Illinois State University
ISU ReD: Research and eData

Theses and Dissertations

6-21-2023

Shifting from a System of Grades to a Culture of Learning: Moving toward a More Equitable Culture of Assessment

Jennifer Pamela McCoy
Illinois State University, jpmccoy18@gmail.com

Follow this and additional works at: https://ir.library.illinoisstate.edu/etd

Recommended Citation
McCoy, Jennifer Pamela, "Shifting from a System of Grades to a Culture of Learning: Moving toward a More Equitable Culture of Assessment" (2023). *Theses and Dissertations*. 1761.
https://ir.library.illinoisstate.edu/etd/1761

This Dissertation is brought to you for free and open access by ISU ReD: Research and eData. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of ISU ReD: Research and eData. For more information, please contact ISUReD@ilstu.edu.

SHIFTING FROM A SYSTEM OF GRADES TO A CULTURE OF LEARNING:

MOVING TOWARD A MORE EQUITABLE CULTURE OF ASSESSMENT

JENNIFER P. MCCOY

124 Pages

Upon identifying a problem of practice centering on students and staff alike focusing on points and grades instead of learning, a small team of teachers were assembled to participate in a cycle of inquiry. The purpose of this cycle of inquiry was to research and test the effectiveness of an intervention related to assessment. The team selected a student self-assessment in the form of an exit slip. Qualitative methods were used to code teacher responses while they implemented the exit slip with students. The team ultimately found that they were able to use the exit slip feedback from the students to inform their instruction. They also found value in the cycle of inquiry process.

KEYWORDS: Assessment; Formative Assessment; Student Self-Assessment; Exit Slips; Cycles of Inquiry; Problem of Practice; Adaptive Problem; Rural Poverty; Small School

SHIFTING FROM A SYSTEM OF GRADES TO A CULTURE OF LEARNING: MOVING TOWARD A MORE EQUITABLE CULTURE OF ASSESSMENT

JENNIFER P. MCCOY

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

2023

© 2023 Jennifer P. McCoy

SHIFTING FROM A SYSTEM OF GRADES TO A CULTURE OF LEARNING:

MOVING TOWARD A MORE EQUITABLE CULTURE OF ASSESSMENT

JENNIFER P. MCCOY

COMMITTEE MEMBERS:

Elizabeth Lugg, Chair

Rudo Tsemunhu

Ryan Smith

Diane Wolf

ACKNOWLEDGMENTS

On December 6, 2018, I had just completed the final projects of my first semester on this journey to earn my doctorate. I submitted my assignments at 11:30pm and breathed a sigh of accomplishment: One semester down. Four hours later, I awoke with a terrible headache and the inability to move my left side. I was having a massive stroke. Thanks to my husband, who acted so quickly and to all of the medical personnel, I walked out of the hospital six days later.

The day I awoke having a stroke also happened to be my husband's birthday; unfortunately for him, he now has to share his birthday with my "strokaversary," but like so many of these hurdles that we have crossed together, he takes it in stride. So, first, foremost, and forever, I am grateful to Scott McCoy for saving my life—*literally*—and figuratively as we traveled this journey together. While I know that my coursework caused many inconveniences for him, he never wavered in his encouragement, support, or belief that I would finish.

My children veritably grew from young adults and teens to fully functional adults during this time. One graduated from high school and two had children of their own, and I am thankful to Joshua, Hannah, and Elizabeth for tolerating the many times I was distracted by work and school. I am excited that Jocelyn and Henry get to see their Mom Mom reach this pinnacle of her education. Thank you, too, to my father-in-law and to my mom and dad for their unfailing support and always giving me a sympathetic ear and soft place to land when I needed a break. From a very young age, my mother and father instilled in me a passion for learning and reminded me of the importance of earning a college degree, and they continued to encourage me through all of my ensuing degrees. Through this final educational accomplishment, I hope that my children and grandchildren and all of my student-children know that they, too, can achieve their dreams.

I would be remiss if I did not acknowledge my school-work family. Melissa Meints and Melissa King were cheerleaders for every dissertation milestone, and Paul Deters, Julie Strating, and Isaac Steidinger will always be more than mere coworkers to me. I am grateful that I get to work with you, and I appreciate that you do what's best for our students every day.

Thank you to my dissertation committee, as well: Dr. Lugg, words cannot express my gratitude for agreeing to accompany me on the final leg of this marathon; Dr. Smith, I appreciated your calm encouragement, knowledge, and positivity. Dr. Wolf, your expertise helped me refine my work, and Dr. Tsemunhu, thank you for joining us in the eleventh hour which allowed me to finish.

I have been blessed to complete this journey—with a husband who has sustained me, a family who supported me, and an extended school family that cheered me through every step, semester, and paper. No achievement is ever accomplished alone, and I thank you for helping me reach this one.

-JPMc

CONTENTS

Page

ACKNOWLEDGMENTS i

TABLES vi

CHAPTER I: A PROBLEM OF PRACTICE 1

CHAPTER II: INTUITIVE THEORIES OF ACTION 5

Traditional Assessment Practices 6

Staff In-service 7

Deficit Mindsets 9

The Team 10

Tacit Assumptions 12

Culture 16

CHAPTER III: UNDERSTANDING THE PROBLEM 20

Vulnerable Students 21

Deficit Thinking 22

Positives of Small Schools 25

Bixby's Culture 25

The Staff 27

Grades and Grading 30

Purpose of Grades 30

Grading Practices 32

Principal Leadership 43

Professional Development 45

Empowering Staff 46

Next Generation Assessments 46

Assessment Literacy 47

Shifting Mindsets 48

Theoretical Framework 49

Meaning-making 49

CHAPTER IV: DESIGNING AN INTERVENTION 54

Exploring Interventions 55

The Intervention: Student Exit Slips 58

CHAPTER V: METHODOLOGY 61

Teacher Use of Student Exit Slips 64

Teacher Reflections 64

Meetings 64

Winnowing the Data 66

Trustworthiness and Ethical Considerations 67

CHAPTER VI: IMPLEMENTING THE INTERVENTION 70

Team Meetings 71

Trouble-shooting Problems 71

Offer Support 72

Adapting the Intervention 73

Intellectual Leadership 73

Teacher Responses to their Reflections 74

CHAPTER VII: EVALUATING THE INTERVENTION 76

Themes 76

Honesty 80

Quality 82

When to give the Assessment 83

Timing 84

Usable Feedback 85

Individual Check-ins 87

Confidence 88

Adjusting Instruction 89

Final Surprises 94

A Book Study without the Book 96

"They didn't Teach Us that in Undergrad" 98

CHAPTER VIII: IMPLICATIONS AND NEXT STEPS 100

The Next Cycle 100

Funds of Knowledge 101

Assessment 101

Bixby's Assessment Future 106

REFERENCES 109

APPENDIX A: STAFF IN-SERVICE ON ASSESSMENT 116

APPENDIX B: THE EXIT SLIP/STUDENT SELF-ASSESSMENT 123

APPENDIX C: TEACHER REFLECTION 124

TABLES

Page

Table 1: Team Demographics 11

Table 2: Summary of Teacher Responses 78

CHAPTER I: A PROBLEM OF PRACTICE

> *"The teacher's job is not to transmit knowledge, nor to facilitate learning. It is to engineer effective learning environments for the students. The key features of effective learning environments are that they create student engagement and allow teachers, learners, and their peers to ensure that the learning is proceeding in the intended direction. The only way we can do this is through assessment. That is why assessment is, indeed, the bridge between teaching and learning"* (Wiliam, 2011, p. 50).

The school district in which I work enjoys a relatively high level of academic success. Our high school was named a US News and World Report "Best"; two of the three schools within the district regularly earned "exemplary" state status (citations withheld to protect the school district's identity). Our graduation rate is high, and our college remediation rate is low. Generally speaking, families feel confident sending their children to our building and believe that their students will emerge from it with an excellent education. We are a "good" district. However, when I became principal in 2019, what I observed in the classroom did not match up with my expectations of a high-functioning, rigorous district. I saw traditional instructional practices (i.e., teacher-centered discussions, desks in rows, and few students actively engaging in learning activities), yet teachers openly remarked that their instructional practices were "cutting edge." On one hand, we engaged in traditional practices, while on the other hand, the state designated us as "exemplary." Just two months before celebrating the end of my first year in the district, the COVID-19 pandemic shut down every school in the state of Illinois. My "hunch" that a deeper problem existed was put on hold.

The Bixby (a pseudonym) school district is a great place to work, and our district benefits

from an encouraging, involved community. It has a small-town atmosphere and a stable economy, where everyone knows everyone else, and everyone is proud of the town and loves our school. Often described as Mayberry-esque by visitors, townspeople frequently refer to it as the little town on the prairie. Despite six-foot distancing and one-way directions in the hallways, we were also fortunate to have community support in returning to face-to-face instruction in the fall of 2020. While our students had not been inside a classroom for five months, the adjustment was not nearly as difficult as we had anticipated, and within two weeks our students remarked that school felt relatively "normal." What was not normal were student grades. As we approached the end of the first quarter, student failures were abnormally high: more than 60% of the students had at least one "F," and many of those students had multiple failures. Even more were approaching failure, by receiving D's in one (and usually more than one) class. Collectively, we began to realize that perhaps our "best" high school was not as good as we thought.

While I was alarmed by the D/F rate, we were expecting some degree of learning loss due to pandemic. Discussions with teachers about student grades revealed that they were giving extra assignments to make up for perceived loss of instructional time. Timelines and due dates were inflexible, and students who failed to meet those deadlines received zeros. Not allowing retakes was viewed as adhering to "our" high academic standards. Quarantined students who fell behind—and they often did—had little hope of learning the material let alone completing all of the late work. In staff meetings, teachers talked about missing assignments, "lazy" pupils, and how students completed work based solely upon the points they could earn just to achieve a certain grade; teachers shrugged that these approaches were standard practice. I pondered whether these practices were just now *becoming* part of our culture or whether they had *always* been a part of it. Regardless, what was becoming evident was that the pandemic was not the

issue. Our grading practices and tacit assumptions about learning were: We were a system of grades, not a culture of learning. Thus, our problem was clear: Staff and students alike focused on grades and points instead of actual *learning*.

Bixby High School's grading problems clearly fall within the criteria for a "problem of practice": These types of problems are "urgent," "actionable," "feasible," "strategic," "tied to a specific practice," and "forward looking" (Mintrop, 2018, p. 30). A problem of practice is also entrenched within the organization (Mintrop, 2018). Our teachers clung to our assessment procedures because they felt that doing so maintained "rigorous standards." Unfortunately, the populations most often negatively impacted by poor assessment practices are the most vulnerable students (Valencia, 2010) within a school's population (Dudley-Marling & Dudley-Marling, 2015). As school districts across the world emerged from the COVID-19 Pandemic, educators and policy makers alike asserted that this time was a chance to change the way we "school" children. Thus, the topic of how we assess our students was a timely and important one in which many thought would be an opportunity to increase learning and equity among our most vulnerable students (Asadullah et al., 2023; Vegas & Winthrop, 2020). In order to address the problem of practice at Bixby, we would need to work within our community of practice to improve not only how we educated and assessed our students but change our perceptions of them.

Underlying our plans were efforts toward working to develop how we approached improving our practice using a science of improvement. As Bryk (2009) explains, "to create usable knowledge about practice" (p. 599) and focus on developing practices that impact day-to-day instruction and student learning. These approaches are beyond the typical professional development activities in which teachers normally participate. In our case, a core group of

teachers became co-designers of how we approached the problem; they were invited to “think like a designer,” to analyze the problem and develop an actionable approach to test a way to improve our assessment practices (Henriksen & Richardson, 2017).

CHAPTER II: INTUITIVE THEORIES OF ACTION

While I had early indications that something was amiss in our instructional practices, the problem centering on student grades came into focus during the immediate weeks after our staff and students had returned for the 2020-2021 school year, while many other districts continued distance learning. With the suspicion that something might be awry with our assessment practices, I remained alert for additional signs of this problem throughout the school year. The steps I took to identify the problem closely followed Mintrop's (2018) process for designing a school improvement process. Specifically, after noting the grading practices and post-COVID issues, I began to frame that a problem existed within how our staff assessed students as well as how they viewed our students who were classified as being of low socio-economic status. While we as a staff returned to face-to-face instruction, I carefully observed my staff's assessment practices and attitudes toward this vulnerable population.

By the end of the school year, the problem had begun to clarify in shape. I had engaged in formal and informal conversations with staff, with these informal conversations about assessment—typically retakes and zeroes—taking place when individual students were at risk for not passing a class. The discussions were one-on-one with staff members or included the student and the student's family. During performance evaluations, we had more formal discourse as certified staff described their assessment habits.

All of these conversations were confirmed when I observed traditional assessment practices in most classrooms—activities like unit tests in the form of extended response and comprehensive finals at the end of the semester. To further shape this picture, I created an end-of-year in-service activity for staff that centered on assessment. The activity revealed that few staff had set policies on retakes and that many viewed retakes as reducing rigor, and a

concerning number had not considered the notion of mastery or were aware of flaws with averaged grading systems. This information pointed to deeply entrenched habits that existed simply because “that’s what we’ve always done”—and not necessarily engaging in activities because they were best for student learning. Other beliefs that arose were centered on not allowing retakes because staff believed that it would reduce rigor and that if students only “cared enough,” they would have tried harder the first time. Unfortunately, these mindsets take away from what should be our central goal: learning, and negatively impacted all students but students of low-income the greatest. Uncovering all of these underlying beliefs was important for gathering a clear picture of our staff’s strengths and areas for improvement. It helped determine which first steps to take them on as we traveled this journey into assessment.

Traditional Assessment Practices

Staff typically used traditional assessment practices in the form of unit tests and essays. The exception to these practices were some of our elective courses; for example, students in our art classes created pieces based upon criteria outlined in rubrics. Throughout the unit, the teacher had students self-assess their progress and reflect on their work when compared to the rubric’s criteria. When their products were finished, they engaged in a reflective process in which they assessed their own and others’ work and verbally described the outcome of their projects. Other electives, like our agriculture and business/technology classes focused on projects, using rubrics as well to assess the final products; however, neither teacher formally integrated a continuous process of self-assessment like the art instructor.

While the majority of staff (especially core teachers) relied upon tests and essays (usually exclusive to English classes), all made efforts to integrate a variety of question types within the assessments themselves. For example, the social studies teacher did not limit his questions to the

multiple choice variety and instead, used multiple choice, short answer, and essay questions. The English teachers rarely placed multiple choice questions on tests, using short answer and essay prompts instead. While this information may see like a minor detail, I feel it points to the teachers' desire to clearly assess the students' understanding of what they had taught. If they were unmotivated or lazy, they simply could have generated multiple choice question tests from their curriculum software. Instead, they chose to write, grade, and give feedback on tests that used a variety of question types.

Staff In-service

I frequently touched on the topics of vulnerable student populations and effective assessment practices in our Leadership Team meetings and even had the entire staff engage in an activity to reflect upon their grading practices during our end-of-year in-service. For the activity, teachers were separated into teams and completed a scavenger hunt (see Appendix A). Clues brought them to specific locations in the building, and at those locations, they worked through examples and questions relating to assessment. The tasks they completed centered on defining grades (an "A," a "B," a "C," etc.), looking at averaged grades, the impact of zeroes in an averaged point system, and evaluating mastery for their content area. The purpose of these prompts was to have staff think critically about what and how they used assessment. At the end of the activity, we engaged in a whole-group discussion about their results.

The feedback from informal conversations and specific feedback from staff on their assessment practices had brought to light a variety of tacit assumptions about grading, rigor, motivation, and more to the surface. These mindsets helped me frame the problem of practice by providing information on how staff's level of understanding of "mastery," as well as their ability to recognize issues with averaged grades. The final prompt about mastery gave insights into the

types of assessment practices they engaged in and whether they had a clear understanding of what that looked like for their content area; this prompt, too, led for some teachers to explain their practice of retakes. For example, staff typically engaged in "one-and-done" assessment practices in which teachers gave a single summative assessment at the end of their instructional units.

Formative Assessment

With the exception of our physical education teachers and the art teacher (because she had taught nearly all of the students prior to high school and had an understanding of their abilities), none of the teachers who submitted their responses to the in-service activity reported pre-testing students *before* instructing them or using regular pre-assessments to *inform* their instruction. Only one teacher who submitted her in-service work, outlined a formal procedure for retakes. In this case, she explained that students could retake a test but the highest they could receive on it was a B minus. This process also required the teacher to create a completely new test as well, and for the student to come in before or after their normal school day.

Retakes and Rigor

Informal conversations continued to reveal that some teachers struggled with choosing the best approaches for assessing students—from retake procedures to test design. Instructors remarked that they struggled with whether retakes were appropriate given that this could be perceived as being "too easy" on our students and lowering our academic rigor. Others were unsure of the best methods for retakes—many felt that they had to create completely new tests, which they saw as burdensome. Permeating these conversations was the notion that if students "cared enough," they would have been successful during the first assessment, and they often pointed to our low-income students as requesting or needing to retake tests most often.

Students "Not Caring"

In more formal conversations, similar attitudes surfaced as well. I am the primary evaluator for all of the secondary teachers at Bixby, and when we discussed the assessment piece of their evaluation, many teachers described a desire to move away from relying on traditional multiple choice or short answer tests to give students grades. They also expressed concern about approaches to retakes—again, reiterating that on one hand they knew that retakes were best practice, but the prevailing attitude at Bixby was that it would somehow lower our rigor. Layering upon this viewpoint, many teachers pointed to our low-income students as "not caring" about their grades or taking advantage of retakes when instructors graciously allowed it to happen. One particular teacher remarked, "They just don't care about their grades, and their parents don't care either. I can't make them *want* to get good grades."

Deficit Mindsets

As a result of conversations like this one, I began "making intuitive theories of action explicit" (Mintrop, 2018, p. 17). The staff had developed a basic cause-and-effect, or heuristic, in which they linked poor grades to student motivation, and more concerning, poor grades of our vulnerable students to a lack of support from their families (Mintrop, 2018). However, what I noted throughout my observations was a deficit attitude as well as lack of educators' connecting their practice to student learning. Instead, teachers had linked a lack of student motivation to their socio-economic status, and both students and staff had focused on work completion and grades instead of what students were actually learning.

The following year, I "consulted the professional knowledge base" (Mintrop, 2018, p. 17) and sought out a team of staff members to help refine and address the potential problem. Research shows the validity of allowing retakes for full credit and that this practice was

especially important for vulnerable populations of students (Jensen & Snider, 2013; Wormeli, 2006; Wormeli, 2011). In discussing low-income populations with our staff, they often referred to what the students lacked in understanding, but research cites the importance of valuing the experience and knowledge that they bring to the classroom (Dudley-Marling & Dudley-Marling, 2015; Valencia, 2010). Research also shows the value of formative and self-assessments—something which many of our staff did not practice (Hattie, 2012; Hattie & Yates, 2014; Jensen & Snider, 2013; McTighe & Willis, 2019; Nolen, 2011; Percell, 2017; Popham, 2001; Popham, 2018; Wiliam, 2011; Wormeli, 2011). In navigating our next steps, I hoped to enlist a team of teachers who would be willing to determine among the variety of improvements we could make to our assessment practice as well as adjust their own practice to test out an intervention.

The Team

Our "team" started with two teachers who had expressed interest in improving the way they assessed students. The first teacher I invited was not only new to the district but very new to teaching as a whole. "Charlie" (a pseudonym) completed his entire student teaching experience during the pandemic and had never taught students face-to-face. The second teacher was in her fourth year in the district when we started this process and had worked previously in a significantly larger school. In all, "Kate" (a pseudonym) had eight years of experience, but she consistently challenged herself to try new approaches and perfect her practice. I elected to begin this work with these two staff members because they had separately expressed a desire to improve how they evaluated student work.

Table 1

Team Demographics

Team Demographics				
Total: 4				
Gender	**Female**	**Male**		
Team Members	3	1		
Years Experience in Education	**1-4**	**5-9**	**20**	
Team Members	1	1	2	
Years Employed at Bixby	**1-4**	**5-9**	**20**	
Team Members	2	1	1	
Content Area	**English**	**Math**	**Science**	**Social Studies**
Team Members	1	0.5	1.5	1

Note. Table 1 consists of the team's demographics, including experience in education as well as their length of employment at Bixby when the team was assembled. The team consists of the teachers and the administrator leading the study and includes the content area that the administrator taught.

I elected to use a snowball sampling method by allowing the participants to select the final member of our team. In our early meetings, I sought out their feedback on who they felt would be a good fourth member. The rationale for having them select the fourth member was two-fold: First, I wanted to engage them in the decision-making process and co-design the cycle of inquiry (Mintrop, 2018), and second, I wanted to follow Shields' (2020) work on transformative leadership by moving beyond superficial change to "accept the need for deep and equitable change and 1) reflect on our own beliefs; 2) examine our school data; and 3) work to understand our community context" (p. 17-30). Having the two primary team members identify the final participant fit into my goal of valuing them as equal participants in the work we were about to complete.

The initial members of the team expressed interest in assessment via previous formal and

informal conversations. Kate and I had discussed better assessment methods during her summative evaluation when she stated that she was frustrated with how to accurately gauge student understanding of her content as well as reduce the burden of retakes. Charlie, on the other hand, knew of my background as an English teacher and wanted more ideas for "alternative" assessments. At the end of his first year of teaching, he had begun to tire of giving the same types of tests and wanted to give more opportunities for students to complete projects or write essays. Kate suggested Anne as our fourth member because the two had had previous conversations about ways to "grade better." Both Kate and Charlie felt comfortable with Anne, so adding her to the group would not disrupt the trust we had already established.

Tacit Assumptions

Discussions with the initial two members of the team led to developing more tacit assumptions, and they began to turn them into explicit theories (Mintrop, 2018). Specifically, they both felt that the core of our "grading problems" centered on teachers not knowing or using effective grading practices. Charlie started exploring the topic of traditional grading practices because he was unsure of how to use or develop rubrics. By the phrase "traditional" or ineffective grading practices that our staff used, we referred to grading on a 100-point averaged scale; deducting points for late work; and issuing zeroes for incomplete or missing assignments. The teachers mentioned at various times that they felt the need to "grade everything," so this introduced another aspect of traditional approaches: they cited assigning points to every assignment issued (grading practice) as well as not allowing retakes; however, researchers advocate for fewer, smaller assessment (Popham, 2001) and focusing mastery instead of casting a wide net with graded assignments, quizzes, and unit tests (O'Connor, 2011; Wormeli, 2006; Wormeli, 2011).

The team debated the need for additional professional development on grading, believing that staff simply did not know about their options when a student was failing their class. However, staff are able to cite other means to demonstrate mastery when students struggle in their classes. The team also discussed standards-based grading (SBG) as a solution to an averaged system, pointing to area school districts who had shifted to this method; however, the notion that teachers lacked background knowledge in strategies for assessing students superseded their belief that we should move to SBG. The greater problem, though, was cultural, a gaming for points instead of a focus on mastery or learning as well as frivolous assignments and the teachers' impulse to grade everything.

Grading

While some of the issues that the team cited were attributable to problems within our system (e.g., an averaged grading policy), others were tied to our culture (e.g., late work policies and zeroes for missing assignments). Our school issues grades based upon a 100-point, averaged system, and the team discussed some area districts' move to standards based grading. They questioned whether we would (or should) make this move as well but rejected standards-based grading very quickly because they felt it would be too complicated to implement at the high school level, referring to college admissions and athletic eligibility requirements (Swan et al., 2012). The team concluded that staff need additional professional development on grading. This response fell in line with the heuristic that a lack of practice should automatically result in more professional development.

Strategies for Assessment

While student grades and a lack of variety of how staff assessed students showed that teachers may not have been using effective grading techniques, conversations with our

instructors as a whole regarding their knowledge of assessment indicated otherwise. The two team members were surprised to hear that when teachers have a student who is failing a class, staff can readily identify specific strategies to improve their assessment practices (e.g., allow retakes for full credit, not grade practice, or focus on evidence of mastery). For example, teachers who have students failing their class will often extend deadlines or allow retakes for full credit. I have met with teachers of students at risk of failing a core class or not achieving the credits they need to pass, and we have discussed what the teacher *needs* to assess whether the student has mastered the content to earn credit for the course (Fisher et al., 2011; Guskey, 2010; Guskey & Anderman, 2013; Jensen & Snider, 2013; O'Connor, 2011; Popham, 2001; Wormeli, 2006). These conversations typically result in the instructor "pruning" a variety of assignments from the student's grade, usually ending with their grade consisting of summative unit assessments. When I revealed this to the team, they concluded that perhaps these practices are reserved for when students struggle and are not part of their daily assessment habits. Rather, staff appeared to know best practices but did not use them unless pressed to do so.

Mastery

We explored why unnecessary assignments—assignments not linked to demonstrating mastery—were included in a student's grade, and it became evident to us what was missing: talk about *learning* as opposed to talk about "points." The team realized that when staff talked about student failures, they frequently mentioned lack of points and how students "game" for grades by selecting specific assessments that help them pass. For example, a teacher and I met to discuss a student's failure in a class that was required for graduation. The teacher explained how the student had not turned in her homework and had failed to come in to retake tests. The student countered that she did not do the homework because it was time consuming and wanted to focus

her attention of getting just enough points on tests to pass the class. In contrast, I spoke to one of our highest achieving students about his learning. After receiving a puzzled look, he described how he had received high scores on tests and always attempted any extra-credit the teachers offered.

Frivolous Assignments

Not all of the frivolous assignments can be placed upon the teachers' shoulders, however. Several of the "seasoned teachers" informed me that previous principals had required at least two grades in the grade book per week. This requirement was intended to help with eligibility for athletics, but some of the staff members have continue the practice of recording at least two assignments each week, with several believing that it is still a requirement. Unfortunately, the results are "fluffy" assignments in students' grades that do not communicate student understanding of the content. Another result of this practice is that students who do not see the value in the activities (e.g., non-athletic students or students not motivated by grades), merely fall behind by receiving zeroes on simple assignment, ultimately resulting in miscommunicating their understanding of the standards.

"Grading Everything"

Building on this problem, one member mentioned—and the other member agreed—that staff felt the need to "grade everything," even small homework activities because they feared that without a grade, students would not complete what staff believed to be important learning activities. However, when teachers permit this type of gaming for points, they send the message that an arbitrary grade is more important than learning (Guskey, 2020a; O'Connor, 2011; O'Connor, 2018; O'Connor & Wormeli, 2011). As Kate pointed out, though, when students get behind or struggle, these "small-yield" activities are often the first to be left incomplete in

exchange for assessments that will guarantee a pass in the course.

This belief speaks to tacit assumptions and traditional beliefs about motivation—that grades themselves are both motivation and consequences for poor (or in some cases, positive) behavior and points to the problem that some of the assignments are perceived (by both the teacher and student) as busy work. And if some teachers grade everything because they think it is the only way to get students to complete the tasks, but students selectively choose which activities to complete based upon the points offered, then *both* sides are "gaming" the system. Adding to the issue, Wiliam (2011) argues that "grading practices prevalent in most U.S. middle schools and high schools are actually lowering achievement. Worse, typical grading practices don't even do one thing they are supposed to do, which is tell us what the students know" (p. 122). In the end, neither side is focusing on learning.

Comments from our team meetings supported that they, too, had fallen victim to this mentality. As the team discussed the issue of points and grades, Charlie asked, "But if I don't put a grade on an assignment, how do I get them to do it?" We recognized that assessment was a necessary component of instruction—and that when done well, it could be a valuable one—but we began to realize that grading every small activity was not only exhausting for the teachers but was a warning sign that something was amiss within our culture. Students were not taking responsibility for their own learning, and neither teachers nor students were focused on learning. Students function within the same culture as our teachers, and it was clear that they, too, would take time to change their views as well.

Culture

The team continued to explore approaches to solving our grading concerns and landed upon what they felt could be a potential solution. They did exactly what Mintrop (2018)

cautions against: jumping to a solution while still early in the process. They initially landed upon standards-based grading, vertical alignment, and motivation as starting points for professional development or district-wide policies.

Top-down Policies

In this case, they asked if standardizing some of our practices could resolve issues for teachers and students alike. Specifically, they cited a nearby district that had shifted to standards-based grading or another neighboring school that had created top-down policies for retakes or no zeros. Kate had worked in a prior district and explained her experience with top-down edicts on grading, describing how her department had locked themselves away in a classroom and spent hours working on vertical alignment and grinding out late work policies.

Alignment

Vertical alignment consists of teachers within the same department working together to develop outlines of topics they agree to cover during each semester as well as how they should sequence the topics as students progress through coursework. The need to communicate curriculum alignment is low among our staff because teachers at Bixby rely on the fact that they will likely have students multiple times as high school students. At the very minimum, most teachers have students for a full year as we have sequenced courses so that most students do not change teachers at the semester. This fact allows us to help students "catch up" when outside influences have impacted their learning sequences.

Late Work

Another top-down policy that the team explored was adopting a school-wide late work policy. Kate had described her work in another district in which she and her department created a departmental late work policy in which they all had to agree and implement. Kate and Charlie

posed that it would be "so easy" to simply implement a school-wide rule on late work: students could not argue with it or play teachers against each other (e.g., "they don't do this in other classes"). One department might deduct 25% for the first day an assignment is late with an additional 10% for every ensuing day; late work would not be accepted after two week

Unlike bigger districts, however, Bixby is not organized into hierarchical departments per se. True, we have departments, but frequently they consist of only one teacher. As a result, staff have a great deal of autonomy, so within that autonomy, we have a wide variety of late work and assessment practices. While most teachers do accept late work in some fashion, some reduce the grade by 50% while others have no penalty. Vertical alignment, on the other hand, is much easier at Bixby since nearly teachers have students for more than one year. The team agreed that while curriculum alignment and late work policies were valuable to explore, forcing teachers to engage in them would more likely impact our culture in a negative way, resulting in just another focus on "grades." Instead, changing teachers' and students' mindset toward one of learning would be a better, sustained direction for our overall school culture (Muñoz & Guskey, 2015; Nolen, 2011; Nordengren & Guskey, 2020; O'Connor et al., 2018; O'Connor & Wormeli, 2011; Wormeli, 2006).

Motivation

While the team reported that teachers often discuss grades on the perimeter of student motivation, I asked whether motivation may be a teacher issue as well. They initially felt that staff may not be driven to change their assessment practices for a variety of reasons. Citing staff motivation, members felt that seasoned teachers may not want to try "effective" grading approaches because they were too close to retirement to implement something new. For others, they claimed that timing could also be an issue by pointing to the fact that high school staff have

one preparation period. On the other hand, newer staff may feel insecure (especially if they had not achieved tenure) or if staff were, in their opinion, at a bad time in their evaluation cycle to try out unfamiliar assessment practices.

I questioned whether our culture supported changing assessment practices, and the team was intrigued. Previous administrators had not encouraged discussions like these; they had little time to after all, since the district recently emerged from poor economic times. I was the first principal in a long time (if ever) to truly work on our culture of instruction and could not help but wonder if I had missed a key opportunity to empower staff to try something new. I also could not overlook the risk teachers would perceive in giving students more autonomy over assessments. By giving students more control, they may believe that they are surrendering their authority instead of empowering students in their own learning. And I did not want to add to the teachers' vulnerability by implementing top-down policies on grading, curriculum alignment, or deductions for late work. Instead, these factors would have to be tempered by arming the teachers with knowledge to make informed decisions on when to release responsibility to students and when to retain that authority as the instructors.

CHAPTER III: UNDERSTANDING THE PROBLEM

Our team had pondered a number of ideas centering on assessment, so it was important to root this work within research. As a result, how we approached our problem of practice was influenced by literature relating to effective (and ineffective) grading practices such as exploring the purpose of grades and grading practice (homework); averaging grades and including zeroes within an averaged grade; including behavior within an academic grade; and encouraging appropriate risk-taking through equitable retake policies. We consulted literature on low-income populations because these vulnerable students represent the greatest diversity within our district. We learned through this process that my leadership as principal was also an important topic to support with research. Mintrop (2018) points out: "Leaders need to acquire knowledge and competence in the domains of problem solving, innovation, design, and research methods, but also in domains related to organizational learning and change, at the individual, group, and systemic levels" (p. 91). As we developed an understanding of our problem, it became clear that culture was an important piece of this issue. Therefore, since we were educators in a small, rural district, it was also important to acknowledge that this fact greatly influenced our culture, too.

In the end, the review of literature helped refine the problem of practice, and assisted the team in determining the direction for our cycle of inquiry. Ultimately, it aided us in understanding the symptoms and causes of our problem as well as provide a firm foundation for the team's decisions for moving forward with this work. In order to understand and shape the problem, we reviewed literature that tied to our culture, specifically exploring vulnerable students and deficit thinking. We also elected to root our problem in research on grades and grading in general—from harmful practices to beneficial ones. Throughout this chapter, we worked through the negatives and build to the positives so that by the end, we had worked up to

professional development for our staff and the positives of a small school.

Vulnerable Students

The team members had their own realizations as well when we discussed staff attitudes toward our most vulnerable students. I prefer to use the term "vulnerable" to describe our students of low socio-economic status (SES) and feel the phrase "low SES" lends to a deficit mindset towards this population (Valencia, 2010). Our district is not particularly diverse in the sense that most of our students are identified as white, but the number of families qualifying for free or reduced lunch is growing. And the notion of struggling financially is significantly different in a rural community with no access to public transportation and limited access to other resources. In fact, it is not unusual to hear of families living in a home with no running water or living conditions so dilapidated that they can see the ground through their flooring.

Staff had previously received some professional development on trauma-informed instruction and working with vulnerable populations. Despite this training, we agreed that our staff sometimes view students who struggle in school and live in vulnerable conditions as lazy, apathetic, or simply coming from families who do not care. Both staff members on the team guiltily admitted to viewing students as lazy, but they also often felt helpless and lacking methods to motivate disinterested students. By holding the view that students or their families do not care, teachers exempt their practice from impacting student learning, choosing instead to place all of the responsibility on the students (Dudley-Marling & Dudley-Marling, 2015; Valencia, 2010). Thus, ineffective assessment practices have the potential to put these vulnerable students at a greater disadvantage.

Deficit Thinking

Deficit thinking is the practice of blaming groups of students' poor academic achievement on their cultural background or low socio-economic status. Essentially, a "student who fails in school does so because of his/her internal deficits or deficiencies. Such deficiencies manifest, adherents allege, in limited intellectual abilities, linguistic shortcomings, lack of motivation to learn, and immoral behavior" (Valencia, 2010, p. 6-7). Instead of seeing the strengths students of diverse backgrounds bring to the classroom, they are labeled as "at-risk" and are often evaluated more harshly than their peers (Dudley-Marling & Dudley-Marling, 2015; Valencia, 2010). Teachers often view these vulnerable students as unmotivated and fail to separate their own practice from student achievement (Valencia, 2010).

Unfortunately, approaches to dealing with and attitudes toward vulnerable students are often wrought with deficit thinking. Researchers found that teachers often assess "lower-track" students more harshly and perceive that they have not learned concepts as well as their "higher-tracked" counterparts (Dudley-Marling & Dudley-Marling, 2015). Because of local demographic projections, our district understands that our population of vulnerable students will likely continue to grow in the coming months and years; as a result, it will be critical to address underlying assumptions and attitudes about our vulnerable students to improve their learning and educational experience.

On the other hand, effective practices, like formative assessments, for example, have the potential to increase student motivation: "Teachers who use formative assessments get more effort out of students. Such assessments tell the teacher exactly what needs to happen next for a given student to succeed" (Jensen & Snider, 2013, p. 77), and while this quote generally centers on using assessment to inform instruction, it also leans to the idea that assessment practices are

linked to motivation. When teachers align their assessments to appropriate student goals and engage in frequent, student-driven tests, they have a greater chance to achieve "flow" (becoming absorbed in the learning process) and motivation is a consequence of both flow and achievement (Wiliam, 2011, p. 149). In essence, both teacher and student shift their focus toward an ongoing dialogue centering on learning and success.

Working through complex processes like this adaptive problem not only required transformative leadership (Shields, 2020) but also a shift in the school culture to address deeply entrenched beliefs (Heifetz et al., 2009; Hill, 2011; Wormeli, 2006). Our problem of practice was not simply about hosting professional development for teachers to use effective assessment practices; it was how to shift mindsets on assessment of staff and students alike to ensure equity for all students, to see the value that all students bring to the learning environment.

Rural Poverty

A key component of our change was also directed at our deficit thinking toward our vulnerable populations (Dudley-Marling & Dudley-Marling, 2015; Valencia, 2010). Unfortunately, issues of rural poverty and education are not well represented in education research; in fact, several researchers (Books, 1997; Chandler, 2014; Irvin et al., 2011) all claim that rural poverty is largely ignored in scholarly work, and they posited that the media has perpetuated a myth that poverty is an urban issue (Books, 1997; Chandler, 2014; Malkus, 2018). Some even argue that because poverty is often seen as an urban "problem," rural poverty is "rendered invisible" and in essence, the rural poor are segregated by geography (Books, 2014, p. 74).

The research that does exist echoes deficit mindsets falling into the category of blaming the victim in which they claim that people believe that the rural poor are "better off" than urban

populations (because they can somehow live off the land), are unwilling to take government handouts (Books, 1997), and the believe that with hard work, impoverished people will overcome their circumstances (Chandler, 2014).

Deficit Thinking & Rural Poverty

Popular deficit attitudes add complexity to the problem when considering that "rural children are more likely to be living in deep poverty" (Chandler, 2014, para. 11). This notion of *deep poverty* is significant for our district as Bixby's low-income rate (free and reduced lunch) is significantly higher than the United States Department of Agriculture's most recent data on the rural poverty. Approaching 30%, Bixby school's free and reduced lunch rate is distinctly greater than the National poverty rate of 16.1%. Even more astonishing, it is more than double the urban rate of 12.6% (U.S. Department of Agriculture, 2020).

The phrase "deep poverty" itself brings visceral images to mind and are accurate descriptions of some of the poverty our most vulnerable families experience in Bixby. While stories like families living without plumbing or existing in deplorable conditions are infrequent even on the prairie of Bixby, it is just as common to hear that the community responds by finding the family another home or someone has anonymously paid to have the plumbing repaired. Books (1997) made a strong claim centering on the idea that the rural poor are "invisible," but our community would argue that when they see a family in need, they respond. However, issues with equity and visibility persist. For example, our teachers' perceptions are that our students generally "do well" and are successful in Bixby, but this view demonstrates that they may only be thinking of ("seeing") certain students and disregard others as expendable. Like the community that surrounds us, staff would argue that they, too, respond when students struggle; unfortunately, it is usually relegated to when they see a student on the cusp of failure that they

react with ways to "save" their grade.

Positives of Small Schools

Empathy and response from our community are not the only positive elements for vulnerable students at Bixby; there is a direct correlation to graduation rates and achievement equity in small schools (Cotton, 1996; Irvin et al., 2011; Malkus, 2018; West et al., 2022). And these positive outcomes extend to facilitating positive relationships and climate as well: "Research also shows that schools that create smaller, more intimate classrooms and learning environments … tend to be successful in fostering positive school-related dispositions, especially for historically underserved students" (Lawson & Lawson, 2013, p. 453; Cotton, 1996; West et al., 2022). For students, a small school ensures that they are "seen."

Being seen and accepted within the community as well as creating seamless learning opportunities holds many positive possibilities in K-12 districts like ours. So much so, that facilities like ours often become the sole place for families to receive any types of services; a fact that allows us to provide a spectrum of supports, described as "cradle-to-career system building [which] offers a variety of opportunities for working with vulnerable populations" (Lawson & Lawson, 2013, p. 464). Not only are our students able to stay in one district—in *one* building—for their entire prekindergarten through high school careers, they often receive vital social services not offered in the community at large (Dahill-Brown & Jochim, 2018; Malkus, 2018). And crucial to effective instruction, teachers have the ability to build positive relationships with each and every student.

Bixby's Culture

Thanks to the structure of our community and school, the physical systems are in place for change, and many of the social pieces are in place as well. Shifting the culture of grading

holds a great deal of promise for each and every student, but achieving this goal remains a complex problem for our culture. While the pandemic did not create our assessment problem, it had spotlighted and magnified it. Like many other schools during the pandemic, the stress of managing online learning while also teaching in-person, social distancing, and a pervasive tension revealed cracks in our culture that had long been ignored or overlooked. Senior staff members reported to me that prior principals in my district had never broached the subject of assessment practices. Any practices teachers engaged in were either self-taught or communicated (explicitly or implicitly) that "this is how it is done here." Staff were used to deciding their curricula as they saw fit, but outside of that realm, they required significant direction.

Similarly, cross-disciplinary leadership was virtually nonexistent. Having a staff who were so accustomed to leading themselves was at the same time a gift and a burden. On one hand, they were used to autonomy, but on the other hand, this independence sometimes gave them a false sense of expertise. Therefore, while the issue of grading stemmed from staff beliefs and misinformation, an adaptive problem like this one not only required teacher-leaders to bring about the needed changes—something of which they had grown unaccustomed—it also required that everyone commit to actively engaging in their own learning.

Whenever administrators work with their staff with the goal of changing the culture, and in my case, one of creating a culture of collaboration, this process can be filled with opportunities and challenges. DuFour (2004), who made professional learning communities (PLCs) the standard among educational professional practice, describes PLCs as the collaboration of groups of educators who are focused on improving student learning and outcomes. The resulting "culture" is results-focused and committed to school improvement.

However, if collaboration is the key to improving student learning, it is not merely a matter of creating the right conditions. To the contrary, principals run the risk of embedding "Contrived Collegiality" within their culture. Blankstein (2013) explains that this type of culture occurs when "teachers collaborate only on the surface without challenging one another's beliefs or approaches to teaching and learning" (p. 147; Hargreaves & Fullan, 2012).

In order to shift our culture, my staff and I needed to be able to give critical feedback, and I needed to model it. Additionally, shifting this culture began with involving staff—in this case, the team, specifically—and co-designing how worked on our problem of practice as well as seeing their expertise and knowledge base as valuable (Mintrop, 2018). These early discussions demonstrated what Mintrop (2018) refers to as "intellectual leadership" where participants are considering "not only what works, but also what they value. They have the courage to use their own judgment and to resist organizational authority when it is used to produce inequity and injustice" (p.88). In essence, my team members were my partners, where they were welcomed to share their insights, contradict each other, and collaborate to work through the heavy lifting of shifting their own mindsets about assessment.

To shift practice and mindset for the better, the team chose to focus on staff strengths (asset-based thinking) instead of issuing rules for grading or top-down policies. We mirrored Mintrop's (2018) conceptualization of effective professional learning by focusing on student artifacts and job-embedded, action research with the hopes that the practices would have the potential to make profound, sustained change in the team of teachers' instructional habits.

The Staff

According to Bolman and Deal (2017), organizations exist to serve human needs, and both, the organization and the people who work in the organization, need each other. In our

case, our report card data show that the three-year retention of teachers district-wide is 91%, and we have many teachers who spend the majority of their careers in the district. At the high school and junior high school level, teacher retention is even higher, at 96% (citation withheld to protect the school district's identity). The administration often joke that people stay for a lifetime when they join our staff. Grading practices are a "people problem," not just a structural one. And because the practices are primarily a people problem, the people must be the ones to resolve it. The issue then became how to shift practice and mindset for the better.

Sustainable Change

Our staff reported a great deal of self-fulfillment in their positions, so moving them toward more equitable grading practices meant building their capacity for this change as well, instead of imposing punishment for ineffective practice (Bolman & Deal, 2017). As a result, I did not want to fall into the trap cautioned by Bolman and Deal when they explained that "employers often fail to invest the time and resources necessary to develop a cadre of committed, talented employees" (p. 131). Instead of issuing rules for grading or top-down policies, my team needed to investigate staff—and their own—strengths (asset-based thinking), and, as a result, the work was much more complex. Building upon staff strengths is in direct contrast with deficit mindsets in which people (staff, teachers, students, etc.) are viewed through the lens of knowledge they lack (Dudley-Marling & Dudley-Marling, 2015; Valencia, 2010). While the work is more complex, it will be more sustainable in the end. We chose to mirror Mintrop's (2018) conceptualization of effective professional learning by focusing on student artifacts and job-embedded, action research with the hopes that the practices we developed would have the potential to make profound, sustained change in the team of teachers' instructional habits.

Complex, systemic change like the one my team and I envisioned were possible because

we have a few advantages from which other districts may not benefit. First, I knew my team very well and worked with two of them for four years. Additionally, I hired the other team member and selected him based upon his "fit" for the district. Next, having been the principal for four years, I was no longer "the new person," having worked hard to gain my staff's collective trust. Staff generally understand my expectations and core values, and I theirs. Finally, while those principles are all important, my staff *want* our students to succeed, and they are motivated to improve when they can see how their actions (positively or negatively) impact students.

Our professional development topics echoed similar themes since I became their principal: vulnerable populations and student learning. And staff were just beginning to connect their everyday practice as well as identifying personal professional goals relating to those two topics. In essence, despite the challenges we faced as a staff, we were ripe for change. Hargraves and Fullan (2012) explain that "successful movements occur when dissatisfactions with and tensions of the current system reach a breaking point" (p. 150). While I would argue that we were not at a breaking point, per se, we were seeing dissatisfaction with our current practice and observing negative consequences on our students' behalf.

Growing Excitement

More importantly, the foundations that we had laid readily aligned to what the authors claim is needed for successful change: "shared experiences, trusting relationships, and personal and social responsibility." They continue, "What pulls people in, teachers all the more so, is doing important work with committed and excited colleagues and leaders engaged in activities that require creativity to solve complex problems and that make a real difference" (p. 151). My small team of teachers were excited about to work on assessment and learning and readily shared

their enthusiasm with others. More importantly, the team understood that the work we discussed can be key to improving student learning and future success. We were doing work that mattered, and it mattered every day.

The excitement and purpose that drove the team was contagious. Early in the process, three additional staff members not involved in the team's work initiated discussions about their assessment practices and referred to discussions with various team members. One teacher offered for members of our team to observe and review her practice, and others created personal professional goals centering on student learning. This spreading engagement and curiosity as well as behaviors among staff indicated a good foundation for sustained, deep learning and had the markings of authentic, embedded professional learning far beyond the contrived collegiality created by many traditional professional development activities (Blankstein, 2013).

Grades and Grading

As the team built its foundational knowledge of grades and grading, we explored several common themes from the literature. We discussed general topics relating to assessment because they were important to ensure that we had a common language and understanding of grades and grading practices. As such, we reviewed the purpose of grades, common grading practices, as well as harmful and effective practices.

Purpose of Grades

Because grades and grading are a bedrock of our public school systems, it seems a little simplistic to define their purpose. However, brief discussions with staff, and even a quick review of literature show varied reasons for why teachers assess students. Generally speaking, assessments indicate a level of mastery on content standards or skills (Fisher et al., 2011; Guskey, 2010; Guskey, 2020a; Jensen & Snider, 2013; O'Connor, 2011; Popham, 2001;

Wormeli, 2006) and their purpose is categorized into two categories: evaluation and feedback (McTighe & Willis, 2019). An assessment may show how students perform in comparison to others (evaluation), or it may communicate a comparison of where they student currently functions to where they need to be to achieve mastery (feedback). *Communication* is at the very core of these definitions, or purposes, for assessment; however, there is some disagreement over who needs to know (or use) that information (Guskey, 2020a; Marzano & Heflebower, 2011; Muñoz & Guskey, 2015; O'Connor, 2011; Popham, 2001).

The above examples were centered on communicating to the student; however, assessments are also key information for the teacher in that they can be used to guide or inform instruction. The team and I chose a more nuanced approach in how we viewed our purpose for grading; we felt that assessments should be a collaboration between the instructor, thus informing instruction and the student, indicating his or her level of mastery and what he or she needed to do next to be successful (Popham, 2001; Popham, 2018).

Our team's view of assessment aligned closer to Next Generation Assessments philosophy which assert that the focus should be on providing clear paths for student mastery, and "assessments should help students understand better who they are, how they learn, and what they need to do to achieve their aspirations" (Conley, 2018, p. 13). Next Generation Assessments do not prescribe specific assessment methods but promote a stance which encourages the use of local assessments that collect data based upon multiple measures of the standard or skills. Instead of hearkening to industrial era assembly line procedures for sorting and ranking students, more current approaches are centered on key (anchor) standards and rooted in how students learn (Conley, 2018). While learning can occur in the absence of a grade, assessments have the potential to contain valuable information on learning, on mastery, and/or

on the instruction (Fisher et al., 2011; Guskey, 2020a). At the very core, effective grading is "feedback" for both the teachers and students (Wiliam, 2011; Hattie & Yates, 2014).

Grading Practices

While new to my district, the topic of grades or grading practices is nothing new among educational research. Research on grading practices has ranged from evaluating the reliability of assessments and report card grades and their related outcomes to teacher perceptions of grades and grading practices in addition to standards-based report cards and large-scale accountability assessments (Brookhart, et. al, 2016). However, there are several experts (Muñoz & Guskey, 2015; O'Connor, 2011; Percell, 2017; Schimmer, 2016; Wormeli, 2006) that believe "the current system" and its assessment practices need to be changed, with some arguing for a complete overhaul, even calling for shifting to standards-based grading (SBG) at all levels (Muñoz & Guskey, 2015; Marzano & Heflebower, 2011; Wormeli, 2006).

The people in my district often point to post-graduation data as evidence that students are successful after they graduate. While many of our students *are* successful, some are not, and those groups are nearly always comprised of our most vulnerable populations. Because of our perceived statistical success, moving away from our current grading system would not only be complex and difficult but also unwelcome by our stakeholders. Like the results of Swan's et al. (2012) findings, our families prefer traditional grade reporting (as opposed to Standards Based Grading, for example) because they feel they "understand" a percentage better than a term descriptor like "exemplary."

As a result of our community's preference and staff attitudes, the work to shift to a system of learning needed to be accomplished within our current, averaged grading structure. Therefore, while some authors call for reconstructing the entire system (Marzano & Heflebower,

2011; Muñoz & Guskey, 2015; O'Connor, 2011; Percell, 2017; Schimmer, 2016; Wormeli, 2006), few, if any, studies actually exist that demonstrate how to shift a high school staff from a system focusing on points to a culture centering on learning while still maintaining a percentage-based framework (Reeves et al., 2017). The fact that we were attempting to create more equity within our current system, increased the complexity of the task set before us, and not only required me to move into unexplored territory as a leader but also rely upon my teacher leaders to extend their leadership beyond anything they had attempted thus far.

Harmful Practices

A review of literature also reveals that ineffective—potentially harmful—assessment reporting practices have been consistently identified among researchers (Fisher et al., 2011; Guskey, 2020a; Marzano & Heflebower, 2011; Muñoz & Guskey, 2015; O'Connor et al., 2018; O'Connor & Wormeli, 2011; Reeves et al., 2017; Wormeli, 2011). In 1989, Grant Wiggins argued the need for fewer standardized tests and more authentic assessments, asserting, "Authentic tests must come to be seen as so essential that they justify disrupting the habits and spending practices of conventional schoolkeeping" (p. 712). Authentic assessments mirror how the content is used in the content area industry. For example, multiple choice standardized tests are not authentic assessments of students' writing ability. *Writing* is an authentic assessment of their ability. Inauthentic assessments muddy the picture of what a student actually knows or can do. Other harmful assessment practices include averaging and aggregate grade, integrating zeroes into a student's grade, including behavior in an aggregate grade, grading "practice," and not allowing retakes.

Unfortunately, little appears to have changed in more than 30 years since his call to action as researchers have consistently called for more change. So consistently, in fact, that

some researchers claim that administrators must immediately stop the following practices in their buildings: (1) using an average to determine an aggregate grade (Marzano & Heflebower, 2011; O'Connor, 2011; O'Connor et al., 2018; O'Connor & Wormeli, 2011; Reeves et al., 2017; Schimmer, 2016; Wormeli, 2006); (2) grading homework (Fisher et al., 2011; Reeves et al., 2017; Wormeli, 2006); (3) including zeros on a 100-point scale (O'Connor & Wormeli, 2011; Reeves et al., 2017; Wormeli, 2006); and (4) grading behavior/including behavior in academic grades (Fisher et al., 2011; Guskey, 2020a; Marzano & Heflebower, 2011; O'Connor, 2011; O'Connor et al., 2018; O'Connor & Wormeli, 2011; Reeves et al., 2017; Wormeli, 2006). To make grades more meaningful and reflective of student learning, on the other hand, schools should *start* allowing: (1) reassessments (redo's and retakes) for full credit (O'Connor, 2011; Wormeli, 2006; Wormeli, 2011); (2) current scores to replace old ones; and (Marzano & Heflebower, 2011; O'Connor, 2011; Wormeli, 2006; Wormeli, 2011) (3) the use of median or mode instead of mean calculations for grades (O'Connor, 2011; O'Connor et al., 2018; O'Connor & Wormeli, 2011; Schimmer, 2016; Wormeli, 2006; Wormeli, 2011).

Averaging.

While much of the respective recent "push" among researchers has been centered on standards-based grading and grade reporting (Schimmer, 2016; Wormeli, 2006; O'Connor, 2011), this shift continues to be an uphill battle for high schools and their stakeholders. Averaging grades on a 100-point scale is the traditional approach, particularly for secondary schools. The notion that an average is somehow more accurate is deeply entrenched not only in high school educators but among their families as well. Parents who preferred traditional grade reporting to standards-based grade reporting asserted that they "understood" was a percentage meant; whereas, an "exemplary," they did not (Swan et al., 2012). However, while the mean

may be mathematically precise (accurate), it is frequently invalid and does not reflect what students actually know (Guskey, 2011; O'Connor et al., 2018; Schimmer, 2016; Wormeli, 2006). Clearly, any shift away from traditional averaging would take a great deal of professional development for staff and education for families alike.

Zeroes.

While any work on grading practices will require learning for all involved (staff, students, and families), one of the most complex elements to understand is that averaging scores are inherently flawed (Wormeli, 2006). And including zeros within that averaged score compounds the issue (O'Connor & Wormeli, 2011; Reeves et al., 2017; Wormeli, 2006). Averaged scores are flawed because they can be heavily influenced or skewed by a single score (Fisher et al., 2011; Guskey, 2020a; Marzano & Heflebower, 2011; O'Connor, 2011; O'Connor et al., 2018; O'Connor & Wormeli, 2011; Reeves et al., 2017; Schimmer, 2016; Wormeli, 2006). Instead of having a picture of the student's understanding which is based upon a preponderance of evidence, an averaged score takes into account every point, small or large.

When zeros are included into this score, discerning what the student knows becomes even less clear. After all, a zero on a single grade record can be exceedingly difficult—if not impossible in many cases—from which to recover. Zeros result in significant flaws in nearly all systems of assessment—whether they are traditional averaged scoring methods or nontraditional scaled evaluative means (such as a four-point rubric scale). Clearly, though, zeros impact averaged systems greatly. So much so, that some authors claim that the integration of zeros into an averaged grade not only detracts from an accurate view of a student's learning, they also invalidate any aggregate score which includes a zero score (Guskey, 2011; O'Connor et al., 2018; O'Connor, 2011; Wormeli, 2006; Wormeli, 2011).

Including Behavior in an Academic Grade.

Similar to averaging zeros into a students' score, including too much in a single letter grade invalidates it as well. Historically speaking, grades were used to sort and rank students (Brookhart et al., 2016; Conley, 2018; Guskey, 2011), and teachers' perceptions of the purpose and meaning of grade continue to parallel historic attitudes. For example, while grades were used to sort students, teachers also used them to compare student abilities in addition to using them as a tool to manage behavior (Guskey, 2011). Even modern-day current practitioners and researchers believe that the "right" forms of assessment can improve student effort and motivate them to do their work (Jensen & Snider, 2013).

Including non-academic factors in students' grades can create an even greater issue of inequity for our vulnerable students. For teachers who view their students through a deficit mindset, they have the potential to grade their students more harshly (Dudley-Marling & Dudley-Marling, 2015) as well as have predisposed negative notions of the students' behavior (Valencia, 2010). Together, these two factors have the power to significantly impact accurately assessing what students have learned about a particular subject or accurately assessing their level of mastery of a specific skill.

Whether grades are used as comparisons or to control behavior, typical practice for assessing and ranking students has been—and currently is—to assign a single grade per content/class (Guskey, 2011). These grades, which are often referred to as omnibus grading are fraught with inaccuracies (Fisher et al., 2011; DuFour & Marzano, 2011; Guskey, 2020a; Marzano & Heflebower, 2011; O'Connor, 2011; O'Connor & Wormeli, 2011; Wormeli, 2006). Because they contain more than just assessment of students' progress toward content area goals or mastery of skills, researchers like O'Connor (2011) and Wormeli (2006) claim that omnibus

grades are "broken." Rather than reflecting a level of mastery, grades are often used to manage student behavior: issuing points for participating in classroom discussion; deducting points from late assignments; and dangling grades as a means to motivate students to do work (Brookhart, 2016; O'Connor, 2011; Wormeli, 2006).

Also adding to the problem of grading are inconsistencies not only between grade levels but even among teachers of the same grades. Fisher and his coauthors (2011), explain that omnibus grading is a "mix of compliance and understanding" and that "the percentage that each contributes to students' grades varies across teachers, schools, districts, and states" (p. 46). Policies like reducing grades for late work; giving extra credit for compliance; or assigning points for class participation are all behaviors and detract from communicating what a student truly knows or understands about the content (Fisher et al., 2011; Guskey, 2020a; O'Connor, 2011; O'Connor & Wormeli, 2011; Wormeli, 2006).

Grading "Practice".

A common practice in averaged systems at the high school is grading homework or "practice." Unfortunately, grading practice creates a culture of compliance, not appropriate risk-taking (O'Connor, 2011; O'Connor & Wormeli, 2011; Wormeli, 2006). A better use of practice is just that, allow students to practice and receive feedback on their progress. Grades should "reflect student understanding of the content" (Fisher et al., 2011, p. 46); thus, *practice* should be reserved for concepts they understand but need to build permanence (Wormeli, 2006). Teachers may then review that practice and use the information they garner from it to adjust (*inform*) their instruction (Popham, 2001; Popham, 2018).

Effective Practices

While removing harmful practices is necessary, just as important to the process is ensuring that effective assessment practices are integrated as well. Some of the most powerful practices centered on formative assessments—assessments that communicate the students' current understanding and how the teacher can shape their instruction to fit the learners' needs. Feedback, too, is part of the formative assessment process and includes feedback from the teacher as well as the student. Feedback from the teacher guides the student on how to improve while feedback from the student guides the teacher on the next steps for instruction. Finally, including the student even further in the assessment process is the use of student self-assessment. This powerful formative assessment gives students ownership of their learning through a reflective process that empowers them by making them aware of their own understanding of the standards or material.

Using Mode instead of Mean.

Instead of averaging scores (mean), calculating a student's grade based upon mode is a much more valid—and fair way—to evaluate and communicate levels of understanding and progress to mastery (Wormeli & O'Connor, 2011; Wormeli, 2006). The use of mode, which identifies the most frequently occurring score, eliminates outlier scores that could be atypical of the student's performance. For example, while a single zero or low score can skew a student's average down to failure or near-failure, when the same scores are evaluated using mode, they are more reflective of what the student has actually learned (O'Connor, 2011; Wormeli, 2006; Wormeli, 2011). Using mode not only gives students opportunity to recover when they struggle, it gives them hope for success all while keeping a focus on learning.

Retakes.

Appropriate risk-taking is vital to the learning process, and if we as educators truly care about our students' learning, then no student should be punished for not understanding content upon the first attempt. To encourage a culture of appropriate, academic risk-taking, all (valid) assessments should be eligible to be retaken. In fact, when students perform poorly on summative assessments (e.g., unit tests), researchers believe that students should be allowed to reassess for full credit (O'Connor, 2011; Wormeli, 2006; Wormeli, 2011). The reason for allowing retakes for full credit is that learning is neither linear nor achieved at the same pace by everyone (Wormeli, 2011). Teachers who engage in this type of practice place student learning above the instructor's possible inconvenience of grading and understand that assessment and learning are inseparable. As a school, when we ground learning through practices like retakes—a philosophy which prioritizes learning above holding on to antiquated grading practices—we are shifting the culture from one of compliance to one of collaboration and move towards assessment for learning instead of assessment as learning.

Formative Assessments.

Another component to improving the culture of assessment is through using formative assessments. Formative assessment has the potential to improve learning two times faster (Popham, 2018), with others claiming "*20-30 times* as much positive effect on learning as the most popular current initiatives" and shows to be "10 times as cost-effective as reducing class size' (Schmoker, 2018, p. 70). Just as important, it may add six- to nine months of learning growth and accounts for students learning four times faster than other interventions (Schmoker, 2018).

In addition to integrating more formative assessments—or assessments *for* learning

(Guskey, 2011; Guskey, 2017; Guskey, 2020a; Guskey & Anderman, 2013; Popham, 2001; Popham, 2018; Wiliam, 2011), assessments must be more authentic. In order to be authentic—and in turn, more equitable—assessments should reflect the real work done by practitioners and experts in the field (Cappuis et al., 2017; Conley, 2018; Wiggins, 1989). Authenticity and equity are increased when they are "responsive to individual students and school contexts" (Wiggins, 1989, p. 704). Using formative assessments work two-fold: one in which the teachers use student performance on frequent, smaller assessments to guide instruction and the other in which students are able to see progress grow to mastering the skills or standards (Popham, 2001; Popham, 2018). For example, a pretest can act as a barometer for student to indicate where they are to where they need to be (McTighe & Willis, 2017; Hattie, 2012). These assessments for learning act as lines of communication between the teacher and the student and interplay to improve both instruction from the teacher and learning by the student (Cappuis et al., 2017; Conley, 2018; Popham, 2001; Popham, 2018).

Formative assessments are not simply giving a pre-test by randomizing a summative test, however. An effective assessment for learning "is a planned instructional *process* in which assessment-collected evidence is used either by teachers or by students to make any necessary adjustments in their current efforts" (Popham, 2018, p. 93). Well-developed forms of this activity are aligned to targets (Popham, 2018), and teachers use the data they collect to adjust their instruction to improve student learning; they are strategic, intuitive, active in nature, and bring learning to the forefront of the instructional and assessment process.

Feedback.

An essential element of formative assessment, then, is referred to as feedback. While assessments of learning by themselves communicate basic information, descriptive feedback can

"communicate useful information and directly reflect specific learning goals" (O'Connor, 2011, p. 4). Feedback supports Constructivist viewpoints because this element of the process sees assessment as an actively engaging the students in thinking about their learning, and as Wiliam (2011) posited, "the whole purpose of feedback should be to increase the extent to which students are owners of their own learning" (p. 132). Constructivism holds that learning is an active process in which learners create their own understanding of new concepts (Greeno et al., 1996; Lambert et al., 2002). Students play a vital role in the formative assessment-feedback process by using or giving feedback to self-assessing their progress to mastery as well as drawing upon their knowledge of how they learn to use different strategies to understand (Popham, 2018; Wiliam, 2011).

In Hattie's (2012) seminal work, he found feedback can be one of the most significant ways to improve student learning. To be effective, however, feedback cannot simply be a score at the top of a paper; students must become *owners* of their learning. Therefore, responding to how the student answers an essay question, evaluates their findings in a lab report, or performs a musical piece engages the student to reflect upon his or her performance and go beyond giving an often meaningless score. Activities like these offer opportunities for self-assessment as well and when combined with a clear path (mapped by the teacher) for how to achieve mastery on specific skills and standards, it creates a quality dialogue between the teacher and student, resulting in a meaningful assessment process and learning (Black & Wiliam, 2010; Popham, 2018).

Self-Assessment.

Feedback is not limited to the teacher and student, however. Other meaningful approaches of integrating feedback and self-assessment are through the use of peer- and self-

assessment. Clearly, the opportunities for feedback are numerous, but "the key for learning lies in the feedback that an assessment offers to both teachers and learners" (McTighe & Willis, 2019, p. 93). Teachers who are adept at creating descriptive assessment "can fold assessment results back into instruction" (Cappuis et al., 2017, p. 2) and utilize data to that are meaningful for students to improve. This communication "includes processes and strategies that encourage and support greater student achievement, especially for struggling learners" (Cappuis et al., 2017, p. 3).

Self-assessment moves feedback out of being the teachers' sole responsibility to partnering with the students and engaging them in the process. Students assess their understanding of concepts or skills and indicate what they need to progress in their learning. In this approach, the activities are shifted to the students reporting their learning and giving the instructor feedback on what needs to be retaught to assist them in their learning. As students become adept at becoming aware of their own understanding and communicating what they do not, this process can increase student learning; Wiliam (2011) described the power of this process: that including students in establishing their own instructional goals and self-assessing their progress towards the, it doubled their learning. However, since our goal was to shift our staff and students' mindsets to a focus on learning, Hattie (2012) summarized the positives best: "This may require a move from talking less about how we teach to more about how we learn, less about reflective teaching and more about reflective learning, and more research about how to embed feedback into the learning process" (p. 152).

There is strong evidence from research that while practices like averaging grades, including zeroes in an aggregate score, and not allowing retakes are harmful to student learning, other assessment activities can significantly improve it. Using mode creates a more accurate

view of a student's grade, and allowing retakes acknowledges that student learning is not linear. However, integrating processes for formative- and student self-assessment have the potential to significantly *increase* student learning.

Principal Leadership

To achieve the type of change that truly impacts learning, literature is clear: School leaders are vital to that process. While I could not find much research with regards to schools revamping grading practices within their existing percentage-based grading systems, researchers urge that "the need to implement healthy grading practices is an urgent one; this action will benefit all learners, but especially those who are struggling" (Reeves et al., 2017, p. 45; see also Jensen & Snider, 2013; Theoharis, 2007). And it is evident that school leaders play important roles in shifting staff mindsets. Specifically, the principals and administrators are key for allowing staff time to work on these topics; providing a focus for improving assessment literacy across the building; and taking active roles as learners of assessment methods as well.

Teachers need time to develop core competencies as well as collaboratively work through assessment issues (Guskey, 2017; Hattie & Yates, 2014; Hill, 2011; Nolen, 2011; Nordengren & Guskey, 2020; O'Connor et al., 2018). Typically, this support is identified as opportunities for collaboration and flexibility in scheduling (O'Connor & Wormeli, 2011; Wormeli, 2006). Deep, sustained change takes significant time and encouragement, and in addition to revising professional development plans, building and district leaders need to support teachers as they work through shifting grading practices (Nordengren & Guskey, 2020; Blankstein, 2013). To provide the needed focus among staff and administration, grading practices should also be a component of district professional development and add that formative assessment practices (assessments for learning) should also be included as part of schools' improvement goals and

culture (Hill, 2011; Nolen, 2011; Blankstein, 2013). Leaders must also be mindful to assess the effectiveness of their current professional development practices as well as maintain a clear focus on the school's goals (Guskey, 2017; Blankstein, 2013).

Hill (2011) and Wormeli (2006) each argue the need for strong school leadership when working to improve assessment practices. Hill goes so far as to say that work to transform how teachers assess is a significant shift in mindset that takes time to become embedded in the culture in order to see sustained changes in practice. Concerns surrounding grading practices are not ones that can be addressed through straightforward—or technical—means, like developing a new bell schedule or creating a system for checking out computer devices to students. The problem of grading practices is much deeper and rooted in school culture and teacher mindsets. Clearly falling within Heifetz's et al. (2009) criteria for an adaptive challenge, it can only be addressed "through changes in people's priorities, beliefs, habits, and loyalties. Making progress requires going beyond any authoritative expertise to mobilize discovery, shedding certain entrenched ways, tolerating losses, and generating the new capacity to thrive anew" (p. 19).

Heiftez et al. (2009) posit that an adaptive challenge requires leaders to diagnose the problem while also determining which elements of the system are worthy of preservation. While this cycle of inquiry is not directly related to leadership, my leadership abilities were a key component of its success. In essence, assessment challenges—a challenge that permeates a school culture—require leadership that "anchors change in the values, competencies, and strategic orientations that should endure in the organization" (p. 15). It requires what Shields (2020) refers to as transformative leadership. A transformative leader is committed to second-order change, and like adaptive change, it is a deeper type of process; however, this mandate focuses on shifting toward a more equitable system and mindset among the participants within it

(Shields, 2020). Furthermore, another result of the pandemic and its impact on local jobs has been that our vulnerable population has grown.

Professional Development

Heifetz (2009) posits that "organizational adaptation occurs through experimentation" (p. 15). And I would argue that teachers (the good ones, at least) consistently try new and better ways to deliver their instruction on a daily basis. However, this type of experimentation is not often acknowledged in school systems. To the contrary, we tend to inundate our staff with new initiatives annually without giving teachers time to master anything (Wiliam, 2011). This "shotgun" approach to professional development results in "little or no real improvement in what happens in the classrooms" (Wiliam, 2011, p. 29). However, effective assessment practices are crucial to improving student learning. When teachers are equipped with the power of using assessment to inform their instruction, students benefit (Wiliam, 2011). However, merely providing an in-service on it is not enough.

Hattie (2012) argues for school leaders that provide opportunities for staff to talk about assessment within the frame as feedback for their instruction. He says, "We need to move from the prepositional divide of assessment as 'assessment of' and assessment for' to assessment as feedback for teachers" (p. 185). Popham (2018), on the other hand, questions why formative assessment is used so infrequently in American schools when so many researchers know the benefits. The reality is, it takes time. It is not a program to implement but a process that must be carefully constructed and integrated (Hattie, 2012; Popham, 2001; Popham, 2018; Wiliam, 2011). These reasons are why we are using a cycle of inquiry. At its core, it is the very best of professional development—job-embedded and collaborative. It is rooted in data and based upon a real, actionable problem within our own school and taps into staff members who take risks and

want to learn, as well as improve their practice and knowledge base.

Empowering Staff

If we are to provide staff with authentic opportunities to evaluate their instruction, then leaders must be prepared to collaborate with them as well (Hattie, 2012; Nordengren & Guskey, 2020). This process requires empowering staff with knowledge of assessment literacy realizes the following tenets of transformative leadership outlined by Shields (2020): Recognizing that our school systems need substantial change regarding equity; working to positively impact belief systems; equitably empowering staff and students; "focusing on democracy, equity, and justice" (p. 95); and "exhibiting moral courage" (p. 164). Building staff understanding of assessment and collaborating with them as they work through the challenges is not only professional learning at its best, it goes to the very heart of the tenets of addressing underlying beliefs and redistributing the power within our school system.

Next Generation Assessments

Working toward Next Generation Assessments (Conley, 2018) fit precisely within Shields' (2020) tenets. Specifically, they encourage educators to focus on meaning-making *with* the student and prepare students for learning beyond high school. It is a new way of thinking about data, a new way of exploring how and why we assess students, and this approach includes the notion of giving valid feedback—something that is vital to our vulnerable populations. Conley (2018) even asserts that assessments should help students achieve their aspirations, not merely reach a level of "success" that is imposed upon them.

Transformative Leadership

This positive outlook on the power of assessment ties with Shields' (2020) assertions regarding transformative leadership. Specifically, Shields (2020) argues for schools that

promote individual student growth, socializes them, and prepares them to be productive citizens in a democratic society. When we use feedback as an opportunity to activate students' funds of knowledge and see students' backgrounds as assets, we are also engaging in culturally responsive actions and mindsets (Dudley-Marling & Dudley-Marling, 2015). Not viewing students as "deficient" because of their background requires moral courage on the part of the instructors; however, this lens also demands courage of the administrators to educate and empower staff to see these students as valuable members of our community (Shields & Hesbol, 2020).

Assessment Literacy

An important component of empowering staff learning and outlook it to build up their knowledge of assessment literacy. Assessment literacy is important for all stakeholders within a school as it allows them to root their decisions in valid assessment data (Cappuis et al., 2017; Guskey, 2020b). Therefore, both principals and teachers should work to improve their vocabulary regarding assessment and collaborate to define specific terms that are key to their emerging culture. Hill (2011) refers to school principals as "conductors of change" and calls for them to be assessment literate as well (p. 356). As assessment literate conductors of change, these leaders will be equipped to help develop clear policies and procedures regarding assessment practices for their schools (O'Connor & Wormeli, 2011).

Researchers call for leaders who are "in the trenches" with staff—engaging in essential conversations with them as well as working to make the school a "unit of evaluation" and collaboratively moving to achieve essential results (Hattie, 2012, p. 191), a sentiment Mintrop (2018) echoes in his designed-based approach for school improvement. For example, in our cycle of inquiry, I purposefully refrained from guiding the team to a specific answer or decision.

I wanted to participate as a team member and communicate that their viewpoints and background knowledge was just as important as my own.

Shifting Mindsets

To improve all students' educational experience, our school must work toward the goal of shifting mindsets towards one of learning. While Hattie points to making the school a collective system of evaluation for all, Wormeli (2006) asserts that leadership should facilitate a culture immersed in assessment where administrators model grading practices, arrange peer observation, affirm early adopters and risk-takers and embed this work within the school's professional development. Sustained change, though, must go beyond the school administrators. While the charge may be taken up by the leaders, the ultimate goal is to positively impact the culture.

Work like the type described here may feel risky to the teachers involved, but there is also distinct risk to the school administrators. In our volatile political climate, topics of equity can be hot buttons, especially in small, rural districts. At a recent conference, several superintendents of rural districts remarked to me that they had had to "re-frame" some of their initiatives because their families had protested the use of some terms that were seen as "too liberal." Within my own district, community members typically follow conservative views and align with other rural schools in the belief that many of the initiatives we must follow come from political outsiders (Dahill-Brown & Jochim, 2018), so engaging in conversations with families and staff about what equity means for us requires moral courage on my part (Shields, 2020). For example, during my supervisory duties where I frequently interact with families, I often carried a book or printout of journal articles to read when I could find a quiet moment. I was very aware that onlookers might read the titles of those texts which could then result in spreading rumors that I was initiating some "liberal agenda."

However, because our students and families are fairly sheltered in their exposure to topics like poverty and equity, I welcomed conversations with anyone who questioned my reading and used that dialogue as a means to educate our community as a whole. These conversations strengthened my commitment to our work on assessment and why it is important. If our goal is to create a culture of learning for all, then as a leader, I must ensure that our aim includes equitable access—to instruction, to assessment, to feedback—for our most vulnerable students as well (Shields & Hesbol, 2020; Theoharis, 2007).

Theoretical Framework

My own experience of working with diverse and vulnerable populations not only impacts how I lead my staff but also my passion to improve the assessment culture for all our students. As a high school English Language Arts teacher, I was greatly influenced by two literary theories: New-Historicism (Eagleton, 1983) and Transactional Reader-Response (Rosenblatt, 1978). Both theories deal with how a reading and a reader's experience with a text are impacted by the era in which it is written as well as by the readers' experiences. As a school leader, this theoretical framework gave me a foundation, of which I am constantly cognizant. I recognize that when presented with new information (professional development, district initiatives, etc.) staff will make their own meaning through the lens of their experiences as well as construct their own understanding and approach to assessments. My goal is not to change their experience or direct their understanding but to participate in the process and collaborate with them to build a better culture of learning for students and staff alike.

Meaning-making

In New-Historicism, a reader not only brings his or her experiences to the text but the text itself can also be read through the lens of historical events that influenced the author and society

at the time. In essence, we read a text with the lens of knowledge that specific events took place and certain beliefs were held in society (or by the author) at the time it was written. Similarly, Reader-Response acknowledges that a text's meaning—or the way a reader creates meaning from the text—is impacted by the reader's experiences, background knowledge, and the society and culture in which he or she lives. In other words, meaning does not come from the text. Instead, readers create their own interpretation of it, and a text can have as many "meanings" as there are readers. Just as when a reader creates their own reading of a text, when working through challenges, people may have multiple understandings and subjectivities related to the problem as well as approaches to resolving it.

The notion that we create our own meaning from a text and have multiple understandings of a problem greatly influenced my philosophical stance of instruction. Because Constructivist Theory (Greeno et al., 1996; Lambert, 2002) takes the position that learners actively create their own meaning when introduced to new concepts, it paralleled elements of the literary theories to which I aligned. How I taught was rooted in helping students create their own understanding of texts and new concepts. This framework lent itself to how I work with staff as a school administrator, and mirrors how Shields (2020) describes transformative leadership as working with and empowering the teachers in my building. By empowering them and acknowledging that they possess their own lenses and create their own understanding of ways to approach problems, is a means to build sustained, effective change for our students. This lens is a firm foundation for how I approach working with staff as well as dealing with an instructional challenge. Working to improve our culture of assessment is no different.

The team's concerns on grading practices are something that will not have an easy or quick solution, requiring me as the school leader to shift organizational frameworks to lead this

change. Organizations often "feel pressure to solve problems quickly, to move to action" (Heifetz et al., 2009, p. 7), and promotes the most important component of adaptive leadership as "diagnosing" the issue at hand. I resisted the temptation to be reactive and spent the first year returning from the pandemic shutdown, carefully observing teachers' grading practices and attitudes toward students.

Working to change deeply entrenched habits and beliefs under normal circumstances is an adaptive challenge, and the pandemic along with our growing vulnerable populations made it even more complex. However, it was not without reward. As a team, we had accepted Shields' (2020) mandate for deep and equitable change, and we felt that revealing underlying assumptions about our vulnerable students was important—not just for improving their learning but to value their contributions to our school community. In many ways, giving the team a space to talk about their feelings of powerlessness when students showed apathy toward an assessment was empowering for them. Sharing their feelings of inadequacy and admitting that they had fallen into the belief that struggling students "just don't care" allowed them to be vulnerable with each other and move beyond those beliefs to reconsider our own values and how we distribute our power—whether it was within the classroom or school as a whole. At the very core, we were "critiquing the status quo" of our school and our own practice (Shields, 2020, p. 141).

Working with Staff

The delicate balance of approaching long-held beliefs and opinions on grading practices is one that definitely reflects the challenge of "how to hold an organization together without holding it back" (Bolman & Deal, 2017, p. 74). Our indicators of success evidenced that our staff did a "good" job of teaching students; therefore, I knew they had the capacity to perform better—in other words, the structures existed for a top performing team. However, the work

required to transform our school into a more equitable place for all students needed to "go beyond trying to ensure a socially just education for all students, to critique, deconstruct, and work to transform structural and societal inequalities and disparities" (Shields, 2020, p. 142). It required us to become agents of transformation. This delicate balance, then, came in working with staff to realize their beliefs and offer alternatives and opportunities to change them. Shields (2020) refers to that component as building collective trust—trust in students, trust in families, trust in staff and "implies a rejection of deficit thinking and a strong belief that everyone is working to the best of his or her ability to fulfill the familial, societal, and educational roles assigned to them" (Shields, 2020, pp. 146-147).

Through this work, we engaged in Triple Loop Learning (D. Renn, personal communication, June 10, 2020): We recognized that a problem existed within our culture, specifically centering on student learning and grades. From there, we identified our tacit assumptions—many of which related to gaming our system of points for grades as opposed to actual learning. Furthermore, our staff viewed inflexible practices as adhering to rigorous academic standards. We used these tacit assumptions to help us clarify our issue as an adaptive problem because it was not one that could be resolved through a tactical change like buying a new curriculum or altering the master schedule. It was clearly embedded within our school culture.

Next, we held conversations with the team as well as with the staff as a whole and then the team reviewed research to understand the problem and identify the change process. In contrast to typical school professional development, we resisted the impulse to jump to a solution as soon as we thought we had identified the problem. Instead, we turned to research-based literature, specifically focusing on effective assessment practice as well as transformative work

for engaging with our vulnerable students (Mintrop, 2018; Mintrop & Zumpe, 2019). Research confirmed best practices regarding assessment and we began to investigate the use of student self-assessments. As a leader, I wanted to prioritize an authentic climate of collaboration within our small team—one where everyone's viewpoint was valued. Since my leadership style is rooted in the belief that we create our own meaning from what we experience, what we learn, and what we read, I felt it was important to develop our team as a collaborative group, that resulted in their ability to take the lead (thus empowering others) in this important work of equitable change (Shields, 2020).

Our next steps included designing a research-based intervention; in our case, we decided to use a student self-assessment which teachers could use to inform their instruction. After implementing the intervention for a quarter, I collected the qualitative data in the form of notes from monthly meetings with the team and team responses to their weekly reflections. In the end, we evaluated the intervention and decided whether it was worthy of further use. We then looped back to the beginning the process and proceeded to plan the next cycle of inquiry in which we would build upon the work we had already completed and plan to include additional staff members.

CHAPTER IV: DESIGNING AN INTERVENTION

Early in this cycle of inquiry, the team agreed that there were concerns about our grading practices, so after consulting the literature, we clarified our problem of practice and used research to guide our creation of an intervention. Our regular meetings served to develop our habits as improvers. If we were to create a more equitable culture of assessment—within our team member's classrooms, at least—then we needed to identify which interventions we wished to test. Perry, Zambo, and Crow (2020) describe this stage of the process as a Theory of Improvement, explaining that as a team, we would be moving from "problem analysis to actually tackling the problem during the testing phase" (p. 90).

Based upon Mintrop's (2018) model of designing school improvement, the eight chapters in this work mirror his cycle of inquiry for equity. We then reviewed research to help us select an appropriate intervention to test. Ultimately, we wanted an intervention that would allow students to assess their level of understanding and provide teachers with feedback on how they could alter their instruction. We settled on using a student self-assessment in the form of a digital exit slip (Wiliam, 2011) that teachers could issue to students and then use the student responses to guide their next steps for teaching. An effective formative assessment is a process (Popham, 2018; Wiliam, 2011), so we wanted to monitor how teachers modified their instruction based upon student responses. As a result, we also developed a reflection for the teachers to complete to collect evidence that described their changes.

The team's first step was to identify a third member for our group. Our staff is exceptionally small, with fewer than 20 teachers at the high school level. Leading such a small staff has advantages and disadvantages. One disadvantage is that the same people often volunteer or are asked to participate in activities outside of their classrooms, like serve on

committees or lead professional development. On the other hand, an advantage is that nearly every instructor teaches every pupil at some point during that student's high school experience, so any positive impact to one teacher's instruction has the potential to touch every student who graduates from our high school. I trusted that the members of my team would decide upon someone who not only needs to improve but wanted to make the changes necessary for that improvement.

The team decided they wanted to invite a teacher with more than 20 years of experience to join our group. The final team member, "Anne" (a pseudonym) is a quiet staff member who does not serve on many committees, so participating on our team was a good opportunity for her to extend her professional capacity. In the end, her length of time as in education and tenure at our district were a good balance for our relatively young team.

Exploring Interventions

We then began to explore designing an intervention to "test." The team had already posited some possible solutions to our "grading problems," so identifying an intervention was not a large step outside of their comfort zone. To ensure that the intervention was research-based, we followed the three criteria Mintrop (2018) outlined: "interventions consist of carefully planned tools, activities, or organizational formats"; "trial and error in accomplishing outcomes is deliberate; it is under-girded by a theory of action drawn from the professional knowledge base"; and "data are collected according to reliable procedures" (p. 135). The intervention consisted of identifying one change—an addition or subtraction of an activity from their already existing assessment practices and document the resulting data.

The team discussed several possible interventions. Previously, team members had remarked that they felt compelled to "grade everything" because they thought that students

would not complete learning activities without points assigned to them. We delved into the notion of why our students did not complete their homework as well as what the research says about not grading practice or formative assessments. The team commented that they often felt overwhelmed by grading; specifically, they cited two issues: grading massive amounts of homework and providing opportunities for retakes (which they thought required creating additional test versions and resulted in even more grading).

"Catching" Students Before Failure

Teachers' policies on retakes vary widely among the high school teachers, and there was some discussion about whether we should attempt to standardize assessment policies throughout the school. However, after Kate described how her experience in a prior district had been a negative one for the entire staff, we decided that we did not want to risk negatively impacting our school's culture by pursuing such a drastic change. The discussion turned to "catching" students *before* the test—before they failed an assessment and needed to retake it. We then pondered whether homework could be a means to evaluate student understanding—perhaps giving feedback and no grade—and considered if we could make homework meaningful while at the same time reducing the grading burden on teachers. After discussing that the most effective purpose of homework is practice (O'Connor, 2011; O'Connor & Wormeli, 2011; Wormeli, 2006; Wormeli, 2011), we explored how we could appropriately use formative assessments—assessments of learning (Guskey, 2011; Guskey, 2017; Guskey, 2020a; Guskey & Anderman, 2013).

The Value of Formative Assessments

Because so much of our conversations skirted the topic of student motivation, I shared the idea that student motivation can be improved through the use of formative self-assessments

(Jensen, 2009). As a result, the team elected to focus on formative assessments; specifically, to develop a student self-assessment in the form of an exit slip, or "exit pass" (Wiliam, 2011) which is a brief set of questions or a response from students during "a natural break in the instruction; the teacher then has time to read through the students' responses and decide what to do next" (Wiliam, 2011, p. 92). By focusing their intervention on adapting their instruction based upon feedback from the students' self-assessment, the team was ultimately achieving multiple tasks in one. First, they had the potential to increase student motivation by garnering exactly what the students needed to be successful (Jensen, 2009; Jensen & Snider, 2013). Second, they were giving students valuable feedback to help them determine their level of mastery and progress to proficiency and achieve a seamless system in which students could communicate their learning needs to the teacher and the teacher would respond with a variety of instructional approaches (Wiliam, 2011).

Feedback

Next, this practice allowed students to reflect upon their own growth and self-assess their work, mirroring Hattie and Yates' (2014) assertion on the power of feedback: "Feedback is powerful if students know (a) what success looks like, (b) appreciate it is aimed at reducing the gap between where they are and where they need to be, and (c) when it is focused on providing them information about where to next" (p. 67). Feedback is even more impactful when students engage in self-assessment and evaluate their own progress of where they are to where they need to be (Popham, 2001; Popham, 2028; Wiliam, 2011).

Finally, teachers were afforded the opportunity to enact one of the most important tools for student learning: informing instruction. Wiliam (2011) explains that regular assessment followed by actionable feedback is more impactful to student learning and has the potential to

increase student learning twice as fast as normal (Popham, 2001; Popham, 2018). Formative assessment can be viewed from two sides: the instructor's and the student's. On one hand, the student completes the assessment, and the teacher uses it to change their instruction. On the other hand, the teacher can individually give the student actionable feedback that the student can use to master the content. Both sides require communication and response—the teacher changes the instruction ("forms the instruction") to fit the student as a result of the student's communication or the student responds to the teacher's new instruction.

The Intervention: Student Exit Slips

The intervention we designed was fairly straightforward, and like many exit slips, it consisted of an opportunity for students to identify concepts they understood as well as material with which they still struggled. As the team considered how to gather student feedback, a quick Google search revealed a variety of options, but they shared the concern that students might write anything down just to finish the activity; therefore, we liked a question from a self-assessment example in Barry Lane's *But How do You Teach Writing?* (2008) in which his model asks for "proof" of the student's learning (p. 221). Our final exit slip consisted of four prompts: "What I learned/know from today's lesson (this unit)", "Evidence of what I learned today", "What I'm still working on", and "What do I need help with?" The team appreciated its simplicity and ability to edit for their courses or specific units.

For the first question, students responded with what they felt they learned for the lesson; the teachers found that students could tie their answers to the day's instructional goals or unit anchor standards. The "Evidence" was supposed to provide information for what they cited as "knowing." As a team, we thought that the evidence students provided in this question would help students give critical responses and avoid overconfidence. The prompt, "What I'm still

working on" asked students to respond with concepts of which they felt vague or unsure. The final question allowed for students to seek help from the teacher as well as guide the teachers in developing the next day's activities.

Teachers decided to use the exit slip in one course. Using it in only one course helped make planning easier for the teachers and allowed them to compare the responses (and teaching) while also removing the variables that multiple courses might bring. All three of the teachers gave the assessment in a course with two sections, and they also agreed to give the self-assessment at least two times a week using an electronic version which I created via a Google Form (Appendix B). The teachers took the first week of the intervention to explain to students why they were trying the self-assessment form, the expectations for it, and how to answer the questions.

The intervention for the students was one piece of the process. We arrived at the exit slip intervention because we recognized that the students were disconnected for their roles in the learning process. Research showed the potential for feedback and self-assessment to improve student learning (Popham, 2001; Popham, 2018; Wiliam, 2011) and even student motivation (Jensen, 2009; Jensen & Snider, 2013), but the teachers on the team were dubious. They were skeptical that the students would take it seriously or be able to accurately assess their own understanding. Creating an exit slip in a Google Form seemed to be the easiest, least-burdensome way by which to collect student feedback. Coincidentally, since the form was digital, it also gave the teachers more flexibility—they could modify it, review responses immediately, and send it out for students to complete at any time during their lesson or unit without disrupting the classroom activities.

Collecting Team Feedback

The other piece entailed garnering feedback from the team, both individually and collectively. Because the focus of this work was shifting perceptions on assessment and learning, I needed insight into the team's perceptions and if—or *how*—they had changed throughout this work. Again, the balance became how to gather this information without burdening them. Therefore, the team decided that they would complete a weekly reflection on how they were using the intervention and what they observed from the students as well as how they were using the student feedback to modify their instruction (see Appendix C).

The intervention we chose to test was carefully selected after reviewing the literature and discussing our options. Our goal was to improve our culture of assessment, so we settled on a simple exit slip, one that we hoped would increase communication between the teachers and their students and improve student attitudes toward assessment and learning. In the end, we wanted an intervention that would allow students to communicate their level of understanding and result in changes to the instruction. While the intervention itself was seen as relatively straightforward to implement, it also had the potential for the greatest impact to student learning and the teachers' instruction. It was for these reasons that the team chose to develop a student self-assessment as one step in our larger process to shift our culture.

CHAPTER V: METHODOLOGY

Mintrop (2018) describes a problem of practices as urgent, local, solvable within a given time frame, and connected to the district's broader strategic goals. Additionally, he asserts that solutions are actionable and practical. The issue of assessment practices falls clearly within Mintrop's criteria, and because poor assessment practices tend to disadvantage vulnerable students greater than others, it is also an issue of equity (Dudley-Marling & Dudley-Marling, 2015). While the team and I were proceeding with our intervention, I continued to plant the seeds for assessment work with the rest of the staff. In essence, I was practicing "Patient Urgency": investing in my small team and leading with patience while also giving them the opportunity to take risks. Additionally, I needed to maintain a sense of urgency because our work was important, with the potential to positively impact some of our most vulnerable populations (Mui, 2015). The work on student self-assessment was only one step towards shifting the culture within our building, but it was a step toward laying the foundation—building our capacity—for more change.

I followed the Carnegie Project on the Education Doctorate (CPED) Framework which is rooted in the practice where scholarly practitioners empower their staff to improve the professional practice within their organizations (CPED, 2022). With that goal in mind, I elected to identify a problem of practice and engage in a cycle of inquiry with a small team of teachers, which is one approach to the CPED framework (CPED, 2022). There are similarities between qualitative action research and the CPED Framework; for example, action research emerges or unfolds throughout its cycle of inquiry and researchers systematically collect and code their data, which is often in the form of qualitative evidence (Creswell & Creswell, 2018; Merriam, 1998). However, while traditional studies endeavor to identify a gap in research or knowledge, the

CPED Framework and problem of practice seeks to identify a gap in practice (Belzer & Ryan, 2013; Mintrop, 2018), with the goal to work with staff to improve their professional practice and embed this work within an organization's culture (CPED, 2022).

The focus of Empowering the staff and embedding this work within our culture fits well with the structure of participatory action research. At its core, this method centers on participatory action research which Coghlan and Brannick (2010) describe as "integrally collaborative and democratic" where the goal of this process is to not merely to conduct academic research with the plan of evaluating phenomenon but to work with and change the results (problems) that are identified (p. 6).

As a full participant on the team, I was deeply cognizant of my involvement in the entire process. Mintrop (2018) explains, "Action research embraces transformative leaders and places them in the center of the research endeavor. It draws its strength from the researchers' insider knowledge of the organization that engages in change" (p. 192). Mintrop also acknowledges the following:

> The strength of action research, namely, insider knowledge and an understanding through the direct experience of it, is also the weakness of action research. When researchers are also change agents, they play multiple roles as initiators, implementers, and evaluators all at the same time (p. 194).

He further explains that it is crucial "to guard against the emergence of advocacy bias" and to "search for competing explanations for observed patterns" (p. 194). Admittedly, no researcher is ever free from bias, so with the goal to reduce it, I collected qualitative data from three areas: weekly written teacher reflections, monthly in-person meetings, and field notes from informal meetings. As a result, I verified the evidence by triangulating the data. These pieces of

data were recorded and then coded. I coded the data by identifying similar phrases and use of synonyms within their responses (Creswell & Creswell, 2018; Merriam, 1998).

I was most interested in the team's perceptions of the cycle of inquiry as well as how they adjusted their instruction. The teachers implemented the student self-assessment in the form of an exit slip and used that feedback to inform their instruction. Collectively, the team agreed to give the exit slips at least twice a week. I collected descriptive qualitative data because it revealed the greatest insight into what the team hoped will lead to sustained change in our culture of assessment. Qualitative data offers the greatest opportunity to capture participant perspective and allowed me to observe them within their own environment. Specifically, these data came from the teachers' responses to their weekly reflections (Appendix C), our monthly meetings, and any informal conversation they had with me about assessment. The weekly reflections were collected using a Google Form, which assisted in sorting their remarks. I removed identifiable information from their responses and then coded their responses. Before our monthly meetings, I reviewed their responses on the weekly reflections, looking for emerging themes, problems, and inconsistencies.

My role as a transformative leader was to work with the team throughout this process, to empower them to solve a problem that we had identified within our organization; however, my role as a researcher was to endeavor to participate in this process while maintaining fidelity and reducing advocacy bias as much as possible. Therefore, I clarified my interpretations of the data (member checking) during our monthly meetings and used the same questions from their weekly reflection to drive their conversations. I did not speak much during these meetings, electing to ask the reflection questions and only interject when I needed clarification or answer a question they asked me directly.

Teacher Use of Student Exit Slips

When teachers began integrating the student self-assessment in the form of the exit slips, the work shifted from exploring mindsets and assessment practices to testing our intervention. But there were variations on how the intervention was used. For example, the team determined that they did not have to implement the slips in exactly the same way, and as a result, all three gave the assessment at different times during their instruction. There were other differences as well. Two of the teachers modified the assessment to have students develop questions relating to the content. One of the team members adjusted when she gave the assessment several times before finally settling on one that worked for her students. The teachers described these adaptations in their weekly reflections and during our monthly meetings.

Teacher Reflections

I did not peruse the student responses as my primarily interested was in the team's development of the intervention and their perception of the process itself and whether it would lead to real-time changes to instruction as a response to the students' self-assessment. Therefore, to gather regular feedback from them, we developed a teacher reflection that individual members completed each week. Echoing the student self-assessment, their reflection form was also short, and contained three prompts: 1) "Reflection (What's going well, what's not going well, surprises, etc.)", 2) "How have you adjusted your instruction in response to student feedback?" and 3) "Other things to note" (see Appendix C). They entered their feedback into a Google Form, too, which I sent out weekly as a gentle reminder to complete.

Meetings

During our regular lunch meetings, I asked questions, took notes, and made careful observations while the teachers ate and talked (Creswell & Creswell, 2018; Merriam, 1998).

Oftentimes, I would engage in the conversation as well when questions arose about best practices or assessment literacy by referring to current research and sharing my own experiences with assessment. Sharing bits of research had become a habit when we worked through our larger process of identifying our problem of practice and developing our intervention. The team of teachers had revealed their struggle with wanting to try out new assessment approaches but not having the time to research and "read up" on them. It made sense, then, for me to share materials relating to the questions they brought posed.

Each week, I reviewed their feedback, looking for common themes among their responses as well as identifying issues that needed to be worked out in person or as a team (Creswell & Creswell, 2018; Merriam, 1998). We met formally one- to two times per month, and two of the team members often sought me out for informal discussions or commentary. However, these regular meetings felt anything but "formal." They were very relaxed occasions where conversation and ideas flowed, with the team often exclaiming surprise when the bell indicated that our time was up. In reality, they were the epitome of good professional development in which we learned from each other, exchanged ideas and engaged in valuable discourse on timely issues (Muñoz & Guskey, 2015; O'Connor et al., 2018). We gathered one- to two times per month during lunch, since the team had selected this time to meet as a group. Our final meetings took place a day before Spring Break, and each team member elected to meet with me individually. We had planned to meet our final time as a group, but when some indicated that they wanted to have lunch with their colleagues due to the approaching break, I gave them the option to have individual meetings at other times during the day. In the individual meetings, I asked the same reflection questions and took notes as usual.

Reviewing Teacher Feedback

Before each team meeting, I reviewed their responses from the teacher feedback form. I used the same questions from the team's form to guide our in-person conversations. Reviewing the team's feedback was important because it served to provide consistency and focus in our conversations; the discussions helped expand upon the teacher's responses and provided points of discourse. During the meetings, I took notes, usually sitting at my desk while the rest of the team ate their meals around the table in my office. The most valuable approach I found was to simply allow them to talk, and while they discussed, shared, and pondered, I typed my notes, occasionally answering their questions or jotting a written note to myself whenever I needed to look up a resource or detail.

After the meetings I reviewed the notes, monitoring for any themes I saw emerging from the team's discussion. To respect the teachers' time, I limited our meetings to school hours, with the plan to use their preparation period and lunch, or other times that they identified as convenient for them. Venables (2018) asserts that to create a high-performing, authentic PLC, facilitators must validate members' voice and "guide the discourse without evaluation or admonishment" (p. 13). Therefore, guiding the meetings as well as determining how often we met was a balance between collecting evidence and respecting the teachers' time and autonomy to conduct the intervention.

Winnowing the Data

I collected data in the form of qualitative responses from the teacher reflections and my notes during our lunch discussions. I had already looked through the reflection responses several times as they were completed by the teachers and initially thought that I would peruse the whole of it for themes. This path proved to be disorganized and muddy. I needed a more

methodological approach to what Creswell and Creswell (2018) call "winnowing the data" (p. 192). Therefore, I finally combed the data sequentially, by taking a more grounded approach and reviewing the data week-by-week and participant by participant, comparing responses and looking for links among the concepts and categories (Merriam, 1998; Creswell & Creswell, 2018). Following Mintrop's (2018) approach as team members being co-designers of the process, we reviewed the themes that emerged as well as discussed any quantitative data the team developed and use the findings to formulate next steps in our process to create a more equitable learning environment for our students.

I noted the similarities from each week before looking at the overarching themes that had developed. A few broad themes had surfaced, but I soon noticed that the themes had evolved over time, therefore it made the most sense to present the data chronologically. If we were to apply our findings to working with the full team, it was crucial to take into account this change over time. Early themes focused on student input quality and honesty; later, teachers focused on how to get better, useful feedback. The team was asked to modify their instruction throughout the process in response to student feedback; as a result, other themes within teachers' instruction adjustments emerged, including communication and growing assessment confidence, and ultimately, a process for changing in their teaching that was specific to each teacher.

Trustworthiness and Ethical Considerations

Having developed an intervention and collected data together, I felt confident in my team and believed that we had worked very well together; however, throughout this process, I was mindful that there could be trustworthiness and ethical considerations at play. Because I evaluate all three of the teachers on the team, I was cognizant that the teachers may feel undue pressure to participate in the study. When I asked them if they would like to participate, I chose

staff with whom I already had good relationships. Ones that I felt that we had already established a level of trust and that they could *and would* offer me critical feedback. For example, Kate is a member of our high school Leadership Team, and she frequently gives her opinions on the directions for the team. Additionally, she has offered candid feedback and advice on some of my proposals and directions for our Leadership Team goals. Charlie and Anne, too, have sought me out to talk about different issues and have been unafraid to disagree with me as well.

The notion of trust is something that I take very seriously and practice daily: I thank people for giving me feedback or disagreeing with me; after all, I believe that we tend to make better decisions when we think critically about the issue at hand. Additionally, I have an open-door policy, and I welcome staff to drop-in and "talk shop"; they know that my office is a safe space to vent, as well. Too, I am no longer a new principal, having finished four years in the district, so I believe my staff know me pretty well by now, and I know them. Despite the fact that we all know each other, I wanted to ensure that the team did not feel pressured to participate, so I reminded the team verbally and in writing that their participation in the study was not mandatory and that there would be no penalty for not participating in it, dropping out of it, or not completing the self-reflections.

Participating in this work required effort, collaboration, and trust from my small team. I relied upon qualitative methods to review the resulting data, which were collected from teacher reflections, in-person meetings, and informal conversations. Upon reviewing the data, I first identified overarching themes then moved on to combing the data sequentially, reviewing it week-by-week and participant by participant, comparing responses and looking for links among the concepts and categories. This process revealed themes centering on student input quality and

honesty, but more importantly, a growing confidence from the teachers as they were able to use the feedback to improve and modify their instruction.

CHAPTER VI: IMPLEMENTING THE INTERVENTION

To begin the process, I had identified our grading practices as a potential problem of practice. After observing my teachers and students for a year, I sought out feedback from a small group of instructors who had expressed interest in improving their assessment methods. We investigated the issue and agreed that it was systemic and then worked to defined a problem of practice. Throughout these early steps, we resisted the impulse to immediately develop solutions—a habit that is a common occurrence in schools (Bryk, 2009; Mintrop, 2018). Once the problem was defined, we turned to research to help us design an intervention. Again, we resisted jumping to solutions, and I shared literature on assessment while also taking our time to evaluate the approaches to test.

Our goal was to gather authentic feedback from a student self-assessment that would inform instruction without overwhelming the teachers; we felt that a successful intervention could not only glean valuable feedback from students but also had the potential to inspire other teachers to improve their assessment habits and make real-time adjustments to instruction. We then tested the intervention by issuing it to students on a regular basis, teachers modified their instruction based upon the student feedback. The team evaluated the intervention through team meetings and teacher reflections. The team meetings were used for discussing the effectiveness of the intervention itself and how the instructors had responded to student needs as well as time for trouble-shooting, offering support, and sharing ideas for adjusting instruction. Collectively, the team experienced a genuine collaboration—"intellectual leadership" (Mintrop, 2018)—and found value in the intervention, and while their responses showed growth in its use, the teacher reflections were not always completed regularly by the team members.

Team Meetings

As the team and I moved ahead and implemented our initial cycle of inquiry, we met monthly to discuss the work. These meetings served multiple purposes. On one hand, the team evaluated whether we were implementing the interventions with fidelity while also allowing me to take notes and monitor our progress. Collectively, we were engaging in "developing our habits of as *improvers*: learning, influencing, resilience, creativity, systems thinking" (Perry et al., 2020, p. 16). While our goals were to ultimately influence the culture of the entire school, we were aware that our work was only one component within the greater system of change that must continue after our cycle was complete. We also used this time to deal with immediate issues like trouble-shoot problems, offer support, and share ideas for improving or how they were adapting their instruction.

Trouble-shooting Problems

The team implemented the student self-assessment intervention at the start of the second semester (see Appendix B). Early on, issues centered on the team brought up concerns about the students not completing the intervention, vague responses, and when to administer the assessment itself. Even though only a few students grumbled about taking completing the self-assessment, most completed it. The teachers, on the other hand, felt that the student responses were vague. The team vacillated between thinking the issue was tied to motivation or to students not understanding their instructors' expectations, so teachers started with ensuring that the students understood the task and reviewed their expectations, what a good response looked like, and how the assessment would be used.

This approach improved the vague responses, but a few students continued to not complete it. The team decided to approach this behavior by offering the students individual

support (meeting to ensure that the student understood the expectation and asking if they needed help) and simply persisting in having them complete it. By the end of the study, this issue became less of a concern, especially as students saw the teacher modifying their instruction to suit the students' responses.

Anne, Kate, and Charlie differentiated when they gave the student self-assessment based on what they felt worked best for their instruction. While all of the students completed the three-question Google Form two- to three times a week, when they took the self-reflection varied teacher by teacher. For example, in Kate's class, she issued it in the middle of the period after a mini lesson while Anne gave her students the form after a unit of instruction. They determined that they did not have to be consistent in when they gave the self-assessment as teaching style and class structure were better determiners of when to give it.

Offer Support

Problem solving like the ones above were good examples of the collegiality that had developed among the team members. The times we met were pleasant ones in which they ate and talked, sharing ideas and offering support. It is not that they were fearful in implementing the intervention, but they were uncertain and did not quite believe that the intervention would be successful. Therefore, the times we met were important for offering support in the forms of reassuring and encouraging each other as they navigated uncertain territory. Charlie was the least experienced of the team and remarked that he did not have a great deal of assessment practices from which to draw. Kate and Anne reassured him when he brought up new ideas (to him) for grading students (e.g., assigning a project instead of giving a test, using a rubric for written responses, etc.).

Adapting the Intervention

Similar to Charlie's concerns about his lack of experience, Kate was the most insecure about the student self-assessment. She wanted to make sure she was "doing it right." Anne modified the exit slip to better suit her classroom's needs, and that change prompted Kate to experiment with it as well (e.g., having students write a question if they responded that they did not struggle with anything, determining at what point in the lesson/unit she should give the slip, etc.).

These meetings also impacted the most important element of the self-assessment: How teachers adjusted their instruction. Charlie shared that he used short videos he found on the internet to reteach the concepts students stated were confusing. Shortly thereafter, Kate answered that she had tried this approach as well. Anne used student responses to generate essential questions that she shared in the student online classroom; Kate liked this idea and used it, too. In essence, these meetings were an opportunity to share what worked and build the teachers' proverbial instructional toolbox.

Intellectual Leadership

We experienced a genuine collaboration through our regular meetings. Following the values of transformative leadership, I worked to empower the team members to navigate through the process of implementing changes in assessment (Shields, 2020). Some of this "empowerment" simply meant giving them permission to try something or adjust the exit slip to fit their needs. Other times, empowering the teachers entailed affirming or challenging their instructional or assessment beliefs based upon what research indicated as best practice. Ultimately, I planned to have them share their experience with staff or lead in-service discussions when future opportunities arose. At the very least, I hoped they would have the

confidence to make assessment decisions when they surfaced or speak to that topic when others brought it up. In addition to providing opportunities for leadership—with the hopes for transforming our culture, empowering the team also allowed me to tap into the professional capital of my members themselves. Hargreaves and Fullan (2012) explain that complex shifts in culture are achieved through the following:

> Shared experiences, trusting relationships, and personal and social responsibility, as well as transparency. What pulls people in, teachers all the more so, is doing important work with committed and excited colleagues and leaders engaged in activities that require creativity to solve complex problems and that make a real difference (p. 151).

The ends described by Hargreaves and Fullan are exactly the kind of goals we were working to accomplish in our group. The team's early discussions demonstrated what Mintrop (2018) refers to as "intellectual leadership" where participants are considering "not only what works, but also what they value. They have the courage to use their own judgment and to resist organizational authority when it is used to produce inequity and injustice" (p.88). This sort of intellectual leadership allowed the team to work the complex problem of assessment.

Teacher Responses to their Reflections

Our in-person meetings took place only once a month, so it was necessary to gather feedback from the teachers while they were gathering the student self-assessments and adjusting their instruction. Therefore, the team had their own reflections to complete, which they did so inconsistently. I recognize that we were adding "one more thing" to their already full instructional plates, and I also understand that they could have concerns that participating or not participating in the research could paint them in a negative light as an employee. As their principal and advisor, I did not want to damage the relationships I had built with each one; I

valued them as members of this team. Therefore, I tread very lightly when requesting the team to complete their reflections. Each week, I resent the Google Form for their reflection with a quick note asking them to complete it. They were voluntary participants, however, who could leave the study at any time and who could complete (or not complete) the reflection. The fact that none of the participants consistently completed the reflection likely speaks to the fact that they took me on my word that there would be no penalties for electing to not participate in all of the components of the study.

The intervention for the students and the reflection for the teachers were both relatively straightforward and easy to implement. Each took only a few minutes to complete. And as we implemented the intervention, we ran into relatively few problems. However, the team used some of our time during our monthly meetings to trouble-shoot issues, offer each other support, and share ideas for modifying their instruction—all good indicators of their growing collegiality. A final issue that I encountered, however, were the teacher's inconsistency in completing the reflections. Despite regular reminders, no teacher completed the reflection every week during the cycle of inquiry. Therefore, as we move forward, we will have to determine whether the reflection is a necessary component of the process. If it is, then we may have to investigate other options for collecting this information.

CHAPTER VII: EVALUATING THE INTERVENTION

It became evident that it was time to compile the data when the team's responses in both the weekly reflections and our lunches became repetitious. I used the information I collected from our in-person meetings to triangulate the data the team shared in their weekly self-reflections. An important component to triangulating this data was observing the correlation between their written comments matched their verbal and nonverbal responses in our meetings (Creswell & Creswell, 2018; Merriam, 1998). Seven themes emerged from the team's responses. Chronologically, teachers' expressed concerns with honesty and quality of student feedback and the timing of when to give the exit slip. The responses then shifted to gathering usable feedback and implementing individual student check-ins. Finally, the team demonstrated a growing confidence and flexibility in using the exit slip, and all resulting adjusting instruction to suit student needs.

Themes

After identifying commonalities from each week, I then looked at the overall themes that had emerged. It made the most sense to present the data chronologically because while overall themes had emerged, I realized that the themes had developed over time. This change over time was important if we were to translate our results to working with the entire staff. The ultimate goal of this work was to fill gaps in practice among the entire staff, so identifying whether—and if so, *how*—this intervention and the teachers' perceptions of assessment and of this professional development would be crucial to our success when introducing the work to the rest of the staff later. The following themes had emerged: early themes centered on honesty and quality of student feedback; later, teachers grew to focusing on how they could garner "better" feedback that was useful to both them and the students. Throughout the process, the team was asked to

adjust their instruction in response to the students' feedback; therefore, within how teachers adjusted instruction, other themes emerged, including communication and a growing confidence in assessment and ultimately, a process for changing in how they taught that was unique to each teacher. I then investigated how the teachers modified their instruction as a result of the students' responses.

For all three of the teachers on this team, I am their direct supervisor and evaluator. Because of those reasons, I was cognizant that there could be other processes at work here. For example, Charlie was very enthusiastic about integrating the intervention; however, since he was one of the newest members of the staff, I had to discern whether this enthusiasm could be motivated by a desire to fit in or unconscious effort to appease me. The fact that he often added an appropriately critical eye to the process, assuaged some of my concerns. Despite his enthusiasm, he often expressed exasperation with the students who, in his opinion, "didn't care." Kate remarked that she had received "vague responses," and Anne described how some students had rushed so quickly through the first feedback form that they had filled out the exit slip on the wrong topic. In response to Kate's and Anne's frustrations, Charlie questioned whether the students would ever take the activity seriously. When Kate struggled with the appropriate time to give the assessment, Charlie encouraged that the self-assessment did not have to come at the end of the period; it could take place after her mini-lesson or even the beginning of the next time the class met.

On the other end of the spectrum, Anne was the most reserved in our conversations. She tends to be one of the more soft-spoken among our staff, so this phenomenon in itself was not alarming. Despite her reserve, she ended up adapting the intervention the most and modified it to fit her classroom needs. By the eighth week of using the assessment, Anne had led students in

her class to use the form to write questions related to the anchor standards for the unit. Kate usually provided the most critique in creating the self-assessment as well as responding with deep reflections upon how she used the student feedback. Each team member brought a unique perspective to the development and implementation of the intervention, and their responses lent to consistent data.

Table 2

Summary of Teacher Responses

Teacher Responses		
January		
Charlie	Kate	Anne
• Answering "truthfully" • Students don't give honest feedback • "I stressed the importance,...so most students have been honest" • Responses not in-depth, lacking specifics	• Some give specifics • Many responses are vague • Difficulty in determining the best time to give the exit slip • End of the period "isn't the best time to give it"	• Some students rush through the form • Students filled out the slip on the wrong topic • Determining when to give the exit slip
	Adjusting Instruction	
• Review the concepts • Find a video	• Had not changed whole group instruction, yet	• Had not changed whole group instruction, yet

(Table Continues)

(Table Continued)

Teacher Responses		
February		
Charlie	**Kate**	**Anne**
• Student responses are "great" when they are honest • Some students seem to be using the exit slip to their advantage • Some students are "honest with me about when they don't understand" • Frustration with reviewing the purpose of the exit slip	• Starting to get better, more thoughtful responses • Several students continue to give low level responses • Working to meet with students individually during work time • Trouble-shooting when to give the exit slip	• Use of assessment as a starting point for specific questions at the beginning of class • Review purpose of the exit slip • 'No one else is seeing this beside me.'
	Adjusting Instruction	
• Whole-group Reteaching	• Meeting with Individual Students • Start class with a discussion of the topics they identified • Short video	• Meeting one-on-one with individual students or small groups • Presenting student questions to the entire group
March		
Charlie	**Kate**	**Anne**
• Students filling out exit slip "mostly honestly" • Told students it would be taken for a grade but didn't assign it points • Students don't take the opportunity to improve their grades • "A 'D' is okay."	• Students gave specifics that the teacher could use as a starting point	• Using student responses to "see their level of understanding • Students create a question that the teacher uses for the whole class • Realization that "There has to be a certain level of understanding to be able to formulate questions"

(Table Continues)

(Table Continued)

Teacher Responses		
March		
Charlie	Kate	Anne
	Adjusting Instruction	
• Whole-group review • Show Video	• Students create a review question for their classmates • Using the feedback to reteach the entire group	• Post student-generated questions in the online classroom • Discuss these questions within small group • All students record information from the class discussion • Student answer an essential question and "prove" it

Note. Table 2 summarizes the individual responses from the team of teachers and is organized chronologically.

Honesty

While I anticipated for the teachers' comfort with issuing the form and knowledge of assessment to evolve over the course of the study, one theme continued to be referenced throughout all their feedback: the notion of "honesty." Early reflections and discussions centered on student honesty in accurately evaluating their learning as well as the quality of the feedback. Charlie especially referred to the students needing to be "honest" in their responses. He said, "The only issue I have been having are [sic] when students don't give honest feedback on what they need help with." The idea that students were being dishonest in their self-assessment was concerning—insinuating that they knew better but chose to not share the information with the teachers. This observation was echoed later when Anne seemingly begged her class to complete the form thoroughly. She reminded them, "No one else is seeing this beside me." Both Charlie and Anne hinted at a looming issue within their classrooms: a concern

that students did not want to share their misunderstandings out of fear of other students seeing them. If students were withholding accurate assessments of their learning, perhaps the teachers had not established a positive learning environment.

The topic of students being honest in their feedback was one that we discussed during our lunches. The written responses from the team made me wonder whether this notion was tied to the quality of feedback or whether it was rooted in a deficit mindset. If the problem was quality and the need for feedback to be specific to best guide their instruction, the concern was valid. However, if the issue centered on the idea that students were being dishonest, we faced a greater challenge within our team—how we perceived our students. Ultimately, it ended up being a little of both and more. Kate remarked, "For many students, I'm getting vague responses that indicate they feel confident about the lesson" when in reality, their understanding was still relatively basic. We concluded that the overconfidence students were demonstrating likely lay in the fact that their understanding of the new material was still topical—they did not comprehend what mastery looked like for the content/concept, so they could not gauge/indicate their skill level. The teachers needed to make the standards explicit so the students could use them in meaningful ways, specifically to guide their self-assessments.

Charlie's outlook, on the other hand, highlighted the importance of modeling appropriate feedback. His reference to honesty really centered on inaccurate feedback from the students, but the true issue was that they did not know *how* to formulate where their misunderstandings lay. As he described the students' shallow responses that lacked specific examples, he made the realization: "I didn't teach to their understanding, so I need to find something else to help them." Unfortunately, our structure of gaming for points had likely exacerbated a deeper problem. Students were so used to complying with the shallow components of our system, that deeper

learning activities were foreign to them, and this point supports the notion that grades really do not communicate a student's level of learning. Conversely, Charlie's response, too, demonstrated a superficial level of understanding how to support students when they do not grasp the concepts the first time a teacher presents them.

Quality

Kate and Anne expressed apprehension early on regarding feedback quality. One described how students rushed through the form while the other explained that students overestimated their mastery of the concepts they had just learned. Both teachers indirectly defined "quality" of student feedback as vague or poor and wanted specific information from the students. For example, Kate explained that in response to the question, "What do you need help with?" Some students answered, "Nothing" or "All of it." In our lunch discussion, we debated whether responses like these were laziness or demonstrated a lack of comprehension—that they were not aware of what they did not understand. In effort to assume positive intent, we went with the latter. All three teachers had explained the how to fill out the self-assessment in the weeks before actually giving the form. In one of our lunches, Kate explained, "I told them, this feedback is really for me, to tell me what I need to do better to help you." Similarly, Anne reminded students, "I'm the only one who sees this, so don't be afraid to tell me what you don't get."

Anne remarked in her weekly reflection that she had reviewed the expectations for filling out the exit slip and hoped that setting aside five- to eight minutes for students to fill out the form would result in "getting more thoughtful responses" from them. Kate, on the other hand, praised the class when some gave "specific feedback they needed help with" and shared them with the students to model what "good looked like." Clearly, the students' and the teachers' struggles

indicated that this type of activity was foreign to both parties; learning how to use the self-assessment effectively took time from both the teachers and the students. Both Kate and Anne found that when they gave examples by which to model, student responses improved.

When to give the Assessment

The perception that students were withholding information or giving inaccurate and poor responses resulted in all three of the teachers identifying the need to reiterate the rationale for the self-assessment activity as well as review their expectations for filling it out appropriately. In addition to these perceptions, Kate and Anne found that they struggled with determining an appropriate time to give the self-assessment, which they concluded played a role in the quality of the student responses. These early procedural hurdles were easy to address with students through quick reminders of why the team was using the intervention as well as showing examples of good responses.

They also began trying out different times in the instruction to give the assessment. The teachers discussed the appropriate time to give the student self-assessment. Charlie was very pragmatic in his approach and felt that the assessment fit best at the end of the period. He typically scheduled it for specific days of the week. In essence, he used it like a dipstick, or a temperature check, monitoring student understanding and combing their responses for what he needed to reteach. Kate and Anne considered whether the quality of student responses was impacted by when the assessment was given. Anne had attempted to give the assessment after entire units of instruction. Her goal was to assess their level of understanding at the end of her instruction and then use the feedback to reteach and review before the summative or unit test. Kate tried a variety of approaches, first starting with giving it at the end of the period and then shifting it to the middle of the period, after a mini-lesson. She even used it to begin the class as a

form of review. In the end, only Charlie remained consistent with his schedule of giving the self-assessment. This testing of when to give the students the self-assessment slip indicated that the teachers were not fully grasping the ultimate goal of the intervention itself, which was to impact their instruction, to cause a change in *how* they were teaching that would meet students' needs.

Timing

A final finding about the student self-assessment centered on timing. Each team member indicated that they struggled to establish the rationale of the intervention with students—something they felt would have lessened if they had started the school year using it. Charlie laughed that the students had realized the three teachers were using the self-assessment form and accused the three teachers of conspiring to develop it over winter break. "I wish we had implemented it sooner, at the beginning of the year. Made it standard practice. Then there would have been no change for the students. I could see the patterns for the entire year instead of just one semester."

Charlie's point was a valid one, as Anne agreed that starting the intervention at the beginning of the year would have helped establish it as habit, as part of the teachers' normal practice instead of "this suspicious change in the middle of the year that three teachers were trying." The team also felt that the longer we use tools like the self-assessment with students, the greater the opportunity to shift the entire staff of teachers' focus on learning. As Charlie asserted, "I feel the more we do these exit slips, [the more] they will help with our conversations on learning collectively. The more we do these, the more we'll all focus on learning."

Usable Feedback

By the third teacher reflection (about five weeks into the semester), a shift seemed to take place, when all three teachers reported that students were giving better feedback. Charlie remarked, “Students are really seeming to use this to their advantage and are honest with me about what they don’t understand.” Anne not only found the feedback to be higher quality, but she was the first to shift how she used the students’ responses: “I am able to use the assessment as a starting point for specific questions to ask at the beginning of the class. I use some questions the students have left for me on the [summative/unit] assessment.” Kate added, “I feel like we’re slowly starting to shift the focus toward ‘Do I understand?’ rather than ‘Will I get an “A”?’” When Kate noticed that students were “quickly filling out the self-assessment saying they knew what they were doing and did not need help,” she decided to individually check students’ work. She also made personal connections with them by “talking to them about what they did well and what they missed.” The teacher discussions, too, changed from a focus on how evaluating the student responses to actually *using* the feedback to inform their instruction.

At the same time Kate was making personal connections with her students, Anne, too, met with students on a more individual basis. Her classes were arranged into pods, so she met with each group to review assignment questions they did not have correct. Charlie did not meet with individual students or groups of students; instead, he reviewed student responses and stated, “If there is a topic that is asked about more than once, I will either reteach the students that portion of the lesson or find an alternative way for them to better understand the content.” The perception that students were giving “good” feedback was a significant shift so early into our intervention; therefore, we explored when we met in person what “good” or “better” student feedback looked like to them.

Interestingly, our discussion on the concept of response quality began with remarks like "not vague" and "specific" as well as having students indicating their level of progress to the goal. But what the teachers were really looking for was specific information that *they* could use to reteach the concepts or find the students resources to explain the topics in different ways. For example, for a student to indicate that they had very limited understanding of the concepts at the beginning of the unit but now understand half of them shows growth, but to report that they did not understand why Naturalism is situated within the Realistic Movement in American Literature communicates to the instructor *how* to fill-in the gaps for that student—it gives them actionable feedback. After Anne told the students "No one else is seeing this besides me," students began to give her "a lot more specific information." She added the following:

> Why the change? I keep stressing the importance: 'I can't help you if you don't give me specifics. I'm using it to adjust my teaching and how I approach it.' I'm sure some of the kids who are seeing that if I put something down, and you adjusted, that reinforced using it [the self-assessment]. 'That worked! She changed; she reviewed the concepts.'

In essence, Anne found that her student feedback became transactional, a means of communication between her and her students. Ironically, this finding parallels the theory behind organizational change in that "in transactional dynamics, behavior is changed through a negotiated give-and-take. In transformational dynamics, new behavior emerges when leaders appeal to, and mobilize, employees' values or commitments" (Mintrop, 2018, p. 122). In our case, the leader was the instructor and her "employees" were the students. The change that took place was transactional in which students indicated their learning needs and the teachers modified their instruction to fit them. This give-and-take that had started to take place was a cycle of communication between the teachers and the student. Teachers received more usable

feedback from the students and the students experienced a response, a change in instruction from the teachers (Popham, 2018).

Mintrop (2018) asserts that changes like the one Anne experienced happens as a result of "drivers"; in our case, they were likely the beginnings of "a self-reinforcing dynamic" (p. 124). As I stated earlier, our staff genuinely want our students to do well, and positive changes like this one were motivating. When students gave them actionable feedback, the teachers were motivated to adjust their instruction. In this case, the change in feedback drove them to continue modifying their instruction for the better which resulted in better, more specific feedback from the students. The transactional change that we were beginning to see was a step toward achieving transformation change. However, achieving deeper, transformational change in which new (better) behaviors emerge would take time. And while there is no doubt that I observed some change in the team's view of assessment and students, this study was one component of the overall change we wanted to see among our staff.

Individual Check-ins

Throughout the study, pupil-to-teacher communication continued to improve—especially as students grew to understand the kinds of feedback their teachers needed to inform their instruction. The feedback eventually led teachers to meet with students individually. One week, Kate added that she was "getting better, more thoughtful responses from some students," and she then used those comments as a springboard for in-person conversations and adding check-ins "with individual students during work time to check over practice problems and help [them] with problems they're stuck on." Anne, too continued to meet with groups of students but by the end of the data collection had shifted to targeting individual student needs by having students keep a document of the essential questions the class had developed from their responses.

Confidence

At this same time, Charlie reported that the students were starting to notice that he had been adapting his instruction to their self-reflections and "how this can help them." While Anne had already modified the assessment to better fit her classroom, Kate followed suit, concluding that a name change would resolve her issue of *when* to give the form to students. "I changed the name from 'exit slip' to 'self-assessment' so that I can assign it whenever it makes sense in our day, whether that's the start of class, or at the end."

Admittedly, at first glance this shift was a logistical one; what was significant, though, was how Kate now *talked* about assessment and instruction. In early discussions, she posed questions, with the tone of her voice revealing uncertainty. She would ask a question to the team and her voice would trail off: "I'm struggling with when to give it. It doesn't seem to work when I give it at the end of the period…" However, by the end of the study, she did not seek out permission to change the name of the assessment from "exit slip" to "self-assessment," She changed the name in her class because she felt it would benefit her students and her instruction. And when she spoke, she made assertions instead of asking questions as her voice trailed off and waited for feedback.

I took the two teachers' initiative in modifying the self-assessment as an indicator of the team's growing confidence in implementing this intervention. Their confidence was also reflected in how they were responding to the student feedback by modifying their instruction. It appeared that the team had achieved a level of transactional change and were on the cusp of some transformational change as well in that student motivation—in their classrooms, at least—was improving (Mintrop, 2018; Wiliam, 2011; Jensen & Snider, 2013).

Adjusting Instruction

All of the work on the student self-assessment would be for naught if it did not result in a change in instruction. Formative assessment is a *process*, one that elicits evidence from the students to alter ("form") how the educator teaches (Hattie, 2012; Popham, 2001; Popham, 2018; Wiliam, 2011). Achieving a seamless system in which students are aware of their own understanding of the material, self-assess their level of mastery, and identify their instructional needs followed by a response in instruction takes time, and we did not see evidence of this process until we were nearing the end of the study. Because the instructional change in response to the students' need is so important, in the following paragraphs, I describe how Charlie, Kate, and Anne developed over the course of this cycle of inquiry.

Charlie

Early responses to how they were modifying instruction indicated that the team members typically limited their modifications to reviewing "missed" concepts with the whole class. Essentially, they were merely repeating how they had taught the previous day. During our in-person meeting, Charlie described feeling at a loss for what else he could do to reteach—to *teach differently*—and admitted that he sometimes spoke fast, so he needed to find a way to ensure that all students could access the key concepts of his lessons. This shift in ability demonstrated a marked change in Charlie as he moved towards transformational practices from his original transaction approach. He found that one of his best—and easiest—tools was to show a video. "It literally takes me a couple of minutes to find a three-minute video that I can show the entire class or share with the kids who didn't get it." Charlie continued the practice of researching and supplying videos to his students throughout the study, and by the end, he had amassed a small cache of resources to use year after year. One morning, he dropped by my office and mentioned

how many resources he had gathered as a result of our work. He remarked, “You know, this [modifying instruction] wasn’t that hard. I just needed to find resources that could, you know, fill-in the gaps. Videos and websites. Now I have a whole bunch that I can keep using and adding to. It just makes sense.”

Charlie was always one of the first to arrive at school in the morning, and he used that time to prepare for the day. Part of his practice was to give review the self-assessment at that time and find resources to share: “If a topic is a common occurrence, I will start the next class with a review discussion of the lesson, as well as the topic that the student may possibly be misunderstanding.”

Kate

While Charlie continued with his practice of examining the students’ responses, identifying common responses from students, reviewing gaps in his instruction, and supplying students with a brief video throughout the intervention, Kate and Anne strove to differentiate how they approached meeting students’ instructional needs. For example, Kate moved from whole-class review to working on specific concepts with individual students then pulling back to use examples to apply the concepts:

> I have been able to check in with individual students during work time and help them with concepts they’re stuck on. After seeing responses from my 7th hour class (who all said they learned the topic but then didn't have much for evidence of learning) I decided to start the next day with a real world example that we built together. They worked together to think through the problem.

Kate utilized the response from the student self-assessments to initiate change in her instruction

but then advanced that change through individual conversations with students. Students saw immediate adjustment in instruction as a result of their feedback.

Anne

Anne, on the other hand, scheduled class time to meet with small groups. By the end of the data collection, she had changed the self-assessment from merely seeking feedback to having students write questions that aligned to the anchor standards for the unit. She used those questions in a variety of ways, including having students keep their own record of the questions and working within their groups to answer them. She also selected some of the questions to use on the unit tests. Finally, she rotated among groups, selectively asking the questions. "I pick specific questions that I think a specific student or group might not know and ask their table. Then they have to give me an answer. If they don't know it, I give them time to research and talk among their group then come back to them to see what their answer is. This sort of helps take the pressure off one kid not knowing the answer since the table has to work through it together."

Thanks to our small class sizes, Kate's and Anne's tactics were achievable where this approach might be challenging in a larger district. Regardless of the class size, though, individualized instruction for each student can be tiresome and time-consuming. Additionally, these two teachers wanted more than a single "rinse and repeat" activity; it was clear that they wanted a variety of tactics from which to draw. Anne's first approach was to have students create questions that she then presented to the class "for whole group discussion as review the next day." While at one of our lunches, Anne described her rationale for having students write questions, and explained that she was trying to release more of the responsibility for their learning to the students. Realizing that student-created questions allowed additional

opportunities to demonstrate understanding, Kate remarked, “I like this idea! The ability to formulate a question on the material already indicates a higher level of understanding.” Later, as Anne continued to explore how she used student responses, she added, “I am thinking about having students keep a doc where they copy/paste the daily questions and record the class responses.” By the end of the study, Anne evolved to giving students the goals for each lesson and had students develop questions for those goals based upon what they had learned.

Anne’s use of questioning emboldened Kate to try a similar approach as well. She wrote the following:

> The last question of the self-reflection said, ‘What do you need help with?’ and I added, ‘If nothing, create a review question from this lesson for your classmates.’ After adding this instruction to the last question of their self-reflection, I've seen more students actually writing down a topic that they need help with. Adding a topic allows me to do a whole group review which benefits all students.

Kate and Anne had pushed themselves to adapt the assessment itself to fit their instructional needs. Charlie, being the least experienced, had found a means to help him manage the day-to-day instruction while also adding another tool to his proverbial instructional tool box. Despite the variety in teaching experience and background knowledge, each of the team members had made the self-assessment valuable for their classroom.

While all three of the teachers had at some point expressed skepticism in whether the students would take the self-assessment seriously—with one teacher even telling the students it would be taken for a grade, but then surreptitiously not entering any points in the grade book, they were all collectively amazed that students finally bought into the self-assessment. “I was really surprised,” remarked Charlie. “Just how responsive the students were since it wasn’t

worth anything. No homework assignment, just 'do the reflection so we both know what we need to work on.' I'm looking at using it for every class next year." Kate added, "The reflection has moved from 'Hey, what's this?' to really using this with students and using it to review with students. Now I have a process that takes no more than a minute to redo the self-assessment. I plan to continue to use it with students. It's helped me as a teacher to know what kids got and didn't get."

But students did not get to that place without some serious shifts in thinking. Kate explained:

> The struggle was getting kids to buy-in in the first place. The first question was 'Is it for a grade?' I responded 'no,' but it's for me to help you learn. What I need to work on. What I didn't do well. After a couple of kids gave me good feedback, it helped them all see what good feedback looked like and could do for them as a class.

Anne, however, did not find the process as quick or straightforward. Perhaps because she modified the assessment the most—or perhaps because she modified it the most often—she found the self-assessment somewhat time-consuming. She also reported that she felt compelled to give students class time to complete it since they had told her that they would not work on it outside of class. Despite these negatives, Anne appreciated the opportunity it gave her to see which students needed extra help.

By the end of the quarter, all three teachers had experienced a shift in how they used the student self-assessments. In their view, the activity had initially begun as a tactical task (Heifetz, 2009), one of which they were highly skeptical. However, each one had moved from simply implementing a task to seeing the value in student feedback and changing how they taught to fit student needs.

Final Surprises

A few final results that had arisen from our discussions bear review. The team brought up concerning student attitudes toward grades and their behaviors toward actively participating in their learning. Early on, when teachers had expressed concern with the students being honest in reporting their understanding of the material, they were surprised when some of the students were blatant about their disregard for their own learning. Other surprises lay in how the teachers came to value the process of the cycle of inquiry itself, with one describing it “like a book study without the book.” Finally, we realized that perhaps another gap in practice could be placed on the shoulders of teacher preparation programs, when two of our less experienced team members reported that their programs had done little build their knowledge of assessments.

Student Behaviors and Beliefs

Putting more of the onus on students by having them write questions marked another turn in the study. It was a shift in teacher mindset not only from feeling solely responsible for assessing student understanding but also *trusting* that students could accurately demonstrate what they learned. All three teachers had expressed frustration with a lack of responses from some students; however, at one lunch meeting Charlie and Kate’s conversation progressed to student motivation and learning. Charlie declared, “Kids admit that they are not trying!” and Kate added, “I’ve flat-out asked them if they studied or took notes, and they tell me that they don’t. I’ve reviewed expectations and shown them systems for how to take notes—how your brain needs multiple exposures to ideas—and they just want to rely on their memory, or me, to give them the information.” They concluded that while our self-assessment is a step in the right direction, we as a staff need to do more to be clear in moving our students toward learning, and not just earning a grade.

Our discussions had always maintained an undercurrent on this notion of our culture of learning, but until that moment, I was unsure if our work would actually bring it to light. The team had realized—just like their expectations on the self-assessment—that we, *collectively*, have to make our expectation explicit. Charlie summarized, “Our students think that a ‘D’ is okay. It’s sad that they just don’t take the opportunity to improve their grades when they can easily do so.” Both Kate and Anne nodded, and Kate asked, “Is it a motivation issue? A parent or family issue?” The two questions were left unanswered, and unlike the discussions from when we were planning our intervention, no one suggested top-down directives or professional development for the staff. Instead, they turned to what they had found from our work on the intervention. Anne added, “Perhaps if we had started this [the self-assessment] at the beginning of the year, we would have ironed out some of the vague feedback early on, and we would be at a point to really have those discussions on learning versus points.” While I hesitate to put so much credence in one simple exit slip, I feel that Anne’s point here was that there is value in opening the dialogue, in communicating with students. This transaction will help them understand the purpose of our classroom activities and the goals for their learning as well as make us better educators by connecting the dots of our instruction.

The revelation that so many of our students are satisfied with a “D” was disappointing. Not solely because our students were complacent but because the team still seemed to be holding on to the notion that grades motivated students to learn, to work. I had begun to see a shift in their attitudes towards assessment, but their perceptions on grades persisted. We had experienced glimmers of transformational change, but we remained at the transactional stage of this process. Heavy work like dislodging embedded beliefs and practice take significant time to change.

A Book Study without the Book

Having concluded the student intervention component of our study, I found myself reflecting on other aspects of this process, specifically, the value of working with this team of teachers. Each teacher expressed that they had enjoyed participating in the work, and their behavior supported their words: They regularly gave the students their exit slips; they completed the teacher reflections; and they participated in our lunch meetings, appearing on time (if not early in some cases) and actively discussed the topics as they arose. While it is possible that they completed the tasks because I am their supervisor, my experience with people tells me otherwise. Personally, I also thoroughly enjoyed collaborating with them in this capacity.

In every district that I have worked, a focus of our hiring practices centered on the applicant's perceived ability to build appropriate relationships with students. I cannot help but wonder if we as administrators sometimes forget the importance of building relationships with our staff. Certainly, I feel that I was "in the trenches" with this team of teachers and that we were working together to make their classrooms, at least, the "units of evaluation" Hattie (2012) describes is important to improving student learning (p. 191).

I cannot deny that being in the trenches with these teachers brought me closer to them, and I definitely feel it positively impacted our communication. About halfway through the third quarter, Charlie sought me out one morning to thank me for having a positive relationship with him. He explained that he felt he could come to me to work through challenges he faced and then described how his friend, who worked in another school district, did not have the same experience with their principal: "He told me that there's no way he could ever have a conversation with his principal about something he was struggling with in his classroom, or work together on a problem like we're doing."

In a separate conversation with Kate, we reflected upon the fact that we had worked on this process for nearly a year. The aspect she liked most about the process was that she was given the chance to problem-solve an issue with colleagues—something she felt was missing in most professional development activities. Instead of a "sit-and-get" as many professional development activities are, she said she valued discussing a real, current problem with colleagues in a focused way. Charlie agreed. "I think it was great and helped my instruction tremendously. When I go to a conference, I've only gotten nil out of PD like that. This is actually developing something to immediately implement in my classroom."

Kate ultimately likened it to "a book study without the book," without having to *do* the homework. "Yeah, you have this focused discussion but you don't have to do the tedious stuff that often comes with a book study." In an age when administrators are compelled to "inspect what they expect" from district professional development efforts, these staff members were asking for more. I had no reason to visit their classrooms to evaluate whether they were implementing the latest initiative. They were coming to me, sharing their insights and struggles. We were working together to improve instruction and student learning.

“They didn’t Teach Us that in Undergrad”

The range of experience among the team was rather significant and done so purposefully. On one end of the spectrum, Anne had 20 plus years of experience while the youngest, Charlie, completed his student teaching during the pandemic. Kate, our third member had under 10 years of instructional experience. Interestingly, the two younger members of the team had made reference to a lack of exposure to effective assessment practices when they were undergraduates. When the two reflected upon their work during this experience, they agreed that undergraduate experience had not prepared them for assessment practices beyond the traditional ones that they had experienced in high school.

I found that as we developed our student self-assessment, not only were their misunderstandings which centered on student motivation (that students will only do something if points are attached to it), their veritable assessment tool box was relatively limited. During our early meetings, either one or both would remark, “Ooh, I’d like to do something like that, but I don’t know what that would look like in my classroom,” or, “I’m tired of giving tests all of the time—and I know the kids get tired of it, too—but how do I use a rubric? I’ve never created one, let alone assessed something with it.” Given their lack of experience with some of the most basic assessment strategies (at no fault of their own) ran counter to the notion that students could give us accurate feedback.

Each team member grew as a result of this cycle of inquiry—whether it was a growing confidence in using assessments to inform their instruction or change in their views about students’ ability to accurate report their understanding. In essence, they expanded their capacity as an instructor—whether it was pedagogy or underlying beliefs. Most gratifying as an administrator, though, I feel that all found the intervention worthy of their time, and clearly, they

valued most the opportunity to discuss, banter, and explore topics with each other.

CHAPTER VIII: IMPLICATIONS AND NEXT STEPS

Is it possible that a simple exit slip or student self-assessment can shift student focus from achieving a certain grade to evaluating their learning? I highly doubt one intervention, no matter how well-designed, can achieve this goal. However, it *is* a step—one step—toward our ultimate goal of improving student learning and shifting our culture to one that focuses on learning. And the process itself revealed some of the staff's implicit beliefs about assessment, grades, and student motivation as well. After modeling their expectations for feedback, the team of teachers were surprised by student responses and their ability to use those responses to improve student learning. Once the communication followed the triple loop model more closely (D. Renn, personal communication, June 10, 2020), and the students were able to see their teachers respond to their needs, that is when the greatest strides were made in the teachers' instruction. In essence, the teachers were learning to "fold the assessment results back into instruction" (Cappuis et al., 2017, p. 2; Wiliam, 2011); they were accessing students' funds of knowledge (Moll, 1992) and began to see what students brought to the assessment (and learning) process as assets (Dudley-Marling & Dudley-Marling, 2015). In the end, all three participants indicated that they wanted to continue to use the student self-assessments as a way to improve student learning. As we look to our next cycle of inquiry, we want to build upon the team's work, and this cycle will include addressing deficit thinking as well as working to shift student mindsets on assessment and learning.

The Next Cycle

We established that deficit thinking is prevalent among our staff, so it is appropriate to begin addressing this problem head-on as part of our next cycle. A component of addressing this problem is to give students the opportunity to participate in improving their school and their

school's culture as well. To that end, I will be creating a Principal's Advisory Board next year to not only give students a voice in the structural "stuff" of school, but also to begin to shift their culture of assessment as well. I know that part of shifting our student culture will require us to break down barriers, perceptions, and stereotypes the students have toward school administrators, so I plan to use positive office referrals to have students meet with me. In this case, I will set aside time each week to meet with several students during our daily activity period, and I will work with the new Advisory Board to plan these meetings.

Funds of Knowledge

From the staff side, we will continue to shift toward more effective assessment practices, but we also need to shift away from deficit thinking. While the team continues to use the student self-assessment/exit slip, we can move into viewing students' funds of knowledge (Moll, 1992) and work through feedback loops with student assessment—folding back to the different layers of the cycle: recognizing our tacit assumptions, understanding the problem, designing a research-based assessment implementing and evaluating the assessment (D. Renn, personal communication April 20, 2022). While all students can benefit when appreciated by what they bring to the classroom, our vulnerable students will benefit greatly since they are most commonly viewed as deficits that need to be repaired.

Assessment

The greatest value has been working with this small team of teachers, and I do not want to underestimate that result. In a small school, few things happen that go unnoticed, especially by the staff. I have no doubt that all of our teachers are aware that I met regularly with Charlie, Anne, and Kate or that they were participating in a study on assessment. It was clear other teachers knew about the study and its topic when two unrelated staff members sought me out to

talk about their assessment practices, and this event took place during their non-evaluation year. Besides the rumors that may be circulating, value lay in the enthusiasm that these teacher share, and enthusiasm is contagious.

At first, Charlie remarked that he did not think the exit slip would work or that it would be too complex to benefit his instruction. At our individual meeting, he made the comment, “I was a bit nervous to implement it initially, but I like it so much that I’m going to keep using it...at least weekly.” In the end, he was the most enthusiastic about implementing it. When Kate met with me individually, she spoke to the value of this process, really in just having the opportunity to share ideas with colleagues. One idea for the next cycle was to hold monthly meetings at lunch in which we would have a focused topic to discuss, but Kate stopped this tangent: “It also needs to be something we’re doing, we’re trying. That sounds great, but we there has to be something else there, too.” I believe what she was looking for was “action.” To be most effective, we need to have an action component as well. After all, we have many committees in which we talk about issues and how to resolve them, but we do not have many in which we are trying out experimenting with different or new methods than our current practice.

The team valued the intervention that we developed and the entire process impacted the team members’ instructional practices, so we will continue using the student self-assessment, modifying it as needed. Part of those modifications will be for the team members to expand its use to all of their classes; additionally, Kate and Anne had already modified the assessment to having students write essential questions for students to use, so they want to explore other ways to focus students on the standards.

We all agreed that the assessment and the process was valuable, but I would pose that all of the team members valued our meetings most of all. This factor, then leads to how can we as a

team continue to meet and still maintain our excitement? A key component of the meetings was sharing ideas for adjusting their instruction, so giving them opportunities to monitor their shifts in instruction would be beneficial. I think it is more likely to happen if we expand our team. In light of Charlie's and Kate's realizations that assessment practices were not addressed in their teacher preparation programs, it stands to reason that our youngest teachers would benefit most from joining our team. While I would not include a first year teacher, I do feel that a second- or third year instructor would be over the "new teacher stress" and would appreciate and benefit from participating in our next cycle of inquiry.

Long-term Goals

Pulling back to the broader view of the Bixby School District, my ultimate goal is to extend this work throughout the entire high school with hopes that it will spread to the junior high school as well. As an administrator, I feel the greater value is in the process itself: identifying a problem, considering causes and investigating theories of action, developing and testing an intervention, then reflecting upon the process (Mintrop, 2018; Mintrop & Zumpe, 2019). Therefore, I hope that we can translate this process to other problems within our school community.

As my teachers and I move forward, the student self-assessment needs to continue to be one component of our cycle of inquiry as well as expanded to include more people and more teams. The three staff members with whom I worked found value not only in the student self-assessment tool but more so in the process. While we had valuable discussions about assessment, I also observed a change in how the team talked about grading practices and our vulnerable students. The team shifted from concerns about giving the assessment "correctly" and at the right time to modifying the exit slip to better suit their and their students' needs. Even

more impressive was the change in how the teachers discussed our vulnerable populations. A year prior, the team had expressed helplessness in assisting students whom they thought were unmotivated, and they sometimes blamed a perceived lack of family support as the root of student apathy. By the time of our final meetings, staff were no longer making comments that "These kids don't care." I hope that we can build upon the growth I have seen among these three teachers an achieve a greater impact—one that can positively influence all learners in our building by following the double-loop learning model and continue to broaden our sphere of influence as we address the rest of the staff members' tacit assumptions (Mintrop, 2018).

The teachers indicated that they would continue to use the student exit slips as part of their assessment process; however, we have more work to do. Unfortunately, our vulnerable students did not explicitly receive special consideration during our cycle of inquiry since precedence was given to initiating the self-assessment in a formative approach. But as we explore learning more about funds of knowledge and addressing deficit thinking, the team will use that opportunity to review vulnerable students' responses to their self-assessment/exit slips.

As the team discussed our next cycle of inquiry, they expressed a desire to welcome more staff members to try integrating exit slips into their instruction as well. Even as we discussed whom they would like to invite, we realized it would not be as simple as giving others the exit slip. To the contrary, our exit slip intervention resulted from identifying a problem of practice and rooting the experiment in literature. Merely giving someone the exit slip and telling them "Try this; it works" would be the equivalent of the sit-and-get professional development the team loathed.

While our intervention may have piqued staff curiosity and even inspired a few to clarify their assessment processes, we did not want to rush into implementing a school-side intervention

for which staff were unprepared (or have any buy-in). The team recognized that they still wanted practice in refining how they used the exit slip, and they wanted more tools for reteaching. As a result, Kate suggested that I take it to the Leadership Team during our annual retreat.

Kate is a member of our school Leadership Team, and our annual retreat serves as an opportunity to plan for the next year's professional development. The Leadership Team aspires to have a continuation of professional development goals from year-to-year and try to avoid jumping from one initiative to another. We use student data (grades, standardized tests, teacher feedback) and staff data (teacher evaluation, interest surveys, and feedback) to determine our goals. Kate suggested building upon our current learning goal, which was to explicitly talk about learning with students to making "space" for teachers to share assessment approaches. Kate was translating what she valued most within our cycle of inquiring and was making it available to other staff members.

We began this journey with the intent to improve student learning—to *shift* student and teacher preoccupation with points and grades to one of learning. However, what I found was that staff themselves shared that focus, so to even begin to initiate that change, it must begin with the adults. The team had identified a problem of practice, but we also found that our teachers had a *gap* in practice (Simmons & Taylor, 2019). The youngest teachers among the participants indicated that this gap had originated at the collegiate level, so in an age when I often feel that administrators are expected to do everything, this is an area that may need to be evaluated by teacher preparation programs. Certainly, school districts may need to consider making assessment practices a component of their new teacher induction and mentoring programs.

Bixby's Assessment Future

Bixby has a relatively high level of success, and that success is likely the result of a proud, supportive community, a collective commitment to learning, and a caring staff. We want our students to learn. We want our students to be successful. Our successes, though, came about despite the fact that our well-intentioned staff practiced traditional, teacher-centered instructional approaches all while claiming that they were "cutting-edge." We realized that something was amiss when we returned to in-person instruction after the pandemic: Nearly 60% of our high school students were receiving D's or F's in their classes. As a small team of teachers and I explored our grading problems, conversations uncovered that our staff held many tacit assumptions about rigor, effectively assessing students, and student motivation. Therefore, we developed a simple intervention in the form of a student self-assessment to see if it could positively impact student learning.

As we worked together to develop, implement, and evaluate the intervention, the team members' knowledge on assessment practices grew, as did their confidence in adapting their instruction in response to student feedback. When this communication between the students and the teacher reached "flow," it became a transactional process that not only proved motivating for the students, the teachers found it motivating as well. Outside of the intervention itself, the greatest results were the impact on the teachers' *perceptions* of students—transitioning from being skeptical on whether the students would take the assessments seriously (let alone giving them useful feedback) to viewing the students as sources of knowledge. This powerful shift has the potential to positively impact student learning, with the greatest potential to impact our district's most vulnerable students. Instead of viewing them with a deficit mindset, going through the process of identifying a problem of practice and developing and using the

intervention helped my team to value their students as sources of knowledge that could positively impact their instruction, and ultimately all student learning.

As an administrator, going through this process was a positive experience for me as well. Being "in the trenches" with the team members grew our relationship as well as increased their access to me (and me to them). I did not have to "inspect what I expected." Instead, they came to me to share results and troubleshoot issues. And because both the team members and I found it enjoyable and valuable, we have already begun considering how we can expand our experience to other staff members. In the end, the team had hoped to improve student learning, and I have no doubt they will consider the process effective if there are fewer D's and F's in their classes at the end of the grading period. However, I hoped to have guided them to a deeper goal, one resulting in more valid assessments and assessment practices as well as cultivating a culture of authentic assessment. At the student level, we will know if we have met our goals if our data show improvement in student learning (local and standardized assessments). At the teachers' level, we hope to see clear assessment practices that are equitable, varied, and that actively engage students; additionally, teachers should be able to root their decisions in current research towards best practice. Finally, from the administrative level, I hope our small group ignites interest among the rest of the staff—something that we are already seeing through various conversations and professional development requests.

This one, small activity has inspired our small team to build upon what we have already established, and I have no doubt that our little school on the prairie will continue to benefit from these efforts. As we are finally reaching normalcy after the pandemic, we continue to learn from the lessons this experience taught us—not the least of these is that education continues evolve and there are always opportunities to improve our practice. I hope that our little school on the

prairie retains all of its positive attributes—including support of student learning and a deep empathy for everyone in our building. But I also hope we are able to see that changes, that growth, can be a positive experience and that all of our students—no matter their background—have value if only we take the time to look.

REFERENCES

Asadullah, M. N., Bouhlila, D. S., Chan, S., Draxler, A., Ha, W., Heyneman, S. P., Lusche, T. F., Semela, T, & Yemini, M. (2023, March 23). A year of missed opportunity: Post-Covid learning loss—a renewed call to action. *International Journal of Educational Development, 99.* https://www.sciencedirect.com/science/article/pii/S0738059323000469?via%3Dihub

Belzer, A. & Ryan, S. (2013). Defining the problem of practice dissertation: Where's the practice, what's the problem? *Planning and Changing, 44*(3/4), 195-207.

Black, P. & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Kappan, 92*(1), 81-90.

Blankstein, A. M. (2013). *Failure is not an option: 6 principles that advance student achievement in highly effective schools* (3rd ed.). Corwin.

Bolman, L. G. & Deal, T. E. (2017). *Reframing organizations* (6th ed.). Jossey-Bass.

Books, S. (1997). The other poor: Rural poverty and education. *Educational Foundations, 11*(1), 73-85.

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *American Education Research Association, 86*(4), 803–848.

Bryk, A. S. (2009). Support a Science of Performance Improvement. *Phi Delta Kappan*, *90*(8), 597–600.

Cappuis, S. Commodore, C. & Stiggins, R. (2017). *Balanced assessment systems: Leadership, quality, & the role of classroom assessment*. Corwin.

Carnegie Project on the Education Doctorate. (2022). *The CPED framework*. CPED. CPEDInitiative.org/the-framework

Chandler, R. (2014). Teachers' beliefs about poverty and the impact on learning disabilities identification in a poor, rural school district. *Rural Educator, 35*(3), 31-39.

Coghlan, D. & Brannick, T. (2010). *Doing action research in your own organization* (3rd ed.). Sage Publications.

Conley, D. T. (2018). *The promise and practice of Next Generation Assessment*. Harvard Education Press.

Cotton, K. (1996). Affective and social benefits of small-scale schooling. ERIC Digest. (ED 401 088).

Creswell, J.W. & Creswell, J.D. (2018). *Research design* (5th ed.). Sage.

Dahill-Brown, S. & Jochim, A. (2018). The power of place: Rural identity and the politics of rural school reform. In M. Q. McShane & A Smarick (Eds.), *No longer forgotten: The triumphs and struggles of rural education in America* (pp. 59-79). Rowan & Littlefield.

Dudley-Marling, C. & Dudley-Marling, A. (2015). Inclusive leadership and poverty. In G. Theoharis & M. Scanlan (Eds.), *Leadership for increasingly diverse schools* (pp. 39-57). Routledge.

DuFour, R. (2004). What is a "Professional Learning Community"? *Educational Leadership, 61*(8), 6-11. https://www.ascd.org/el/articles/what-is-a-professional-learning-community

Eagleton, T. (1983). *Literary theory: An introduction*. University of Minneapolis Press.

Fisher, D., Frey, N., & Pumpian, I. (2011). No penalties for practice. *Educational leadership, 69*(3), 46–51.

Greeno, J. G., Collins, A. M. and Resnick, L. B. (1996). Cognition and learning. In Berliner, D. and Calfee, R. (Eds.), *Handbook of Educational Psychology* (pp. 15-41). MacMillian.

Guskey, T. R. (2010). Lessons of mastery learning. *Educational leadership, 68*(2), 52–57.

Guskey, T. R. (2011). Five obstacles to grading reform. *Educational Leadership, 69*(3), 16–21.

Guskey, T. R. (2017). Where do you want to get to? Effective professional learning begins with a clear destination in mind. *Learning Professional, 38*(2), 32–37.

Guskey, T. R. (2020a). Breaking up the grade. *Educational Leadership, 78*(1), 40–46.

Guskey, T. R. (2020b). The dark side of assessment literacy: Avoiding the perils of accountability. *AASA Journal of Scholarship and Practice, 17*(1), 7-15.

Guskey, T. R., & Anderman, E. M. (2013). In search of a useful definition of mastery. *Educational Leadership, 71*(4), 18–23.

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

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.

Hatti, J. & Yates, G. (2014). *Visible learning and the science of how we learn*. Routledge.

Hill, M. F. (2011). 'Getting traction': Enablers and barriers to implementing assessment for learning in secondary schools. *Assessment in Education: Principles, Policy & Practice, 18*(4), 347–364. https://doi.org/10.1080/0969594X.2011.600247

Heifetz, R., Grashow, A., & Linsky, M. (2009). *The practice of adaptive leadership: Tools and tactics for changing your organization and the world*. Harvard Business Press.

Henriksen, D., & Richardson, C. (2017). Teachers Are Designers: Addressing Problems of Practice in Education. *Phi Delta Kappan*, *99*(2), 60–64.

Irvin, M. J., Meece, J. L., Byun, S., Farmer, T. W., Hutchins, B. C. (2011). Relationship of school context to rural youth's educational achievement and aspirations. *Journal of Youth & Adolescence, 40*(9), 1225-1242.

Jensen, E. (2009). *Teaching with poverty in mind.* ASCD.

Jensen, E. & Snider, C. (2013). *Turnaround tools for the teenage brain*. Jossey-Bass.

Lambert, L., Walker, D., Zimmerman, D. P., Cooper, J. E., Lambert, M. D., Gardner, M. E., & Szabo, M. (2002). *The constructivist leader* (2nd ed.). Teachers College Press.

Lane, B. (2008). *But how do you teach writing?* Scholastic.

Lawson, M. A. & Lawson, H. A. (2013). New conceptual frameworks for student engagement research, policy, and practice. *Review of Educational Research, 83*(3), 432-479.

Malkus, N. (2018). A statistical portrait of rural education in America. In M. Q. McShane & A. Smarick (Eds.), *No longer forgotten: The triumphs and struggles of rural education in America* (pp. 9-28). Rowan & Littlefield.

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

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. Jossey-Bass.

Mintrop, R. (2018). *Design-based school improvement*. Harvard Education Press.

Mintrop, R. & Zumpe, E. (2019, March). Solving real-life problems of practice and education leaders' school improvement mind-set. *American Journal of Education, 125*(3), 295-344.

McTighe, J. & Willis, J. (2019). *Upgrade your teaching: Understanding by design meets neuroscience*. ASCD.

Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory into Practice, 31*(2), 132-141.

Mui, C. (2015, March 23). Do you have the patient urgency required for successful innovation? *Forbes*. https://www.forbes.com/sites/chunkamui/2015/03/23/patient-urgency/?sh=574e16f54057

Muñoz, M. A., & Guskey, T. R. (2015). Standards-based grading and reporting will improve education. *Phi Delta Kappan, 96*(7), 64–68.

Nolen, S. B. (2011). The role of educational systems in the link between formative assessment and motivation. *Theory into Practice, 50*, 319-326.

Nordengren, C., & Guskey, T. R. (2020). Chart a clear course: Evaluation is key to building better, more relevant learning. *Learning Professional, 41*(5), 46–50.

O'Connor, K. (2011). *A repair kit for grading: 15 fixes for broken grades*. Pearson.

O'Connor, K., Jung, L. A., & Reeves, D. (2018). Gearing up for fast grading and reporting. *Phi Delta Kappan, 99*(8), 67–71.

O'Connor, K., & Wormeli, R. (2011). Reporting student learning. *Educational Leadership, 69*(3), 40–44.

Percell, J. C. (2017). Lessons from alternative grading: Essential qualities of teacher feedback. *Clearing House: A Journal of Educational Strategies, Issues and Ideas, 90*(4), 111–115. Routledge.

Perry, J. A., Zambo, D. & Crow, R. (2020). *The improvement science dissertation in practice*. Myers Education Press.

Popham, W. J. (2001). *The truth about testing: An educator's call to action*. ASCD.

Popham, W. J. (2018). *Assessment literacy for educators in a hurry*. ASCD.

Reeves, D., Jung, L. A., & O'Connor, K. (2017). Special topic: What's worth fighting against in grading? *Educational Leadership, 74*(8), 42–45.

Rosenblatt, L. M. (1978). *The reader, the text, the poem*. Southern Illinois University Press.

Schimmer, T. (2016). *Grading from the inside out: Bringing accuracy to student assessment through a standards-based mindset*. Solution Tree Press.

Schmoker, M. (2018). *Focus: Elevating the essentials to radically improve student learning* (2nd ed.). ASCD.

Shields, C. M. (2020). *Becoming a transformative leader: A guide to creating equitable schools.* Routledge.

Shields, C. M. & Hesbol, K. A. (2020). Transformative leadership approaches to inclusion, equity, and social justice. *Journal of School Leadership, 20*(1), 3-22.

Simmons, N. & Taylor, K. L. (Spring 2019). Leadership for the scholarship of teaching and learning: Understanding bridges and gaps in practice. *The Canadian Journal for the Scholarship of Teaching and Learning, 10*(1), 1-18, Article 2.

Swan, G. M., Guskey, T. R., & Jung, L. A. (2014). Parents' and teachers' perceptions of standards-based and traditional report cards. *Educational Assessment, Evaluation and Accountability, 26*(3), 289–299.

Theoharis, G. (2007). Social justice educational leaders and resistance: Toward a theory of social justice leadership. *Educational Administration Quarterly, 43*(2). 221-258.

U.S. Department of Agriculture (2020). *Rural America at a Glance*, 2020 Edition. United Stated Department of Agriculture, Economic Research Service, Economic Information Bulletin, 221, December 2020.

Vegas, E. & Winthrop, R. (2020, September 8). Beyond reopening schools: How education can emerge stronger than before COVID-19. Brookings Institute. https://www.brookings.edu/research/beyond-reopening-schools-how-education-can-emerge-stronger-than-before-covid-19/

Valencia, R. R. (2010). *Dismantling contemporary deficit thinking*. Routledge.

Venables, D. R. (2018). *Facilitating teacher teams and authentic PLCs: The human side of leading people, protocols, and practices*. ASCD.

West, E. M., Zolkoski, S. M., Lockhart, J. R., Holm, J. M., & Tremont, J. (2022). "Everybody knows everybody": Adolescents' perceptions of what helps them succeed in a rural Title I school. *Journal of Adolescent Research, 27*(5). 672-690.

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *The Phi Delta Kappan, 70*(9). 703-713.

Wiliam, D. (2011). *Embedded formative assessment*. Solution Tree Press.

Wormeli, R. (2006). *Fair isn't always equal: Assessing & grading in the differentiated classroom*. Stenhouse.

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

Wednesday, May 26th, 2021

HS Teachers

Groups:

1	2	3	4
[redacted]	[redacted]	[redacted]	[redacted]

Complete Task 1 as a group

After you complete task one,
use the clue to find two...,
And continue on until all are through.

No stealing, no peeking, or hiding anyone else's paper,
Even though everyone has a similar caper.

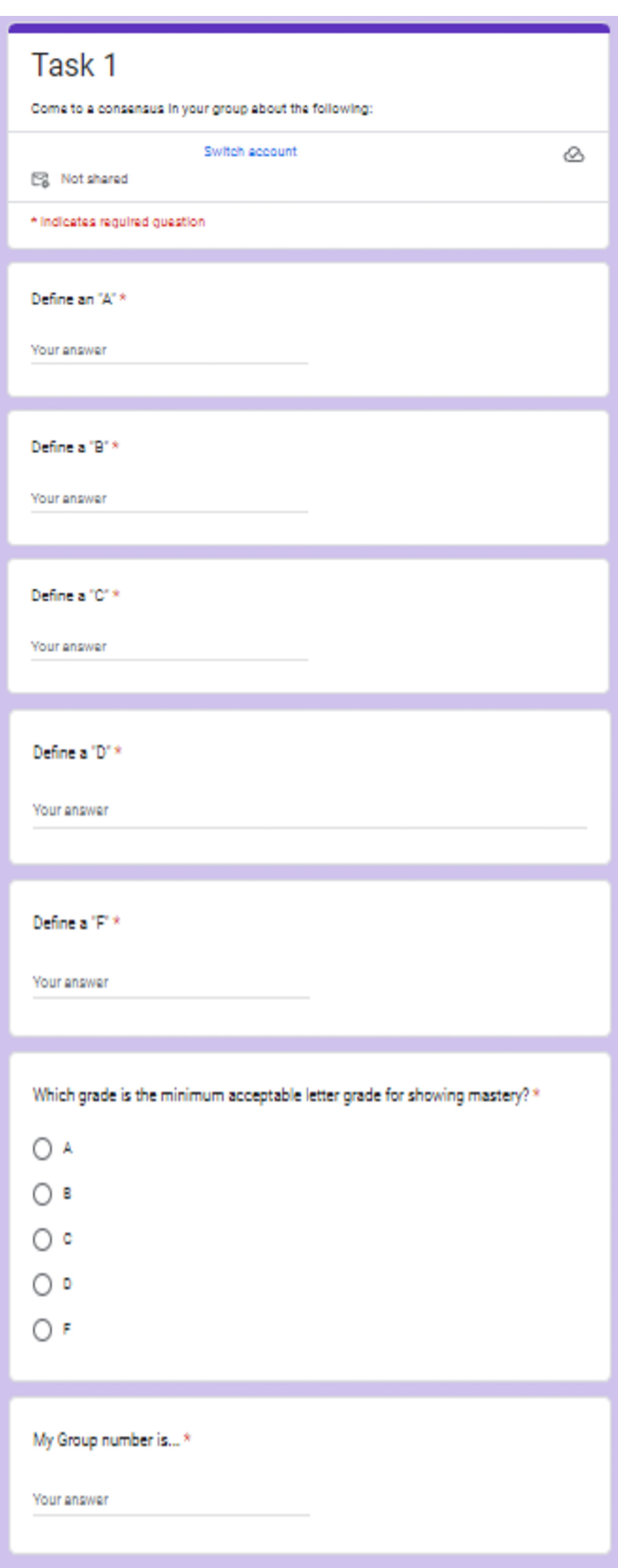
See the keeper of books for task two,
She can't give you answers or direct you.

There's lots to do, you'll be tempted to run,
But turn in your tasks when they're all done.

We'll return in the end,
To where this school year began.

I hope you find this game to be fun,
In the end, we'll see if by knowledge you've won...

Task 1:


Task 1

Come to a consensus in your group about the following:

Switch account

Not shared

* Indicates required question

Define an "A" *

Your answer

Define a "B" *

Your answer

Define a "C" *

Your answer

Define a "D" *

Your answer

Define a "F" *

Your answer

Which grade is the minimum acceptable letter grade for showing mastery? *

- A
- B
- C
- D
- F

My Group number is... *

Your answer

Task 2:

A

The daily high temperature readings for a week in central Illinois during July:

90, 87, 92, 88, 0*, 85, 86

*reading was not recorded

What is the average temperature for the week above?

Task 2:

B

The following are the daily balance of your checking account:

$1,200; $1,550; $1,350; 0*; $1,250; $1,150; $1,100

*computer glitch, failed to record balance

What is the average daily balance of your account?

Task 3:

In December and January, assignments were placed in the Weekly Updates. Some were to respond to questions, and others were to complete an activity in our Staff Google Classroom. Here are the scores for staff members (order randomized):

	#1	#2	#3*
A	0	0	1
B	2	0	1
C	2	2	1
D	2	2	0
E	2	2	1
F	2	0	0
G	0	0	0
H	0	2	2
I	0	0	0
J	2	2	2
K	2	2	2
L	0	2	1
M	2	0	1
N	2	0	2
O	2	0	1
P	2	0	1
Q	2	2	1
R	2	0	1
S	2	2	1
T	2	2	1
*2=complete; 1=partial or failed to upload; 0=did not complete			

Would it be fair to record teachers' averaged scores in their evaluations? Why or why not?

Task 4:

According to researchers*, “A grade represents a valid and undiluted indicator of what a student knows and is able to do—mastery.” Consider the following averaged grades for a student. Which student has the most valid grade?

Student	Hmwk /10	Hmwk /20	Hmwk /20	Project /50	Test /100	Total Points	Avg
A	10	20	20	10	70	130	65%
B	0	0	0	45	95	140	70%
C	8	19	15	35	85	162	81%

Next Clue: Mirror, Mirror, next to the stall; tell me my next task on the wall.

Task 5:

Individually define “Mastery” in your content, and turn your definition into Jennifer, along with your other tasks.

APPENDIX B: THE EXIT SLIP/STUDENT SELF-ASSESSMENT

Student Self-Assessment

Form description

This form is automatically collecting emails for users. Change settings

Name (last, first) *

Short answer text

What I learned/know from today's lesson (this unit): *

Short answer text

Evidence of what I learned today: *

Long answer text

What I'm still working on from today's lesson (this unit): *

Short answer text

What do I need help with? *

Short answer text

APPENDIX C: TEACHER REFLECTION

Teacher Reflection

Form description

Name

Short answer text

Reflection of using the student self-assessment (what's going well, what's not going well, surprises, etc.):

Long answer text

How have you adjusted your instruction due to the feedback on the student self-assessment?

Long answer text

Anything else to note?

Long answer text