third phase, *ideate*, the team began to discuss and develop possible solutions based on the information they had gathered about the user experiences of teachers and parents, and on the feedback they received when the staff shared their vision for an ideal grading and assessment system. During this phase, the team engaged in brainstorming to develop ideas. Jill suggested establishing a consistent protocol for student re-assessments as a response to questions from the teachers about how to hold students accountable while also allowing them to re-assess. Julie pointed out that it might be best for them to start by making sure their assessments are designed to provide specific information that would inform the re-assessments. She shared that the parent responses on the survey indicate that the parents understand that when they know how their child is doing in specific skill areas, they are more prepared to support their learning. She suggested

86

creating assessments that are delineated by skill so that the teachers would be supported in their ability to provide grades that were specific and well-defined. Jill agreed with Julie, saying, "We know we want grades that are specific and accurate. We probably do need to start by making sure our assessments are specific and accurate. We can start by developing strategies to assess specific skills." The team noted that the parent survey supports this priority. When asked which would be most helpful in supporting their child in learning, *knowing my child's current academic performance on a specific skill* had the highest number of responses. Julie and Lacy brought up the fact that parents clearly want to know about their children's behaviors and work habits. On the parent survey, the second highest response to the question about the information that would be most helpful for them in supporting their child's learning was *knowing their child's behavior at school*. On a similar question on the parent survey, *information about student respect at school* was chosen as the information parents most wanted to see on a report card. Lacy pointed out that the staff expressed a desire to have clear guidelines for assessing and reporting behavior and work habits.

After brainstorming solution ideas, they worked to narrow their list based on the goals and targets that were set during the *illustrate* phase. The team decided to start with the creation of clear protocols and communication tools to use when assessing and reporting student behavior and work habits separately from student academic progress. They developed a document to support teachers in monitoring behavior and work habits (Appendix C). They also decided to focus on assessing and grading specific skills and concepts. At the district level, they had already begun to identify Ashland Essentials, which are the standards that must be mastered at each grade level. Their idea was to use the Ashland Essentials to begin narrowing the focus of

their assessments and grading as a first step towards creating grades that are specific to skills. (Appendix D).

Phase 4- Iterate

Once they decided on the ideas they wanted to implement, they entered the *iterate* phase. During this phase, they began to implement the solutions they developed, which included the staff-wide use of the document created by the team to support teachers in the reporting of behaviors and the development of grade level essential skills and concepts that will then be used to narrow the focus of instruction and assessments.   The plan was for the team to implement the use of both tools and to encourage volunteers in the staff to also try them so they can make adjustments based on their experiences and on the feedback they receive from staff and parents. The team shared the two documents with the staff and asked them to try using them as they assess their students. During the session to plan this phase, the team decided that they did not want to rush this phase but wanted to spend as much time as necessary for trial, error and adjustment. One point that was made by all three members of the team was that the shifts the teachers will be making are changes in thinking and practice that have been developed and internalized into their standard operating procedures over many years, thinking and practice that have become very comfortable and automatic. The team wanted to allow enough time for the changes to become embedded, and they wanted the teachers to have the time and mental space necessary to adjust and to problem-solve when concerns or roadblocks developed. I chose to conduct the final interviews to look for changes in thinking and practice as a result of their work with the Learner Centered Design model while they were in the early stages of the *iterate* phase.

The purpose of the study was to assess whether there was a change in the educators'

beliefs and practices related to grading and assessment as a result of their participation in a LCD

model. I sought to gather evidence of change or lack of change through teacher reflections, final

interviews, and a document analysis. The interviews took place on April 16, 17, and 20, 2020.

The interviews were conducted via video conference calls because of social distancing as a result

of the 2020 COVID-19 pandemic. I recorded the interviews in the same way that I recorded the

face-to-face interviews, using an app on my phone to record audio only, after which I transcribed

the interviews into Word.

In order to analyze the data, I highlighted the teacher comments I targeted for coding and

categorization. I started with the same two main big ideas I used with the pre-intervention

interviews: purpose and meaning of grades and assessment and grading practices. I created a

separate category for statements that did not fall into either of these categories but referenced a

shift in thinking about the change process itself. I placed the participants' comments side by side

in a table and extracted phrases from the post-intervention interview that deviated significantly

from the pre-intervention interview, suggesting a change in thinking or practice. I then

categorized these statements into two main themes: changes in practice, and changes in thinking

about communication and efficacy.

Change Process

The participants specifically mentioned shifts in their thinking about the change process

itself. Julie, who was on the LCD team expressed appreciation regarding the pace with which

they have decided to approach the change. She said, "I do like that we're going slow. I do love

the slow process of this because I feel like we need to change a lot of what we're actually doing

in the classroom first, and the ideas that we have about it and the culture around it before actually doing the documentation and all the logistics." This aligns well with the team's decision to stay in the iterate phase for an extended period of time in order to allow for a change in the culture and for teachers to have time to make changes in their classroom.

Lacy, who was also on the LCD team, also mentioned the slower pace of the change process and indicated that they actually took a step back in one sense. That is, she reported, the decision had been made to move forward with a change in the reporting of grades only for essential skills as well as a separation of behaviors and academics for kindergarten and first grade, while second and third grade will only focus on the separation of behaviors and academics when reporting grades. Jill, also a member of the LCD team, echoed this point and said "I think that kindergarten and first (grade) were already doing a version of it. It just wasn't quite the way it's going to be, but it still wasn't as big of a jump for us to change in kindergarten and first (grade), but second (grade), third (grade) are in a different place." Tim, who is not on the LCD team, pointed out that his team appreciates the fact that they will have more time to work on the identification of essential skills.

Julie and Lacy shared during their interviews that one change in the process that stood out for them was the inclusion of *parent voice*. Julie said, "I honestly had not even thought about surveying the parents to get their thoughts, but I am so glad that we did." Lacy reported that she was worried about surveying parents because she thought it might send them in another direction but she was "pleasantly surprised to realize that the parents' thinking was actually very much in line with ours."

90

Change in Practices

The documents created by the LCD team to support the separation of academics and behavior (Appendix C) and to identify essential skills (Appendix D) are evidence of attempts being made to change grading and assessment practices. There was also evidence of changes in practice in the statements made by the participants during the post-intervention interviews. In some cases, the participants specifically mentioned changes they had made, but in others, I inferred a change in practice based on differences in their responses from the pre-intervention interview to the post-intervention interview.

Julie said that she previously had very loose guidelines on re-assessing and would let students re-assess at any time because she was cognizant of the fact that the goal is for students to learn all of the targeted skills. She has since changed her practice to align with the work her team has done identifying essential skills. She indicated that they plan to recycle the essential skills throughout the year, which means that students will have multiple opportunities throughout the year to show that they have mastered one or more of the identified skills. Therefore, rather than having endless opportunities to re-assess a skill, she has established a consistent timeline. She said that based on the pattern of one new skill per week, that "within two weeks; that week and then the following week, the students have time to reassess." She also shared in her reflection that she used to grade a lot of formative assessments that were given right after the students learned new skills or concepts. She has realized that it makes more sense to wait and grade the cumulative assessments that occur after the students have had more time to practice and internalize skills.

In her first interview, when asked how she determines a student's grade, Lacy talked about grades coming from quizzes, weekly assessments, and spelling tests. In the second

interview, her answer centered more around informal assessments. In fact, she said that because she has a smaller group of students, her assessments do not have to be formal written tests. Instead, she discussed the use of "informal assessments like a game or a matching type process of things or an activity that we can do in a small group." In her reflection, Lacy also discussed the work her team is doing to change the way they assess and report student progress to parents. She wrote that they are "looking at changing assessments to reflect the skills we are working on." Similarly, Jill and Jenna talked about the use of checklists or some kind of organized way to informally collect student data. Jill shared that her team is using more checklists and informal observations to assess students on specific skills and Jenna talked about trying to use documents to help organize her observation and anecdotal records. They both pointed out, however, that "how to do this effectively will take some work," suggesting that the change in practice is still a work in progress. Jenna specifically mentioned how hard it is to "change her way of doing things" and that she has not yet found a tool she can use consistently.

In the pre-intervention interview, Tim mentioned that he "would like to see report card grades as a breakdown of the skills," but that right now, "the grade on the report card is just an average." During the post-intervention interview, he shared that his team has spent time identifying the most important skills they want their students to master before moving on to the next grade level so they can narrow their instructional practices as well as developing a way to more efficiently and accurately reporting to families and students. At this time, their gradebook and their report card have not changed to reflect this work, but there are suggestions within their work that point to a change of practice that will be implemented over time. Another change of practice that Tim discussed in his reflection was the separation of behavior skills and academic skills. He wrote that it will be a big change for his team, but that "It will really make a

difference to focus on those skills separately and make sure we aren't counting them into the main grade." Jill also talked about how separating behavior from academics will be a great way to celebrate hard-working students.

Communication and Efficacy

In observing the LCD meetings, reading the reflections from the participants and listening to their responses to the pre-intervention and post-intervention interview questions, there was a common theme around communication. They talked about grading and assessment as a form of communication, about keeping lines of communication open with staff during the change process, and about clearly communicating the change with staff and the Ashland community. During the course of the study, however, I detected a subtle change in their thinking from a focus on the communication of facts about student achievement to a focus on communication that is more efficacious in promoting student learning.

Lacy shared how her thinking about what a grade is communicating to her as a teacher has changed. Previously, she thought, "Well, if they fail it, they just don't know it and that's just how it is." Her thinking has now shifted to, "Well, what can I do to help them reach the point at which they are more capable of expressing that they do know it?" She now uses the information to write new goals and to change her instruction in order to provide her students with more and different ways to be successful and to show that they are learning. Jenna also verbalized a change in her thinking. She said that, whereas she used to think that students who showed mastery of the content earlier should get higher grades because it shows a higher ability, she now realizes that might not be accurate and says, "If we think about what the goal is, which is for students to learn, then maybe it doesn't matter how quickly they show that they have learned it."

Julie and Tim both talked about how this change process has led to more team discussions about what they are doing with grading and assessment and why they do things the way they do. For example, Julie said that her team has been talking about why they double points on some assessments, why they take grades for homework, or why they take grades the day they teach the material. She said that the important thing is "Why are we doing all of it? Well, to tell parents and to tell students and to tell myself where the student's at and how we can help them." She expressed appreciation for the conversations her team is having about this purpose and how their grading and assessment align with the purpose. Tim also shared that his team has had conversations about what they are assessing and whether it is communicating what they actually want it to communicate. Jill talked about how the essential skills her team has identified will help them know as teachers if students are learning what they intended to teach.

Julie and Tim also talked about how their communication can be more specific and effective for parents and students. Tim said that from the work his team is doing to identify essential skills "we hope to create a more informative grade report for families to identify how their students are doing based on specific skills." He also mentioned that he thinks they can do a better job of using that information to teach students to gauge their own learning. Julie has worked to provide more communication with parents about how their students are doing on specific skills by using a 3-point scale, from a 1 that indicates little or no understanding to a 3 that indicates mastery, on all work that is sent home. One of her takeaways, however, is that she needs to be more consistent in how she uses the language of this new approach with parents because she realizes it is new for them, and they will need support in order to use the information effectively to help their children. She also expressed how she has realized that the best thing she

94

can do is to help her students understand how they are doing on a given skill and what they can do to improve.

Summary

The Ashland team's participation in the LCD model appears to have prompted some changes in thinking and practice related to grading and assessment. The teachers are beginning to make changes in how they assess their students, especially in their focus on specific essential skills and in their efforts to separate behavior characteristics and academic skills. They are increasing their use of informal assessments and they are examining their grading and assessment practices to ensure that they are communicating accurate information.

The teachers have also changed their thinking about the purpose and scope of assessments and grades as communication tools. Rather than grades being used to report how students compare to their peers or whether or not they are succeeding in general, they are looking at grades as a way to identify each student's strengths and weaknesses and how learning can be improved, thereby building self-efficacy for each student. They are using discussions about grading and assessment to increase communication within their teams about what is important and how they should structure their instruction and next steps, which is developing their collective efficacy.

The team is hopeful that their decision to slow down the process will allow for more time for a change in the overall culture in the building and for all of the staff members to implement these changes with integrity. The structures they have in place, including documents to support teachers in assessing behaviors separately and the work they have done to identify essential skills at each grade level will support this change in culture and in the implementation of new grading practices.

CHAPTER V: ANALYSIS AND NEXT STEPS

The purpose of this study was to determine if the use of a collaborative and user-centered process will have an increased impact on changes in thinking and practice around grading and assessment. The process used in the study is the Learner Centered Design Model, developed by Cale Birk and Charity Allen (2017). The Learner Centered Design Model is a series of 5 phases designed to support a change or initiative by involving the members of the organization who are most impacted by the change.

The conceptual framework supporting the study is a constructivist learning framework that combines Fullan's Educational Change theory, Hargreaves and Fullan's Professional Capital theory, and Heiftz's theory of technical verses adaptive change. The conceptual framework consists of a driver or goal as the seed that drives the change. The driver is supported by educators working in a space of social capital constructing new knowledge and meaning about what it means to assess students and assign grades in ways that are equitable and accurate. Through this constructivist learning and collaborative process, decisional capital is produced. The decisional capital, or the expertise of the educators, is utilized to form accurate and equitable decisions about student learning, thereby creating new grading and assessment practices.

The intent of the study was to look at this problem of practice through this framework and Learner Center Design Model in order to help answer these questions:

- In what ways does a learner design system approach work to support teachers in changing their perceptions about grading and assessment practices?

- What is the evidence of changes in grading and assessment practices as a result of a teacher's participation in a learner design system?

96

In order to truly analyze the results and determine the possible and appropriate next steps in creating meaningful change in grading and assessment, it makes sense to place research and its findings into the history of grading and assessment reform efforts as well as to discuss how to use the findings of research to guide future studies and inquiries. The history of any practice will always be an important part of a discussion related to change in education because of the fact that all educators experience school for most of their lives and develop their paradigms based on their experience and the experience of their families. According to Guskey and Brookhart (2019), we have maintained practices "not because they are effective, but because they are steeped in traditions that many stakeholders believe have served us well" (p. 4). In addition to understanding the history of a practice when considering a change, it is also imperative that we understand where we are now. If we do not have a thorough understanding of our current reality, we cannot possibly know how to achieve our desired results. It is the understanding and acceptance of the gap between our current reality and our vision for the future that will provide us with insight into the actual changes that must take place.

Where Have We Been?

The initial interviews with teachers at Ashland revealed that they have used traditional grading practices such as the calculation of one grade from scores representing a wide range of learning topics, skills, behavioral characteristics, and outcomes. Another reveal was that grades have been reported in completely different ways from one grade level to the next. There is consistency within each grade level in regard to which grades are entered into the gradebook but there was at least one instance where one teacher in a grade level used a different calculation for the grade by changing the value of the math skill assessments to make them more comparable to

the math fluency assessments. The team in this study expressed concerns about the accuracy of their grades and the consistency from grade to grade.

Concerns about the reliability and variation of grades have been the focus of studies almost from the time grades became ingrained into the school experience in the 1850s. Guskey and Brookhart (2019) found sixteen individual studies of grading reliability from the early 20th century, ranging from 1912 to 1935. The earliest statistical investigation they found, however, was a study published in the *Journal of the Royal Statistical Societ*y in the United Kingdom in 1888. The author of the study, Professor F.Y. Edgeworth, set out to make things fairer for students by using the normal curve theory to measure the degree of error in grades due to chance. What made his research notable, however, was the fact that he went on to measure the degree of error due to other factors, such as the difference in the graders' views of which knowledge and skills are most important and whether or not items on the exam were representative of the desired learning outcomes. The idea that errors in grading are attributed to more than chance, for which it is difficult to control, opened the door to the idea that educators can and should work to control factors that contribute to errors in the reliability of grades.

The studies in the early 1900s that Guskey and Brookhart (2019) analyzed examined variations in grading, including whether grades are affected by the quality of the student work or the format, or whether different grading scales would decrease variation. These studies consistently found great variation in the grades teachers assign to student work, and the authors generally attributed the inconsistency to one or more of the following sources: criteria for evaluating the work, the quality of the student work, teacher severity or leniency, differences in the tasks, the grading scale, or teacher error. All of these studies helped to draw attention to

factors inherent in grading that can be addressed when educators make a concerted effort to improve consistency.

In an effort to address multiple sources of inconsistency, a number of reform efforts in the area of grading and assessment have focused on achieving consistency by addressing the grading scale and or the calculations used to determine grades. Programs have been developed that use technology to apply algorithms to calculate grades and to limit teacher-decision making (Guskey and Brookhart, 2019). These programs and algorithms offer the pretense of objectivity as educators seek consistency and reliability in the grades they assign. If this objectivity could be realized, issues related to teacher severity or leniency and teacher error would be adequately addressed. Any efforts to develop an objective and accurate algorithm must also address the wide and extreme variances inherent in grading procedures and protocols. Educators have included a variety of contributing factors to their grades, including achievement, current performance, past performance, behaviors, effort, compliance, etc… and they have included them in very different proportions, and with very different criteria (O'Connor, 2007a). Accounting for all of these factors in order to develop a consistent algorithm that eliminates subjectivity has proven to be incredibly challenging.

One element of traditional grading has been the practice of assigning one summary grade or percentage for each subject at the conclusion of a grading period (O'Connor, 2007a). McTighe and Curtis (2019) makes the point that this practice diminishes the value of a student's achievements by distilling all learning experiences and demonstrations of growth and performance into a single point, thereby ignoring the richness of student performance. O'Connor (2017a) raises the concern that the practice of adding up scores and calculating a grade is counterproductive to the objective of learning for all and that it produces a culture of point

99

accumulation. When students are concerned about accumulating points, the focus shifts from learning to competition and playing the academic game because of a focus on the final score or grade and not on the processes of learning and mastery. This also causes all assessments to become summative, which reduces the chances that students will take risks and engage in difficult tasks for the purpose of growing and learning.

All of these inconsistencies, reform efforts, and concerns have led educators throughout the world to consider sweeping changes in how we execute grading and assessment, changes that have the potential to transform the entire landscape of calculating and reporting grades. There has been an evolving discovery about the power of grades to change the trajectory of a child's education and future. Grades influence not only how students perceive school, but also how they see themselves as learners. Grades are predictive of whether students drop out of school, whether they apply to and attend college, and their success in college (Brookhart & Guskey, 2019). For this reason, we find ourselves at present in a place in which grading and assessment are at the very heart of many education change efforts and at the center of many professional learning opportunities.

Where Are We Now?

The use of the Learner Centered Design model in Ashland created opportunities for collaboration among teachers as they worked together to develop a deeper and clearer understanding of the underlying context of their current practices and to use that understanding to identify possible solutions to implementing necessary changes. The process toward this understanding included contributions from teachers and parents and was not directed by any mandates or pre-determined outcomes. During one of the LCD sessions, one of the participants, Julie, pointed out that it was the superintendent who initiated their journey because of concerns

100

about inconsistencies in the content of grades and the method of reporting, but he has charged teachers with the task of charting a pathway to addressing those issues. The team was given the autonomy to work through the LCD process to find solutions that honor the user experiences of the teachers and parents.

Brookhart and Guskey (2019) identified seven studies that focused on principals who use a nondirective, collaborative approach with teachers, parents, district leaders, and other stakeholders to shape and change grading and reporting practices. One of the studies specifically noted the importance of "shared control" when transitioning grading and reporting practices. Another study identified the negative and limiting effects of principals using a directive approach. An examination of the seven studies, made it clear that a common factor for success was centered on collaboration, from teachers having time to collaborate and engage in shared decisions to collaborative problem-solving and decision-making between teachers, principals, and district leaders. Other factors that contributed to success gleaned by Brookhart and Guskey (2019) in their analysis of these seven studies related to making decisions based on guidance from stakeholders and not mandates from the leaders and the use of coalitions with teachers, parents, and school stakeholders early in the process.

The introduction of learning standards in the mid-1990s has provided a structure for possible change in the areas of grading and assessment. There is more consistency in content and there are common expectations for what students should be learning. The standards also provide an avenue for communicating assessment data that is more specific and more well defined. Systems like standards-based grading, skills-based grading and competency-based grading have emerged as methods to assess and report on student performance that are specific to stated goals or targets. Educators are recognizing the importance of separating behaviors from

skills when reporting grades if their goal is to truly report what students know and can do. The challenge has been, and continues to be, how to support educators in making these changes in practice. Without common language and understanding, a shared desire to change, and a responsive plan for change, educational communities are finding themselves spinning their wheels in traditions that are no longer meeting educational or societal needs and expectations. O'Connor (2007) rightly points out that grades are not likely to disappear any time soon and suggests that a critical component of any transformation in the area of grading will require a concomitant cultural evolution.

A school district about 100 miles from Ashland was able to utilize the learning standards to collect specific data on student achievement in the late 1990's through an extensive school improvement effort with a goal of improving learning (Sappington, 2008). Between 1997-2001, this district realigned their curriculum with the Illinois Learning Standards, while identifying "Academic Expectations", which was the content and skills that they expected children to know and be able to do. In doing this, they were able to develop formative assessments that provided them with very specific data about student learning. In turn, they used this information to guide instruction and communicate with parents. This school district had seven elementary schools with student populations that had a range of 44 to 83 percent low income (Sappington, 2008). During a time when schools struggled to close the low-income achievement gap, this district was able to do so. They were successful not because they taught to the test, but because they focused on learning.

Through the collaborative LCD process, one of the solutions the team decided to pursue was the establishment of essential skills and standards on which each grade level will focus when assessing and generating grades. Each grade level team collaborated to identify the essential

skills students must master before progressing to the next grade level, and then they created a common document to support their efforts to shift their practice in focusing on these skills and standards as they produce skills-based grades. The LCD team created specific guidance for teachers, in the form of identified behaviors along with success criteria, to use in their efforts to separate academics from behaviors. They have not yet made significant changes in their practice, but they have structures in place to support the changes they hope and expect to see.

<div align="center">**Where Are We Going- Next Steps**</div>

Through the use of the LCD model, the team at Ashland decided to focus on two aspects of grading and assessment. The first is the separation of academic achievement and behavior when assessing students and assigning grades, and the second is assessing individual skills in order to report specific grades. The solutions they developed for both focus areas involved the creation of guiding documents to support teachers. For the separation of achievement and behavior, the team developed a document that identifies specific behaviors and work habits along with rubrics to clarify the expectations. In order to support teachers in assessing individual skills, the team established priority standards and created a document that defines the standards and the expectation for student mastery. Use of these tools and continued dialogue within their teams will support the changes in practice Ashland would like to see in order to improve the accuracy and consistency of reported grades.

In their analysis of research related to grading and assessment, Guskey and Brookhart (2019) suggested the following implications regarding reliable grades: the importance of being clear about criteria, and the need to be consistent. According to Guskey and Brookhart, the criteria need to be clarified in such a way that they define learning rather than how well a student follows directions. Grades will be more consistent, they argue, when educators evaluate qualities

of the work separate from the criteria. Similarly, O'Connor (2007a) proposes that grading is effective when grades are accurate, meaningful, consistent, and supportive of learning. The work the Ashland team put into separating academic achievement and behavior will support them in clarifying their criteria and increasing the consistency of grades because they will be focused on learning, not on rules and strategies for capturing grades. Their identification of essential learning targets will increase the accuracy of grades and will enhance their meaning because the grades will communicate specific information about student performance on a defined skill or concept rather than on an arbitrary combination of skills or concepts. Identifying essential learning targets will also support learning because teachers, students and parents will have information about where students are in their learning progression for specific skills, allowing them to establish next steps.

Robert Marzano (2003) suggests that a "guaranteed and viable curriculum", or GVC, is a school-level factor that has significant impact on student achievement. He defines GVC as a combination of the opportunity to learn and time. By this, he means that students should have the opportunity to learn the content expected of them as opposed to being tested on skills that they may or may not have had the time to learn. The challenge, according to Marzano (2003), lies in predominately in the time factor. He calculated that the time needed to teach the 200 standards and 3,093 benchmarks that are contained in the national and state-level standards, it would take 15,465 hours for teachers to adequately address the standards for students in kindergarten through twelfth grade. However, he also calculated that a reasonable estimate of the amount of instructional time during that time span is only 9,042. Therefore, in order to have a GVC, it is very important to identify the content and skills that can be addressed and ensure that they can be addressed in the time available for instruction. The work that Ashland has done

104

in identifying essential learning targets will aid them in providing a "guaranteed and viable curriculum."

In discussing effective change leadership, Michael Fullan (2016a) suggests that leaders must become the masters of content and process given that "all effective change processes shape and reshape ideas as they develop capacity and ownership" (p. 9). The ideas are shaped and reshaped because other people within the organization have unique and varying knowledge bases and insights about how to improve the solution. The collection and synthesis of knowledge and insights from a cross-section of those associated with change is what creates social capital (Hargreaves & Fullan, 2012). Fullan (2016a) contends that "working together in focused, specific ways to learn from others to accomplish something of value attracts and promotes human capital that, in turn, further develops the group" (p. 35). Through the LCD process, the Ashland team has leveraged human and social capital as they have collected ideas and insights from the staff and from parents to develop and improve the solutions they have proposed. A missing piece, however, may be the voice of students, given that students represent a human capital perspective that is personally and directly impacted by grading and assessment. Their input, therefore, could and probably will influence teachers' thinking about grading and assessment in ways that are preeminently constructive.

Student Voice

In his book, *Indelible Leadership: Always Leave them Learning*, Michael Fullan (2016a) discusses the phenomenon of "students as change agents and proteges, which makes them lifelong learners, capable of helping themselves and humanity" (p. 31). He observes that students are more interested in changing the world and are less emotionally committed to the status quo. Perhaps they are in a better position than others to provide the momentum needed to

make changes to educational systems that are historically bound. Fullan (2016a) asserts that deterioration often occurs when efforts to transform the education profession and save students are large-scale, policy-driven efforts and suggests that we would be better served to let students save students through a partnership with professional educators.

Furthermore, specific to the changes in grading and assessment that were the focus of this study, there is evidence that student involvement in their own learning leads to higher levels of achievement (O'Connor, 2017). "Research and best practice tell us what some teachers have known for a while: When students are involved in the classroom assessment process, they are more engaged and motivated and they learn more" (Davies, 2007, p. 31). Students achieve at higher levels when they share in the process of constructing criteria, and therefore, have a clear understanding of what success looks like and what the expectations are for their learning (Black and Wiliam, 1998, Davies, 2017). In fact, Black and Wiliam found that one of the highest reported educational interventions involves students co-constructing the criteria, self-assessing in relation to the criteria and presenting evidence of their own learning.

In discussing the challenge of shifting legacies, Fullan (2016a) points to the fact that, even when they want to change, people often return to their old ways or they lack the clarity and confidence to act in new ways. Engaging the human and decisional capital of students and including them in the construction of new ideas about grading and assessment may lead to more effective and longer-lasting changes. Their collective voice will provide an important perspective, and their willingness to embrace change and abandon status quo will provide impetus to move a system forward that has had a lot of inertia.

Next Steps

  This study of the use of an LCD model has already produced some success in changing the thinking and practice of educators at Ashland, notably their work to separate behavior characteristics and academic skills and their focus on essential skills in their grading practices. Another change of thinking that occurred as a result of the process is related to the value of including parents in the process. Julie and Lacy both mentioned in their post-intervention interviews that they had not previously thought about including the parent voice in deciding how to shape the change process, but they highly valued the input they received from parents, and they believe it will make a difference in the level of support they receive from parents as they implement change. A next step would be to also gather student feedback throughout the *iterate* phase while the teachers are engaging with and implementing solutions suggested by the LCD team. The teachers could dialogue with students about their behaviors and how those behaviors may be impacting their overall success. They could also engage students in self-reflection about their assessments and discuss student performance on specific skills and then ask participating students to share whether or not they have noticed any differences in the process and ask them if or how they are using information from the applicable assessment.

  Ashland Elementary also may want to engage in a follow-up LCD cycle and include students and parents in the entire process to consider their user experience, but specifically to include them in the *illustrate* phase during which they help to shape what a successful solution might look like, the *ideate* phase during which they would have input on possible solutions, and the *proliferate* phase during which they would provide feedback on the success of the solutions and possible changes for the future.

In the case of Ashland Elementary, where the students are young, their involvement in the LCD cycle may be periodic and generated from questions about how they learn and what kind of support and feedback helps them in their learning. Their perspectives, though, are incredibly valuable because they have had very little time to become institutionalized in their thinking by historical and possibly outdated practices. It is also possible that Ashland Elementary might want to utilize older students who have moved out of the elementary grades to garner perspectives that include their elementary experience as well as their current reality. Jill mentioned that her own children have had a variety of experiences with teachers who have experimented with a skills-based approach to grading and assessment and would be in a good position to provide constructive feedback.

While this study was conducted at the elementary level, future studies could look at how this design model would work for potential change in grading and assessment at the secondary level. There are significant differences in what instruction looks like in an elementary building secondary, so the outcomes will likely look very different. However, the LCD process may be helpful in enacting change at this level as well.

If other schools or districts want to utilize the LCD model to approach a problem of practice like the one presented in this study related to grading and assessment, or any educational problem of practice, I would encourage them to utilize student voice and include students' participation from the very beginning. The act of liberating students and enabling them to be co-agents of change may be an incredibly powerful approach to making changes in education. Including students in the process will help empower students, but their inclusion will also challenge and expand the perspectives of the adult educators.

It is my hope that other educational organizations can use the principles of this study to engage in their own change, whether it be a change in grading and assessment practices, or any other aspect of the educational setting. While the nature and purpose of the organization will make a difference in how the process is approached, the idea that the human capital of all stakeholders can be used to enhance the social and decisional capital of the organization is key. The context in which the organization functions will dictate the details of the LCD team and the process timeline, but the co-construction of new solutions that are collaboratively built and tested can be applied in any setting.

Discussion

While I set out to study a problem of practice that was specific to much-needed change in the grading and assessment practices that are currently predominate in schools, I found myself being drawn to educational change in general. I am convinced now, more than ever, that we must re-imagine what school looks like and how we view learning. We continue to see young, curious, innovative, creative, and excited learners who become disenfranchised from learning for many reasons, of course, but being tethered to historical practices that have outlived their usefulness is one of them. In modern America, we talk a great deal about thinking outside the box. It's time to jump completely outside the box and dedicate ourselves to finding revolutionary ways to nurture children in what should be for them the exciting pursuit of learning.

The lack of a clear vision and back mapping, along with desperately holding onto the attitude of "we have always done it this way," undoubtedly contribute to the difficulties we have instituting meaningful change in education. When we engage in change, it is often in response to a concerning event or problems we have uncovered in our current structures. Rather than taking

109

the time to become very clear about what we want to see as the outcome and then back mapping and developing a plan to achieve that vision, we "react" to problems and try to fix them. Even in the act of trying to make change, we are too often eager to grasp onto the life ring of past practices without realizing that a life ring in this circumstance might be more anchor than life preserver. Sometimes when making a change, it's better to risk the swim and see if we can reach a new, better, shore.

It is my hope that the use of a model that is not only collaborative, but also intentional in establishing a clear vision for success will help to change how we engage in change. When a team starts with a vision of what they want a solution to look like before going right into problem-solving mode, it is more likely they will be able to develop a clear plan to achieve the vision. Similarly, if the process is seen as the construction of new thinking and if human capital is used effectively to promote a more collaborative social capital, I believe that teams will be less likely to fall into the pattern of embracing a status quo that no longer works for the greater good of all the students we are charged to teach.

REFERENCES

Abbott, J. (2015). Battling for the soul of education: Moving beyond school reform to educational transformation. *NAMTA Journal, 40*(3), 93-105.

Armstrong, T. (2006). *The best schools: How human development research should inform educational practice*. Alexandria, VA: ASCD.

Banks, J. (1998). The lives and values of researchers: Implications for educating citizens in a multicultural society. *Educational Researcher, 27*(7), 4-17.

Birk, C. & Allen, C. (2017). *Changing change: Using learner-centered design, from failed initiatives to a change process that connects, empowers, and actually works*. Seatle*, WA:* PBL Consulting.

Black, P., & Wiliam, D. (1998). Inside the box. Raising standards through classroom assessment. London: King's College London.

Black, P., & Wiliam, D. (2003). "In praise of educational research": Formative assessment. *British Educational Research Journal, 29*(5), 623-637. doi:10.1080/0141192032000133721

Bradley, D. F., & Calvin, M. B. (1998). Grading modified assignments: Equity or compromise? *Teaching Exceptional Children, 31*(2), 24-29.

Brodinsky, B. (1972). *Grading and reporting: Current trends in school policies & programs*. Arlington, VA: National School Public Relations Association.

Brookhart, S. M. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice*, *10*(1), 35-36.

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

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., . . . Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research, 86*(4), 803-848. doi: 10.3102/0034654316672069Caine, R. N., & Caine, G. (1991). *Making connections: teaching and the human brain*. Alexandria, VA: ASCD.

Brooks, J. G., & Brooks, M. G. (1999). *In search of understanding: The case for constructivist classrooms*. Alexandria, VA: ASCD.

Creswell, J. W (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: Sage publications.

Cox, K. B. (2011). Putting classroom grading on the table: A reform in progress. *American Secondary Education*, 67-87.

Cross, C. T. (2015). *The shaping of federal education policy over time. The progress of education reform. 16(2),* Education Commission of the States.

Curry, M. W. (2012). In pursuit of reciprocity: Researchers, teachers, and school reformers engaged in collaborative analysis of video records. *Theory Into Practice*, *51*(2), 91-98.

Darling-Hammond, L. (2010). *The flat world and education: How America's commitment to equity will determine our future*. New York, NY: Teachers College Press.

Davies, A. (2007).   Involving students in the classroom assessment process. In Reeves, D. B. (Ed.). (2009). *Ahead of the curve: The power of assessment to transform teaching and learning*. Solution Tree Press.

DeWitt, P. M. (2016). *Collaborative leadership: Six influences that matter most*. Corwin Press.

Disibio, R. A. (1971). Different approach to grading. *Education, 92*, 34-35.

Donohoo, J. (2016). *Collective efficacy: How educators' beliefs impact student learning*. Corwin Press.

Erkens, C., Schimmer, T., & Dimich Vagle, N. (2017). *Essential assessment: Six tenets for bringing hope, efficacy, and achievement to the classroom.* Bloomington, IN: Solution Tree Press.

Fullan, M. (2001). *Leading in a culture of change*. San Francisco: Jossey-Bass.

Fullan, M. (2016). *The new meaning of educational change*. Routledge.

Fullan, M. (2016a). *Indelible leadership: Always leave them learning*. Corwin Press.

Gardner, D. P., Larsen, Y. W., Baker, W.O., Campbell, A., Crosby, E. A., Foster C. A., Jr., et al. (1983). A Nation At Risk: The Imperative For Educational Reform. An Open Letter to the American People. A Report to the Nation and the Secretary of Education. Washington, DC: National Commission on Excellence in Education.

Glesne, C. (2016). *Becoming qualitative researchers*. Boston, MA: Pearson.

Guskey, T. R. (2009). *Practical solutions for serious problems in SBG*. Thousand Oaks, CA: Corwin Press.

Guskey, T. R. (2011). Five obstacles to grading reform. *Educational Leadership, 69*(3), 16.

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

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

Guskey, T. R., & Brookhart, S. M. (2019). *What We Know about Grading: What Works, what Doesn't, and What's Next*. ASCD.

Hancock, D. R., & Algozzine, B. (2016). *Doing case study research: A practical guide for beginning researchers*. Teachers College Press.

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. New York, NY: Teachers College Press.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. doi: 10.3102/003465430298487

Heifetz, R. A., & Linsky, M. (2002). A survival guide for leaders. *Harvard business review, 80*(6), 65-74.

Heifetz, R. A., & Linsky, M. (2004). When leadership spells danger. *Educational leadership, 61*(7), 33-37.

Heifetz, R. A., & Laurie, D. L. (2001). The work of leadership. *Harvard Business Review, 79*(11).

Hochbein, C., & Pollio, M. (2016). Making grades more meaningful. *Phi Delta Kappan, 98*(3), 49-54.

Hursh, D. (2007). Exacerbating inequality: the failed promise of the No Child Left Behind Act. *Race Ethnicity and Education, 10*(3), 295-308.

Johnson, B., & Christensen, L. (2017) *Educational research; Quantitative, qualitative, and mixed approaches* (6th edition). Thousand Oaks, CA: Sage.

Jung, L. A., & Guskey, T. R. (2009). *Grading and reporting in a standards-based environment: Implications for students with special needs.* Paper presented at the Annual Meeting of the American Educational Research Association in San Diego, CA. Retrieved from http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED509343

Jung, L. A., & Guskey, T. R. (2012). *Grading exceptional and struggling learners.* Thousand

    Oaks, CA: Corwin Press.

Keni, B. C. (2011). Putting classroom grading on the table: A reform in progress. *American*

    Secondary Education, (1), 67. Retrieved from

    http://www.jstor.org.libproxy.lib.ilstu.edu/stable/23100415

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

Lambert, L. (2002). *The constructivist leader*. Teachers College Press.

Lortie, D. C., & Clement, D. (1975). *Schoolteacher: A sociological study* (Vol. 21). Chicago:

    University of Chicago Press.

Marzano, R.J. (2003). *What works in schools: Translating research into action.* ASCD

Marzano, R.J. (2017). *A handbook for personalized competency-based education*. Bloomington,

    IN: Marzano Research.

Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and*

    implementation. John Wiley & Sons.

McTighe, J., & Curtis, G. (2019). *Leading modern learning: A blueprint for vision-driven*

    schools. ASCD.

Mintrop, R. (2016). *Design-based school improvement: A practical guide for education leaders*.

    Harvard Education Press.

Nottingham, J., & Nottingham, J. (2017). *Challenging learning through feedback : How to get*

    the type, tone and quality of feedback right every time. Thousand Oaks, CA: Corwin.

O'Connor, K. (2007). *A repair kit for grading: 15 fixes for broken grades.* Portland, OR:

    Educational Testing Service.

O'Connor, K (2007).   The last frontier: Tackling the grading dilemma. In Reeves, D. B. (Ed.). (2009). *Ahead of the curve: The power of assessment to transform teaching and learning*. Solution Tree Press.

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

Popham, W. J. (2011). *Transformative assessment in action: An inside look at applying the process*. Alexandria, VA: ASCD.

Reeves, D. B. (2004). The case against the zero. *Phi Delta Kappan, 86*(4), 324.

Reeves, D. B. (2011). Taking the grading conversation public. *Educational Leadership, 69*(3), 76-79. Retrieved from http://www.ascd.org/publications/educational-leadership/nov11/vol69/num03/abstract.aspx

Reeves, D., Jung, L. A., & O'Connor, K. (2017). Special topic: What's worth fighting against in grading? *Educational Leadership, 74*(8), 42-45. Retrieved from http://www.ascd.org/publications/educational-leadership/may17/vol74/num08/What's-Worth-Fighting-Against-in-Grading%C2%A2.aspx

Saldaña, J. (2016). The coding manual for qualitative researchers. Los Angeles: Sage.

Sappington, N. (2008). Don't ask teachers to teach to the test. *The Illinois School Board Journal,* 76(6), 33-47.

Schimmer, T. (2016). *Grading from the inside out.* Bloomington IN: Solution Tree Press.

Schneider, J., Feldman, J., & French, D. (2016). The best of both worlds. *Phi Delta Kappan, 98*(3), 60-67. doi:10.1177/0031721716677264

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

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4-14.

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.

Starch, D.A., & Elliott, E. C. (1912). Reliability of the grading of high-school work in english. *The School Review,* (7), 442. Retrieved from

http://www.jstor.org.libproxy.lib.ilstu.edu/stable/1076706

Tierney, R. D., Simon, M., & Charland, J. (2011). Being fair: Teachers' interpretations of principles for SBG. In *The Educational Forum 75*(3), pp. 210-227.

Trainor, A., & Bouchard, K. A. (2013). Exploring and developing reciprocity in research design. *International journal of qualitative studies in education*, *26*(8), 986-1003.

Vatterott, C. (2015). *Rethinking grading: Meaningful assessment for standards-based learning*. Alexandria, VA:ASCD.

Westerberg, T. (2016). *Charting a course to SBG: What to stop, what to start, and why it matters.* Alexandria, VA: ASCD

Wiliam, D., & Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal, 22*(5), 537-48. Retrieved from https://www-jstor-org.libproxy.lib.ilstu.edu/stable/1501668

Wormeli, R. (2006). Accountability: Teaching through assessment and feedback, not grading. *American Secondary Education, 34*(3), 14-27.

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

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

117

In what format do you prefer to receive your child's grade reports?

Checking Skyward (online or using the app)
- o A hard copy sent home
- o A digital copy e-mailed home

How often do you check Skyward to see your child's current grades?
- o At least weekly
- o Bi-weekly
- o Monthly
- o Once per quarter
- o I don't check Skyward

What information about your child's performance in school and their learning is most important to you in a grade report? Please rank from 1 (most important) to 7 (least important)
- o A score that is an average of all graded skills within a subject (i.e. for math- an overall grade of B)
- o Student performance on skills specific to a subject (i.e- for math- adding 2 digit numbers, measuring to the nearest inch etc…)
- o Information about student effort (i.e- your student is persistent)
- o Information about student respect at school (i.e- honest, caring, etc…)
- o Information about student involvement in learning (i.e.- engaged in class activities)
- o Ongoing improvement or growth over time
- o How your child compares to others

Which of the following would be most helpful to you in supporting your child in their learning? (Choose no more than three.)
- o Knowing my child's current academic performance on a specific skill
- o Knowing my child's grade in a subject
- o Knowing my child's behavior in the classroom
- o Knowing if my child gets along well with others
- o Knowing if my child is socially- and self-aware
- o Knowing how well m y child is able to focus in class

Learner Centered Design Model

- Supporting change by viewing and planning the change through the lens of the "learners"- those who will be most impacted by the change
- We want to ensure we have agreement within the staff about why we are considering a change in grading/assessment
- We need ALL voices and ideas

Considerations for exploring changes in Grading and Assessment?

- How AR fits in grading and assessment. Should it count for a grade?
- Academic consistency **within** grade levels (grades and reporting, rubrics, etc…)
- Grade reporting consistency **between** grade levels
- Assessments that cover multiple concepts/skills, but produce one grade or are reported as one concept/skill- need for **authentic** and accurate assessments

119

Considerations for exploring changes in Grading and Assessment?

- Developing policies for re-takes
- Using multiple probes/assessments before reporting student performance or progress- not using averages
- Noting modifications and accommodations, especially for non-IEP students
- Separating learning from behaviors

The ideal Grading and Assessment plan will....

- Be clearly articulated and communicated

- Be based on Viking Essentials that are expressed in understandable language

- Include the use of multiple probes to assess student learning and progress

The ideal Grading and Assessment plan will….

- Accurately describe student performance at a given point in time
- Identify specific areas of need for individual students
- Demonstrate student growth and progress over time

APPENDIX C: SUPPORTING DOCUMENT FOR ASSESSING BEHAVIOR

An Ashland Elementary Learner is...

Striving for Excellence	Evidence	Consistently Sometimes Rarely
On Task	Listening to instruction, following directions, attentive in class, working towards completion of classroom tasks, using class time effectively	
Prepared	Organized, prompt, attending regularly, ready with materials, completing required tasks	
Persistent	Putting forth effort, learning from mistakes, motivated, persevering through tasks the student finds challenging	
Comments:		

Respectful	Evidence	Consistently Sometimes Rarely
Mindful	Using self control, empathizing with others, socially and self aware, open to other ideas and opinions	
Collaborative	A respectful communicator with peers and adults, supportive of other learners, a team player, appropriate in interactions with staff	
Kind & Caring	Thoughtful, inclusive, compassionate	
Honest	Truthful, trustworthy, academically honesty, demonstrating integrity	
Comments:		

Involved in Learning	Evidence	Consistently Sometimes Rarely

122

Academically responsible	Demonstrating study skills, self-regulated, intrinsically motivated, committed to learning	
Opportunistic	Asking questions, seeking help when needed, advocating for themselves, using available resources, empowered to learn	
Engaged	Intellectually curious, involved in classroom activities through actions and discussion	
Comments:		

123

APPENDIX D: SAMPLE ESSENTIAL SKILLS DOCUMENT

2nd Grade Reading Curriculum Map for August/Early September– Essential Skills in Blue

Month/Mentor Text	Skills Assessed	Materials/Assessments
August/ Early September The Mixed up Chameleon Get Up and Go Henry and Mudge Days with Frog and Toad	Questioning RL 2.1 **Main Idea RL 2.10*** I can identify the main idea and supporting details to better understand a fictional text. **Narrative Elements (setting) RL 2.7*** I can use the words and pictures from a fictional text to identify the characters, setting and plot. **Compare and Contrast RL 2.9** I can compare and contrast elements from different stories to better understand a fictional text.	Jessica Tobin-RL 2.1 Use book to find Main Idea Jessica Tobin-RL 2.7 Jessica Tobin-RL 2.9 (use Frog and Toad stories)

* Skills Repeated