Fall 2019

Using Standards-Based Grading Practices to Support Student Self-Assessment

Alison Espinosa

USING STANDARDS-BASED GRADING PRACTICES TO SUPPORT STUDENT SELF-ASSESSMENT

by

Alison Espinosa

Bachelor of Science
University of South Carolina, 2006

Master of Teaching
University of South Carolina, 2007

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Education in

Curriculum and Instruction

College of Education

University of South Carolina

2019

Accepted by:

Leigh D'Amico, Major Professor

Christopher Bogiages, Committee Member

Suha Tamim, Committee Member

Natalie Dutrow, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

Dedication

This work is dedicated to my husband, Hugo, for creating the space in our lives for it to take place; to my parents, Ron and Sandy, for flaming my desire to pursue it; and to Bruce, for keeping my feet warm.

Abstract

The purpose of this action research study was to determine the impact that four standards-based (SBG) grading practices have on students' ability to critically and accurately self-assess their understanding of secondary mathematics learning objectives, as well as the overall impact on student learning. This research took place over twelve instructional weeks in a ninth-grade mathematics class. During the first four weeks, the baseline cycle, students completed weekly formative assessment items and an accompanying assessment reflection; traditional assessment practices were enforced during the baseline cycle. During the final eight weeks, the treatment cycle, students continued to complete weekly formative assessment items and accompanying assessment reflection, but four SBG practices were introduced. The four SBG practices are 1) alignment of formative assessment items to content descriptors, 2) opportunity to be assessed at varying levels of difficulty, 3) student choice of assessment difficulty, and 4) opportunity to improve assessment results.

A concurrent triangulation mixed methods methodology was used, and multiple forms of data were collected. The impact of the four SBG practices on the accuracy of students' self-assessment (SA) and the overall impact on student learning was analyzed quantitatively. The impact on the quality of students' SA was analyzed qualitatively. Data analysis revealed that the treatment positively impacted the accuracy of students' SA and student achievement, but no significant impact on the quality of students' SA. However,

throughout this action research study I engaged in a cycle of systematic inquiry by observing student behaviors and reflecting on the qualitative data. By doing so I learned that what and how students think about the opportunity to self-assess, or the value they have for self-assessing their performance, impacted the quality of their SA.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AR .... Assessment Reflection

CCSM .... Common Core Mathematics Standards

DoK .... Depth of Knowledge

NBPTS .... National Board for Professional Teaching Standards

NCLB .... No Child Left Behind

MGO .... Mastery Goal Orientation

PGO .... Performance Goal Orientation

PLC .... Professional Learning Community

RTT .... Race to the Top

SA .... Self-Assessment

SBG .... Standards-Based Grading

SRL .... Self-Regulated Learning

Chapter 1:

Nature and Significance of the Problem

Introduction

Over the past 30 years two major policy streams have converged in American K-12 education: standards-based curriculum and accountability reform, and teacher quality reform efforts (Polikoff & Porter, 2014). Federal legislation including No Child Left Behind and the Race to the Top Fund has increased scrutiny of assessment data and emphasized state and national standards (Rosales, 2013). As teacher evaluation is increasingly linked to student assessment data (Hoernke, 2014), and grades impact students' college acceptance, career trajectory, and scholarship potential, teachers, students, and parents are concerned about student assessment data (Hochbein & Pollio, 2016). Critics of test-based accountability argue that emphasizing assessment data negatively impacts the quality of instruction, as many accountability tests and aligned test preparation materials require low cognitive demand (Blazar & Pollard, 2017). Conversely, if accountability assessments are aligned to rigorous standards, and teachers align their teaching to these standards, using assessment items to inform instruction could have a positive impact on the quality of instructional strategies (Polikoff & Porter, 2014).

Teachers are at the center of this work to implement standards and assess student mastery of those standards. Most of my career as a secondary level mathematics teacher has been spent trying to determine what students know. Student learning frames every pedagogical decision, from planning and delivering instruction, to reflecting on instruction. I have spent extensive time assessing what students do and do not know, often communicating this information to students in the form of numeric marks on formative and summative assessments. While feedback from teachers is crucial to promote student understanding and mastery of information, recently I began considering how students can construct this knowledge themselves by critically self-assessing their thinking and understanding.

Self-assessment (SA) challenges students to evaluate their strengths and weaknesses with the intention of improving future performances (Desjarlais & Smith, 2011). Stallings and Tascione (1996) found that regular SA resulted in improved vocabulary and communication skills, and increased learning autonomy in mathematics students. In order to self-assess their mastery of material, students must first understand the purpose of the SA, a clear interpretation of grades, and the criteria required for success (Desjarlais & Smith, 2011). Thus, it is important that students develop a clear understanding of the assessment criteria and aligned content standards.

Problem of Practice/Purpose of Study

In 2016 the National Board for Professional Teaching Standards (NBPTS) endorsed the practice of supporting SA by asking candidates to articulate this practice in their written commentary for component four of a four component process. In my experience as a National Board candidate support provider, I found that candidates struggled to identify and articulate instructional strategies that support SA. I further realized that I shared their struggle; consequently I began investigating strategies that I could implement to better support students' SA of their performance and understanding.

Standards-based grading (SBG) practices improve standardized assessment scores and communication between teachers, students, and parents about the meaning of grades and the criteria for success (Hochbein & Pollio, 2016). It follows that implementing SBG practices may support students' ability to self-assess their thinking and understanding. Wondering whether SA practices are supported by a better understanding of classroom assessments practices and interpretation of grades, this study seeks to determine the impact that SBG practices have on students' ability to self-assess their performance and understanding.

This practical action research study was motivated by a desire to support the development of students' ability to self-assess their understanding and performance, while also considering the impact that these practices have on student learning. I investigated whether implementing specific SBG practices improved Secondary Mathematics 1 Honors students' ability to critically and accurately self-assess their understanding and measured the impact that these practices have on student achievement.

Research Questions

The following research questions frame this practical action research study:

1. What impact do standards-based grading practices have on Secondary Math 1 Honors students' ability to *accurately* self-assess their performance on assessments items correlated to unique content standards?
2. What impact do standards-based grading practices have on the *quality* of Secondary Math 1 Honors students' self-assessment of their understanding of the content standards?

3. What impact do standards-based grading practices have on student *achievement* in Secondary Math 1 Honors?

Research Design

Using a case study design, this practical action research study occurred over a twelve-week research cycle. Students completed one formative assessment per week, as well as an accompanying assessment reflection, throughout this twelve-week study. During the first four weeks, the baseline cycle, traditional assessment practices were maintained. During the final eight weeks of the research cycle, the treatment cycle, four identified SBG practices were introduced; all other practices were consistent throughout the baseline and treatment cycles. The four SBG practices are 1) alignment of formative assessment items to content descriptors, 2) opportunity to be assessed at varying levels of difficulty, 3) student choice of assessment difficulty, and 4) opportunity to improve assessment results. To ensure that students were willing to challenge themselves by selecting high difficulty level assessments while accommodating the traditional grading system dictated by school policy, successful completion of each assessment level resulted in different academic marks. Students who completed a tier 1 assessment at the *proficient* level earned a maximum mark of 80%. Students who completed an a*pproaching mastery* level tier 2 assessment earned a maximum mark of 90%. Students who completed a *mastery* level tier 3 assessment can earn up to 100%. Efforts were made to maintain consistency of all other instructional and assessment practices throughout the entire twelve-week research cycle to increase the trustworthiness and rigor of the research results.

As the research questions require analysis of changes in the accuracy of students' SA and student achievement, as well as analysis of changes to the quality of students' SA of their understanding, I chose a mixed method research design (Efron & Ravid, 2013). The nature of the research questions designates a unique need for either qualitative or quantitative measures, thus a concurrent triangulation method design is appropriate (Efron & Ravid, 2013). Multiple forms of data were collected, and multiple opportunities were provided for students to demonstrate their reflective thinking, increasing the trustworthiness of the research results (Merriam & Tisdell, 2016). Once per week, throughout the baseline and treatment cycles, both qualitative and quantitative data was collected and analyzed.

Qualitative research is used to understand how students make sense of their experiences (Merriam & Tisdell, 2016); thus, qualitative methods were used to collect and analyze students' ability to critically self-assess through written narrative. After completing a formative assessment item, students completed an assessment reflection instrument. Student responses to the qualitative items were analyzed inductively throughout the data collection period to determine common themes among responses and monitor student thinking about their learning and progress towards accurate SA. Qualitative data was categorized by emergent themes using an analytic approach to coding.

The assessment reflection questionnaire also required students to predict the score they will receive after completing each assessment item. The accuracy of students' self-assessment was measured quantitatively as the difference between the predicted and achieved score. The achieved score was determined by the researcher by comparing student work against a pre-established rubric. A negative measure of accuracy indicates that a student underestimated their learning, while a positive measure of accuracy indicates that a student over-estimated their learning. The closer the measure of accuracy is to zero, the more accurate the SA. Descriptive statistics were used to determine the significance of the treatment results.

Lastly, quantitative methods were used to measure the impact that SBG practices have on student achievement. Students completed a pre-assessment at the beginning of both the baseline and treatment cycles, as well as a post-assessment at the conclusion of each cycle. Descriptive statistics were used to analyze changes in student performance between pre- and post-assessments to measure changes in student achievement.

Multiple data sources were used to increase the rigor and trustworthiness of the research results. A reflective journal was maintained by the researcher in order to maintain consistency, increase reflexivity, and reduce bias (Merriam & Tisdell, 2016). Efforts were made to limit extraneous variables and maintain consistent instructional practices throughout the research cycle in order to increase research validity.

Concepts and Definitions

Standards-Based Grading (SBG)

Implementing a SBG system requires a shift in teaching and instruction. Teachers must use standards to drive learner-centered instruction, and assessments must be aligned to individual standards or content descriptors (Colby, 1999). Established learning goals at different levels of difficulty should clearly communicate what students need to know and are required to do; providing opportunities to demonstrate understanding at progressive levels of complexity enables all students to work towards attainable goals while reducing boredom and frustration (Clayton & Shores, 2015).

A variety of instructional strategies and assessment practices are associated with SBG; some of which conflict with the assessment practices at my school. Thus, selecting a limited number of practices associated with SBG to implement during the research cycle was necessary. The following practices were selected as common assessment practices associated with SBG (Colby, 1999; Hochbein & Pollio, 2016; Marzano, 2009; Miller, 2013):

1. All assessment items are explicitly aligned to unique content descriptors.
2. All content descriptors are assessed at three tiers of difficulty (*proficient, approaching mastery, mastery*).
3. Students are able to self-select the assessment difficulty level for each content descriptor.
4. Students are able to improve the score they receive on all assessments, choosing to take assessments at a higher difficulty level to reflect their growth in understanding.

Self-Assessment (SA)

The construct SA may be confused with other related, or intertwined, practices. Some researchers use related practices interchangeably or overlap practices within each construct. For example, Leise's (2010) process of reflecting on performance closely aligns with the practice of SA. Requiring the articulation of specific learning criteria so that subjects may judge their performance or position in a rubric, Leise's (2010) theory of reflection prompts students to reflect on their performance in order to support growth and future learning. Generally, however, this research denotes several distinctions between the practices of reflection and self-assessment.

In this study, student reflection is defined as the process students take when thinking about their learning; before, during, or after a lesson or assessment. Self-assessment may occur during reflection, but specifically refers to the process students take when relating their level of understanding to the expectations established by the content descriptors and evaluating their performance on previously completed assessment items. The purpose of SA is to improve future performance by identifying strengths and areas of improvement. Self-assessment may occur while students are reflecting on their performance, but SA is done with unique goals and intentions. Whereas reflection is reactive, intended to gain insight and help students make sense of an experience, SA is proactive; a practice intended to support student growth in understanding (Desjarlais & Smith, 2011). "The goal of reflection is 'knowing'; the goal of self-assessment is 'growing'" (Desjarlais & Smith, 2011, p.96).

Metacognitive learning strategies, self-efficacy, and self-regulated learning strategies share similar and often overlapping characteristics with SA; but SA has differentiating practices or intentions. Metacognition is a process that incorporates elements of SA or reflection on performance but is a much broader process that includes other cognitive processes, including planning learning strategies and monitoring understanding while learning (Leise, 2010). Self-efficacy is a belief that one is able to perform successfully on an upcoming task, providing a judgment on ability; SA evaluates performance on a task that has already been completed, providing a judgment on knowledge (Sung, 2006).

Significance of Study

The results of this action research study provide insight into whether implementing specific SBG practices support the quality and accuracy of students' SA. Evidence suggests that SA results in an increased value of learning and belief that learning is impacted by intrinsic factors, rather than luck or other unknown variables (Brookhart et al., 2004). When students are required to self-assess regularly, Stallings and Tascione (1996) found that they began to self-assess their performance on tasks even when SA was not required. This resulted in improved communication, increased perseverance, and motivation compared to students who were not supported in self-assessing their performance (Stallings & Tascione, 1996). Further, accurate SA promotes self-regulated learning, a learning theory linked to increased learner autonomy and achievement. Thus, determining classroom practices that support self-assessment is valuable to teachers and students.

Limitations/Delimitations

Action research is a deliberate, cyclical, reflective research process in which the research is conducted by insiders, usually within the researcher's local context (Herr & Anderson, 2015). Whereas traditional researchers typically conduct their research by imposing carefully constructed treatments within a controlled setting, action researchers engage as participants in the research setting (Herr & Anderson, 2015). Focusing more on local application than global theory, action research is a process of inquiry, usually initiated to solve an immediate local problem (Herr & Anderson, 2015).

As the classroom teacher I implemented the treatment during regular classroom instruction, collected the resulting data, and analyzed the results. The collection and analysis of qualitative data required a level of subjectivity; thus, it was necessary that I monitor bias while analyzing and reporting the data. Though the purpose of this research is to inform my personal practice, the results of this study are applicable to secondary level mathematics teachers. Characteristics of the research participants should be considered when determining the transferability of the research results. Focus should be given to mathematics students who display characteristics similar to those included in the research sample. Further research is needed to fully determine the impact of implementing these practices across different content areas.

Organization of the Study

This action research dissertation provides insight into whether implementing SBG practices supports the accuracy and quality of students' SA, as well as the impact on students' understanding of mathematics content. The remaining chapters consist of a thorough review of related literature, description of data collection methods and procedures, data analysis, and discussion of the research findings. Chapter two, the literature review, provides a concise summary of research related to SBG and SA. Chapter three, methods and procedures, thoroughly describes the research design, the instruments and procedures used to collect data, and methods of data analysis. Chapter four, data analysis, provides an overview of the data, as well as the quantitative and qualitative methods used to interpret the data. Chapter five, research findings, provides a summary of findings, an overview of implications, and recommendations for future research. Lastly, instruments are provided in the appendices.

Summary of the Chapter

Supporting students' ability to self-assess their learning further supports the development of positive student affect (Brookhart et al., 2004). As students who are regularly provided opportunities to self-assess their learning are more likely to develop feelings of empowerment, autonomy, and authority over their learning (Brookhart et al., 2004), supporting SA is a valuable pedagogical strategy. However, it is unknown what instructional strategies are most effective for developing these skills in students. The results of this action research study provide evidence of whether specific instructional strategies support SA, and the impact that this has on student learning of mathematics.

This action research study introduces four identified practices of SBG in a ninth grade Secondary Math 1 Honors class. The quality and accuracy of students' ability to self-assess their performance on criteria aligned formative assessments was analyzed using quantitative and qualitative data analysis, along with the impact that these practices have on student achievement. The results were used immediately to inform my personal practice; consideration of the research design, participants, data collection, and analysis should be considered when implementing these practices in other contexts.

Chapter 2:

Review of the Literature

Introduction

In 2016, the National Board for Professional Teaching Standards (NBPTS) released the guidelines for Component 4, the final component of the National Board Certification application process. By doing so, the NBPTS endorsed the practice of self-assessment (SA), as this component included a written prompt that required National Board candidates to describe how they support students' use of SA to promote student learning (NBPTS, 2016). In my role as a National Board candidate support provider, I found that teachers struggled to identify intentional strategies to support this practice; our conversations revealed that these candidates and I were not providing opportunities for students to regularly self-assess in our classrooms. Even more, that we were unable to identify specific activities or strategies that we implemented to effectively do so. However, we accepted the NBPTS' endorsement of SA as indication of its positive impact. Consequently, I began investigating strategies in to develop and support students' SA skills.

In order to self-assess, students need to better understand the assessment process, a practice that has traditionally been left solely to teachers (Clayton & Shores, 2015). Standards-based grading (SBG) practices are known to increase student understanding of grades and grading criteria (Hochbein & Pollio, 2016; Iamarino, 2014), but limitations of the traditional grading system restrict full implementation of SBG practices in some contexts. Thus, this research seeks to determine whether certain SBG practices can be implemented in a traditional classroom in order to support SA. This action research study will employ a mixed methods case study design over the course of twelve instructional weeks.

The following research questions frame this practical action research study:

1. What impact do standards-based grading practices have on Secondary Math 1 Honors students' ability to *accurately* self-assess their performance on assessments items correlated to unique content standards?
2. What impact do standards-based grading practices have on the *quality* of Secondary Math 1 Honors students' self-assessment of their understanding of the content standards?
3. What impact do standards-based grading practices have on student *achievement* in Secondary Math 1 Honors?

During this action research study, four identified SBG practices were implemented in a secondary level Mathematics 1 Honors classroom to determine the impact on the accuracy and quality of students' SA. The following section outlines the purpose of this endeavor, along with the development of the theoretical framework of this research. The remainder of this chapter reviews literature related to SA and SBG. The history, development, impact, and relevant research related to each construct is reviewed. Suggestions for and the impact of implementation of both SA and SBG is discussed.

Purpose of the Literature Review

A review of related research is needed in order to better understand the research constructs (Machi & McEvoy, 2016). The purpose of this research is to determine effective strategies for supporting SA; it is constructive to examine the impact of supporting SA in order to discern whether this is a worthwhile endeavor. Further, it is important to examine relevant research in order to examine existing strategies that have already been identified and discussed. Although there is evidence that SBG practices support student understanding of grades and grading criteria (Hochbein & Pollio, 2016; Iamarino, 2014), more information is needed about SBG to determine its associated practices, how to effectively implement these practices, and the known impact on implementation on student thinking and learning.

By engaging in a literature review, researchers develop their understanding of the research topic (Machi & McEvoy, 2016). Reviewing related research supports the development of my theoretical framework, which relates theories on SA and SBG in order to support the research thesis. The review of literature consists of a three-part process for both SA and SBG; the history and development of related theories are examined, suggestions for implementation are reviewed, and the impact of implementation is discussed. By examining these three aspects of each construct I develop a well-rounded perspective of both SA and SBG as separate theories, as well as how they are related.

Research was conducted using the University of South Carolina's online article database, Google Scholar, and ERIC. Sources for this literature review include peer-reviewed journals, published dissertations, websites, and books.

Theoretical and Historical Foundations

Several learning theories and instructional strategies inform this study and construct the theoretical framework guiding this research. The practice of SA is examined to determine strategies that support its implementation and impact. Self-assessment requires students to reflect on their performance and understanding, a practice that inherently values a constructivist view of student perception of knowledge and knowledge creation; thus, constructivist learning theory shapes this researcher's view of student learning and instructional practices. Self-regulated learning theory is rooted in constructivist learning theory, arguing that students use metacognitive strategies to evaluate their learning process and motivation; SA plays a role in this process. The remainder of this chapter examines constructivist learning theory, self-regulated learning theory, SA, and SBG to explain the theoretical framework that drives this research.

Constructivist Learning Theory

Unsatisfied with the unique emphasis that traditional education placed on content and the abundance of freedom provided by progressives, John Dewey believed that educators needed to better understand the nature of human experience and its impact on learning (1938). Dewey (1938), who is often credited as being the philosophical founder of constructivism, emphasized the role that prior knowledge and personal experiences play when learners construct new knowledge. Dewey's theories paved the way for other constructivists. Bruner's (1962) theories extend on Dewey's, insisting that "education must also seek to develop the process of intelligence so that the individual is capable of going beyond the cultural ways of this social world...so that he can create an interior culture of his own" (p.116). Also influenced by Dewey, Piaget's (1964) Cognitive Development Theory outlines the four stages of child development and was unique in that it focused on the learning processes of children, rather than all learners. Piaget rejected theories that child thinking was a simplified version of adult process, instead suggesting that children think in ways that are drastically different than adults.

Though constructivist-learning theories differ in focus, there is a foundational guideline that unites them: all human knowledge is constructed. Rather than adopt pre-existing knowledge structures, learners develop their own interpretations and organization schemes for approaching problems (Shepard, 2000). Thus, students should be active participants in the learning process. Rather than transmitting knowledge to students, teachers must design and facilitate learning opportunities for students to construct their own understanding of content (Fernando & Marikar, 2017).

Student self-monitoring is a key element of knowledge construction, as students need to self-assess in order to know when they are learning (McMillan & Hearn, 2008). In order to construct new knowledge, students must determine whether new information aligns with prior knowledge. If new information contradicts prior knowledge, students must reconcile the differences (Shepard, 2000). This process requires students to organize, evaluate, and internalize new knowledge; SA is a natural part of these processes (McMillan & Hearn, 2008).

Constructivist learning theory views assessment as an integral component of the instructional process (Sharikzadeh, 2003). For assessment practices to support student learning, changes to the content, structure, and use of assessment data must be made (Shepard, 2000). Constructivist learning theory supports these changes; assessments must prompt students to make sense of their results by developing understanding of assessment criteria, which supports the construction of new knowledge. Self-assessment can serve as a vehicle to promote this practice and develop these understandings. However, strategies for supporting SA are ambiguous and need further examination. Self-regulated learning theory provides the framework to investigate and inform these strategies.

Self-Regulated Learning Theory

Self-regulated learners are confident, diligent, resourceful, proactive, and they are aware of what they do and do not know (Zimmerman, 1990). They plan, set goals, self-monitor and evaluate throughout the learning process. These behaviors help self-regulated learners maintain high self-efficacy, be intrinsically motivated, and decisive (Zimmerman, 1990).

Self-assessment is a core element of self-regulated learning (SRL) theory (Andrade & Valtcheva, 2009), an umbrella that covers the cognitive, metacognitive, behavioral, motivational, and emotional aspects of learning. Self-regulated learning theory consists of fourteen self-regulated learning strategies: self-evaluation, organization and transformation, goal setting and planning, information seeking, record keeping, self-monitoring, environmental structuring, giving self-consequences, rehearsing and memorizing, seeking social assistance, and reviewing (Zimmerman, 1990). Four of these strategies require students to self-assess in some capacity.

Zimmerman's original model of SRL has undergone several revisions, but many of the core elements of SRL are maintained. The most recent model, the Cyclical Phases

Model, includes three phases: forethought, performance, and self-reflection (Zimmerman, 2013). Self-assessment is an active element of the performance and self-reflection phases.

During the performance phase students monitor their progress while learning. This phase relies on two major processes, self-control and self-observation. Self-control refers to the techniques that students use to direct learning, self-observation refers to metacognitive monitoring processes that students employ throughout task completion to monitor their work (Zimmerman, 2013).

In the self-reflection phase, students self-assess learning to determine their success (Zimmerman, 2013). This phase includes two processes, self-judgments and self-reactions. Self-judgments assess the effectiveness of one's performance, strategies used during the performance phase, and prior learning experiences that may have affected their current performance (Zimmerman, 2013). Self-reaction refers to the learner's conclusion about whether or not they need to adapt their approach to learning during future cycles (Zimmerman, 2013). These processes subsequently inform student thinking during the forethought phase.

Processes in the forethought phase, task analysis, goal setting, and strategic planning, prepare students for learning (Zimmerman, 2013). The forethought phase initiates the learning cycle and informs student thinking and processes during the subsequent phases. The effect of these processes is heavily influenced by the performance and self-reflection phases. Superficial processes during the performance and self-reflection phases limits students' ability to critically analyze tasks and plan effective strategies. This further inhibits students' ability to create specific and challenging goals.

The Cyclical Phases Model attempts to distinguish between learners who are characterized by high quality processes during the forethought and performance phases, proactive learners, and those who rely on peer comparisons during the self-reflection phase to measure learning, reactive learners (Zimmerman, 2013). Although SRL theory posits that all learners self-regulate learning, proactive self-regulators exhibit a more effective and efficient cyclical pattern of processes than reactive self-regulators (Zimmerman, 2013).

Proactive self-regulators. Proactive self-regulators are characterized by higher self-efficacy beliefs, outcome expectancies, mastery learning goals, and a high value of tasks (Zimmerman, 2013). These qualities support superior functioning in the forethought phase, during which they develop strategies to complete tasks and establish specific, challenging goals (Zimmerman, 2013). Self-control processes are enacted during the performance phase to execute these strategies, relying on metacognitive monitoring and self-recording to guide students' work. Metacognitive monitoring is an informal tracking of performance and outcomes against the established goals; self-recording is the creation of formal records of learning processes and outcomes. Informed by higher functioning in the forethought phase, proactive learners are able to assess their performance based on their mastery of established goals. Students' judgments about their learning and errors are mostly causal in nature, and they are able to react adaptively (Zimmerman, 2013). This thinking during the self-reflection phase informs future task analysis, goal setting, and strategic phases. Thus, effective processes during the self-reflection phase result in superior processes during the subsequent forethought phase (Zimmerman, 2013).

Reactive self-regulators. Reactive self-regulators are characterized by lower quality functioning during the forethought phase. Lacking the self-motivation needed to engage in critical task analysis, these learners set vague, unchallenging goals for themselves, and fail to develop a strategy for task completion (Zimmerman, 2013). This limits their processes during the performance phase. Instead of metacognitively monitoring their progress and self-recording their processes, these learners engage in learning tasks without clear direction. During the self-reflection phase, reactive self-regulators are unable to evaluate their progress against previously established goals. Thus, if they do self-evaluate, reactive self-regulators are limited to comparing their progress against their peers in order to assume a level of personal growth (Zimmerman, 2013). Consequently, reactive self-regulators respond to errors defensively. Evidence of error leads these learners to feel dissatisfied with their performance and may lead them to disengage from their work (Zimmerman, 2013).

Next steps. Compared to reactive self-regulators, proactive self-regulators display many of the characteristics of independent, persistent problem-solvers (Zimmerman, 2013); thus, supporting the development of proactive self-regulatory practices is a valuable pedagogical endeavor. These learners are characterized by their divergent quality of processes during the forethought phase, which is informed by student reflection on their learning (Zimmerman, 2013). Supporting students' capacity to engage in higher quality reflection and assessment on their learning could develop processes necessary for effective forethought phase experiences. It is the goal of this research to determine whether specific assessment practices support students' ability to critically and accurately self-assess their learning and performance. Thus, the results of this research may suggest strategies that support the development of Zimmerman's (2013) proactive self-regulators.

Self-Assessment

The concept of SA permeates multiple learning theories associated with constructivism and has been defined using a variety of terms and behaviors (Andrade & Valtcheva, 2009; Leise, 2010; Panadero et al., 2015; Black et al., 2004; Stallings & Tascione, 1996; Mazloomi & Khabiri, 2018; Brown & Harris, 2014; Bruce, 2001). Leise (2010) discusses SA using the terms "reflection on performance" and "reflective learning" (p.66) to describe the process of reflecting on performance as a means to support learning. In their research with college and secondary level mathematics students, Stallings and Tascione (1996) differentiate between SA and self-evaluation by their respective scale of evidence of student learning. In their research, SA provides a local view of student learning, defining SA as the "process in which a student determines the types of errors made in her or his mathematics work" (p.548). Student self-evaluation offers a global view, defined as "the student's reflections about his or her general understanding of the mathematics explored up to that point" (p. 548). Andrade and Valtcheva (2009) also differentiate between assessment and evaluation, noting that SA is an iterative process, intended to inform progress and revision. Self-evaluation on the other hand, involves students determining a numerical grade for their work (Andrade & Valtcheva, 2009). Brown and Harris (2014) promote a more inclusive definition, defining SA as an "evaluation of a students' own work products and processes in classroom settings" (p. 22).

Recognizing the variety of intentions and terminology, other researchers have attempted to define SA using more structured classification systems. Panadero, Brown and Strijbos (2015) summarize five different SA typologies. These typologies classify and differentiate SA based on students' knowledge and interests, teacher involvement, the balance of power between student and teacher, the presence of assessment criteria, use of SA results, the SA format and its transparency (Panadero et al., 2015). Recognizing that SA is a multidimensional construct, Panadero et al. (2015) advise consideration of the purpose of conducting SA, the SA medium, time between instruction and assessment, student expectations, and learner access to assessment criteria, as these factors may impact the effect of SA on student performance.

Research and literature that involve SA reference a variety of overlapping and often conflated terms. Most notably, literature uses the terms self-reflection, self-evaluation, and SA interchangeably. Thus, this research attempts to clearly distinguish between these practices. In this study, the practice of self-assessing aligns with Leise's (2010) practice of reflection, and Stallings and Tascione's (1996) practice of self-evaluation. It does not however, align with Andrade and Valtcheva's (2009) practice of self-evaluation. Specifically, in this study, self-reflection refers to the reactive process of thinking about learning; reflection may occur at any time during the learning process. Students may self-assess while they reflect on their learning, but SA refers specifically to the process that students take when comparing their performance against established learning criteria. The purpose of SA is to improve future performance by identifying strengths and areas that need improvement. Students may reflect on their performance in order to gain insight, but students proactively self-assess in order to support growth in understanding against the assessment criteria.

Impact of self-assessment. Numerous benefits of supporting SA practices have been reported, including an increase in student achievement (Andrade & Valtcheva, 2009; Brown & Harris, 2014; Kelberlau-Berks, 2006; Leise, 2010; Noonan & Duncan, 2005; Panadero et al., 2015; Shepard, 2000; Stiggins, 2002; Stiggins & Chappuis, 2010), stronger of self-regulatory skills (Brown & Harris, 2014; Kelberlau-Berks, 2006; Panadero et al., 2015), increased motivation (Brown & Harris, 2014; Kelberlau-Berks, 2006; Panadero et al., 2015), improved self-efficacy (Brown & Harris, 2014; Kelberlau-Berks, 2006; Panadero et al., 2015; Shepard, 2000; Stallings & Tascione, 1996; Stiggins, 2002), and stronger student engagement in group tasks (Steinkruger, 2007). Studies also show that SA resulted in improved attitudes towards learning mathematics (Andrade & Valtcheva, 2009; Steinkruger, 2007).

Black, Harrison, Marshall, and Wiliam (2004) found that students who self-assess are more aware of when they are and are not learning. This awareness helped students identify when they needed additional support, time, and resources, and were able to better communicate their confusion and understandings (Black et al., 2004). Andrade and Valtcheva (2009), Steinkruger (2007), and Stallings and Tascione (1996) also concluded that regularly self-assessing their understanding supported student communication. When students self-assessed they were more willing to give attention to peer feedback, analyze the reasoning of their peers (Steinkruger, 2007), and use content appropriate vocabulary (Stallings & Tascione, 1996).

Supporting SA results in a better understanding of assessment criteria. Shepard (2000) found that when students self-assess, they become more concerned with grade criteria and feedback than with the grades themselves. These findings support Stallings and Tascione's (1996) observations that when they self-assessed, students began to view assessments as instruments for measuring their current understanding, and not as terminal measurements of their ability. Even when grades were low, or when they disagreed on assessment results, students expressed greater satisfaction and appreciation for the opportunity for feedback (Stallings & Tascione, 1996). Additionally, students began self-assessing their performance regularly, without prompting or guidance, and used SA results to inform their future study plans (Stallings & Tascione, 1996).

Though the benefits of supporting SA were reported for a variety of demographics, some are unique to specific population subgroups. Controlling for race and income, Andrzejewski, David, Bruening, and Poirier (2016) found that when non-white or low-income students learn to reflect on their science learning, they increased their rate of achievement and performed significantly higher on standardized assessments than their majority peers. Andrzejewski et al. (2016) hypothesize that SRL behaviors may already be associated with white, middle-class school culture, thus students outside of this culture may need explicit instruction and opportunities to develop these skills and benefit from their impact. Similarly, Black and Wiliam (as cited in Stiggins & Chappuis, 2010) found that student-involved assessment practices support low achieving students more than high achieving students. Lastly, engaging in regular SA supports language learning students' writing ability and language proficiency (Mazloomi & Khabiri, 2018).

Supporting self-assessment. Accepting the cognitive and affective impact of engaging in SA practices, it is important to review researchers' recommendations for how to effectively and efficiently support SA practices in classroom settings. Researchers agree that students need instruction on how to self-assess (Andrade & Valtcheva, 2009; Brown et al., 2015; Brown & Harris, 2014; Bruce, 2001; Collett, 2014; Hoernke, 2014; Mazloomi & Khabiri, 2018; Stiggins, 2002; Zimmerman, 1990). Teachers should model SA behaviors and provide regular feedback on SA, transitioning from simple, concrete SA techniques before incorporating complex techniques that include holistic judgments of performance (Brown & Harris, 2014). SA practices should be supported throughout the learning cycle (Brown & Harris, 2014), and should be used only with formative assessments (Brown et al., 2015). SA should not be associated with summative assessments, as SA is most effective when associated with an opportunity to improve or revise.

Students should be given the opportunity to self-assess as a means to improve, or self-correct, their work (Bruce, 2001), as students will not value SA unless it leads to an opportunity to improve performance (Andrade & Valtcheva, 2009). Excellence in any area requires a cyclical process of refinement; by providing students the opportunity to revise their work communicates a faith in students' capability for success and gives students greater control over their performance (Bruce, 2001). Opportunities for revision must be intentional and well communicated to students and include goal-setting and post-performance reflection components (Bruce, 2001). When establishing goals, students evaluate their needs and ability to create specific plans. When reflecting on their performance students identify strengths and weaknesses and their performance, which informs later goal-setting opportunities (Bruce, 2001). Through these experiences students learn how to improve their work, the final element of SA (Bruce, 2001).

In order to accurately self-assess, students must have a clear understanding of the assessment criteria (Andrade & Valtcheva, 2009; Brown et al., 2015; Bruce, 2001; Hoernke, 2014). Otherwise students may self-assess their performance based on their effort, their interpretation of their performance against their peers, or optimism of their ability, rather than their ability (Brown et al., 2015). Bridging explicit connections between standards, activities, and rubrics, facilitates student understanding of assessment criteria (Hoernke, 2014), thus it is not surprising that the use of rubrics is heavily supported (Andrade & Valtcheva, 2009; Brown et al., 2015; Bruce, 2001; Hoernke, 2014). Bruce (2001) further recommends having students translate standards into their own words to produce goals that are specific and appropriately challenging.

Effective SA is difficult to support in conjunction with traditional, teacher-centered instruction, within a competitive classroom culture, or exclusively with traditional, standard assessment items (Butler & Lee, 2010). Changes to the classroom structure are necessary, beginning with increased student involvement in the assessment process (Stiggins & Chappuis, 2010). By inviting students to participate in establishing assessment criteria, involving students in record keeping to monitor their progress, and allowing students to communicate about their performance with others, teachers give students authority over their assessment process and consequently their learning (Andrade & Valtcheva, 2009; Stiggins, 2002; Stiggins & Chappuis, 2010). This authority allows students to benefit from their assessment results, rather than view themselves as victims of it (Stiggins & Chappuis, 2010).

Validity of self-assessment. The attention given to supporting the accuracy of SA is dependent on the purpose of SA. If SA is used to motivate students to review their work against established criteria, the accuracy of their SA may not matter (Brown et al., 2015). However, if SA is used to guide revision and inform improvement, the accuracy of the SA is important; otherwise students may make decisions based on inaccurate information (Brown et al., 2015; Panadero et al., 2015). Leise (2010) found that improvement in the quality of SA is positively correlated with improvement in performance and improving the accuracy of SA supported student self-efficacy and transfer of learning. Thus, efforts to increase the accuracy of SA have cognitive and affective benefits for students.

Student SA reports vary in accuracy when compared to standardized assessment data and teacher assessment data. Brown et al. (2015) found that students consistently assessed themselves higher than their teacher's rating, with correlation between SA ratings and teacher ratings ranging from .48 to .98. They also found a weaker correlation between SA rating and standardized test results, ranging from .26 to .66 (Brown et al., 2015). Several factors relate SA accuracy and consistency, including age, ability, and product (Brown et al., 2015). Student age was negatively related to average SA score, and positively related to accuracy of SA. Students' academic ability was positively related to SA accuracy, and academically strong students self-assessed their performance more severely than low achieving students. Lastly, students were able to self-assess simple, concrete tasks with clear criteria more accurately than complex, abstract tasks (Brown et al., 2015).

The use of SA data has a significant impact on SA validity. Though SA may involve having students assign themselves grades for their work, many researchers discourage teachers from including these marks in the calculation of students' course grades (Black et al., 2004; Hoernke, 204; Leise, 2010; O'Connor, 2011). Leise (2010) found that when SA feedback is evaluative, SA accuracy decreased. Brown and Harris (2014) expressed similar concerns; sharing and using SA marks may reduce student honesty and assessment against the established criteria. Panadero et al. (2015) suggest that teachers allow students to initially keep SA results private to promote honest SA practices. However, as students benefit from teacher feedback on SA, eventually SA should be shared (Brown et al., 2015; Panadero et al., 2015).

Research indicates several practices that support the accuracy and validity of SA. Brookhart et al. (2004) and Stallings and Tascione (1996) found that the accuracy of SA improved with exposure and practice. Stallings and Tascione (1996) found that with practice, students increased their ability to identify and classify the type of error made in their mathematics work. Panadero et al. (2015) found that students who are involved in establishing assessment criteria are able to self-assess more accurately, and that SA accuracy improves when students receive clear feedback related to the accuracy of their SA. Providing rewards in conjunction with feedback may also increase SA accuracy (Brown et al., 2015).

It is important to note that not all measures of SA are valid with all populations, and that how well students self-assess with one method may not transfer to another SA method (Brown & Harris, 2014). For example, the Self-Regulated Learning I v.5 (SRLI V.5) is a tool used to identify student characteristics that result in academic success. The

SRLI V.5 has been successfully tested in a Midwestern university with predominantly white students. However, Goodpasture, Reinhardt, and Thomas (2007) found evidence that this pattern did not hold when the SRLI V.5 was used with African American students.

Next steps. Research on supporting SA agrees on several key aspects, which inform this action research study. First, expectations must be clearly articulated (Andrade & Valtcheva, 2009; Brown et al., 2015; Bruce, 2001; Hoernke, 2014). Students must have a clear understanding of the assessment criteria, as well as what is expected of them when they self-assess. By involving students when establishing assessment criteria, and modeling SA behaviors for students, teachers can support progress towards these goals. Second, once students have self-assessed their performance, feedback on their SA is necessary (Andrade & Valtcheva, 2009; Brown et al., 2015; Brown & Harris, 2014; Bruce, 2001; Collett, 2014; Hoernke, 2014; Mazloomi & Khabiri, 2018; Panadero et al., 2015; Stiggins, 2002; Zimmerman, 1990). Feedback should reference students' goal, provide evidence of their performance in reference to that goal, and include a plan to reduce the gap (Stiggins, 2002). Lastly, SA is an integral element of formative assessment (Stiggins, 2002), and students should use SA results to inform revisions of their work (Andrade & Valtcheva, 2009; Black et al., 2004; Bruce, 2001). Thus, opportunities to revise their work and reassess are necessary.

In order for assessment data to be used to improve learning, we need to shift from assessment of learning (e.g. standardized summative assessments), to assessments for learning (Stiggins, 2002). Assessments for learning must involve students in the learning process by providing clear criteria throughout instruction, translate assessment results into discernable feedback, and use assessment results to increase student independence and ownership over their learning (Stiggins, 2002); practices that either support or are supported by SA. However, little research exists on how to effectively support SA, and specific strategies are needed (Noonan & Duncan, 2005).

Standards-Based Grading

The standards-based reform movement originated in the mid-nineteenth century when urban schools adopted systematic grade level advancement practices (Hoernke, 2014). Students had to successfully complete formal examinations based on a uniform curriculum in order to be promoted to the next grade level. As a result, instruction and testing became more organized, but the need to evaluate student performance varied widely between teachers. In the early twentieth century, two researchers, Starch and Elliot (as cited in Hoernke, 2014), collected student examination papers in English, Math, and History. Teachers in over 200 schools marked this student work and the resulting scores were compared. On a 100-point scale, grades awarded to the English papers had a range of forty points, and one of the history papers had a range of ninety-two points, evidence of wide variability in assessment practices. Starch and Elliott (as cited in Hoernke, 2014) concluded that the wide discrepancy in assessment scores diminished the equity and validity of teacher assessment practices. Their findings began discussions about the benefits and challenges of norm-referenced assessment practices (traditional grading) and criterion-referenced practices (SBG).

Towards the end of the twentieth century the federal government began using standards and high stakes assessments as a means of holding schools accountable.

President George H. W. Bush (1993) expressed support for school accountability reform efforts in *America 2000: An Education Strategy*. President Bush called for improved standards in core subjects, and public report of assessment data evaluating student achievement against these standards. Later, President George W. Bush further promoted school accountability with the enactment of the No Child Left Behind Act of 2001 (NCLB) (U.S. Department of Education, 2002). In addition to publicly reporting school accountability measures, NCLB established high goals of continuous improvement for American public schools (U.S. Department of Education, 2002). President Obama's 2009 Race to the Top (RTT) competition continued this trend by encouraging states to adopt common standards and evaluate teacher performance based on student outcomes on high stakes assessments (U.S. Department of Education, 2010).

Federal legislation like NCLB and RTT increased attention on high stakes testing and adoption of standards-based instruction, but there has been insufficient discussion of how assessment practices and grade reporting should change (Clayton & Shores, 2015). Though grades were originally intended to signify students' level of understanding, they have evolved to serve as gateways in students' educational career. Strong marks are needed for admittance into competitive academic programs, thus norm-referenced grading practices promote competition between students (Black et al., 2004). This has a positive effect on high achieving students, but victimizes low achievers (Black et al, 2004; Stiggins & Chappuis, 2010). Extrinsically motivating factors and competitive classroom cultures develop feelings of hopelessness in low achieving students, further disconnecting students from the educational process.

Efforts made to reduce the achievement gap between white and non-white students by focusing on students' self-concept have minimal long-term effect. Boosting student self-concept provides temporary increases in student affect, but the feelings of hopelessness are reinstituted when low performance recurs (Stiggins & Chappuis, 2010). In order to suspend this cycle, students and teachers need to redefine success as continued improvement instead of immediate high achievement (Stiggins & Chappuis, 2010). Grading practices that increase student ownership over the assessment process are also needed to give students a feeling of control over their learning and understanding of assessment criteria.

Bradbury-Bailey (2011) describes two orientations for interpreting grades: the mastery goal orientation (MGO) and performance goal orientation (PGO). Mastery goal orientation motivates achievement focused on mastery and competence, thus developing intrinsically motivating factors. Conversely, PGO motivates achievement focused on achievement relative to others, which develops extrinsically motivating factors (Bradbury-Bailey, 2011). Federal legislation that mandates standardized assessments contributes to the development of PGO, which results in competition between students where grades are used as extrinsic motivators (Bradbury-Bailey, 2011). To increase the value of grades and make them more meaningful for students, teachers must promote a MGO of the grading process. Standards-based grading is offered to provide a framework to shift from PGO to MGO.

How to implement standards-based grading practices. Assessing student learning serves two purposes: to inform classroom decisions and motivate learners (Stiggins, 2002). Standardized testing for school accountability accomplishes neither purpose, instead legitimizing legislators and policy planners to make decisions. Involving students in the assessment process empowers students, creating feelings of control and greater understanding of assessment criteria and meaning (Iamarino, 2014; Stiggins & Chappuis, 2010). Whereas traditional grading practices position teachers as the rule-setter and expert, SBG practices allow students to actively participate in the construction of grading criteria, and subsequently the construction of knowledge (Iamarino, 2014). When students design assessment criteria they demonstrate knowledge of content, as well as variance in difficulty (Marzano & Heflebower, 2011).

Effective implementation of SBG practices provides students choice and authority over their assessment schedule. Students should be given the opportunity to assess when they feel ready, as well as the opportunity to reassess when they have grown in their understanding (Clayton & Shores, 2015; Colby, 1999; Marzano & Heflebower, 2011). This practice ensures that grades reflect student mastery of learning objectives, rather than the culmination of grades that accrue throughout the learning process. Even more, allowing students to reassess shifts student perception of grades; rather than seeing grades as terminal measures of ability, grades are opportunities for feedback to inform future growth (Colby, 1999).

Grades should only measure student understanding of the learning objectives. However, grading practices often reward or punish students for other behaviors, including attendance, punctuality, neatness, effort, and efficiency in following directions (Hoernke, 2014). This practice leads to inconsistent grading between teachers, which limits the validity and reliability of grades. Standards-based grading practices reduce this conflict. Teachers should eliminate "omnibus grades" (Marzano & Heflebower, 2011, p.34), only scoring student work in relationship to specific measurement topics. All assessment items should be clearly linked to aligned standards (Colby, 1999; Iamarino, 2014). This practice has a positive impact on pedagogy as well. When teachers explicitly link assessment items to standards, they must reflect on their perception of the standard, giving consideration to the type of knowledge that would provide evidence of mastering the standard (Hoernke, 2014). This reflective teacher practice provides unique benefits to classroom instruction, independent of the benefits to student understanding.

Targeted teacher feedback on student progress is essential. Unclear grading criteria lead to grading inconsistencies between teachers, which increases students' confusion over grading practices. Providing clear feedback on assessment items that are explicitly linked to standards informs students of their progress and areas for improvement.

Marzano and Heflebower (2011) acknowledge that traditional grading systems are unable to reflect student progress towards mastering objectives at increasing levels of difficulty. For example, a student may earn 90% of points awarded on a simple assessment on a particular learning objective. Later, this student may earn 80% of points awarded on a more challenging assessment on the same learning objective. Communication and use of these grades do not reflect potential growth in student understanding. Thus, effective adoption of SBG practices requires teachers to create a grading scale that informs student mastery at varying levels of complexity (Marzano & Heflebower, 2011). Rather than changing the grading scale, Iamarino (2014) incorporates code words for four different levels of mastery: beginning, approaching, meeting, above. These terms inform student progress in relation to each standard at various points throughout the learning process. Colby (1999) adopted a similar system, noting that when students are allowed to select and reassess standards at different levels of difficulty, their perception of assessments shifted. Students began to see assessments as a continuation of learning and instruction.

Impact of standards-based grading practices. Several positive effects of implementing SBG practices have been reported. Clear alignment of assessment items to standards increases student understanding of assessment criteria, which helps students identify their strengths and weaknesses (Iamarino, 2014). Colby (1999) and Iamarino (2014) reported that increased understanding of assessment criteria supports transition from a teacher-centered classroom to a student-centered classroom based in constructivist learning theory; increased understanding of criteria empowers students to take control of their learning, as students begin to view assessments as an opportunity for feedback and growth. SBG practices help students develop a better understanding of assessment expectations, rubrics, and the alignment between instructional periods, activities, and assessment items; consequently, assessments are perceived as a continuation of instruction (Colby, 1999).

Bradbury-Bailey (2011) determined that this effect is maintained with different student populations. Bradbury-Bailey (2011) compared the academic achievement of African American students in a traditional classroom against African American students in an SBG classroom. This research was motivated by a perceived disconnect between African American students and the educational process. In the SBG classroom, students scored higher on the state standardized end of course (EOC) assessment; the correlation between students' grades and the score they received on the EOC assessment was stronger in the SBG class as well. In the traditional classroom, students' course grade and EOC score had a correlation coefficient of .421. In the SBG classroom the correlation coefficient was .713. Additionally, students' perception of assessments shifted from being teacher-controlled to being self-controlled, suggesting a connection between SBG practices and mastery goal orientation (Bradbury-Bailey, 2011).

Adoption of SBG practices increases intrinsic motivation that is mastery oriented, which strengthens students' feeling of self-efficacy (Clayton & Shores, 2015). Students no longer view their confusion as lack of learning, but instead as a starting point for new learning (Iamarino, 2014). Providing specific, immediate, targeted feedback to students supports their self-monitoring and self-reflecting behaviors. Standards-based grading practices "create an environment where students are empowered and encouraged to be self-monitoring" (Clayton & Shores, 2015, p.172).

Standards-based grading practices help teachers manage students' needs throughout the learning process. Clearly aligning assessment items at varying levels of difficulty allows teachers to efficiently determine which students need additional support on specific learning targets, thus facilitating quick, customized remediation (Colby, 1999). Standards-based grading practices also enables emphasis on more important standards; teachers are able to see how frequently they assess each standard and can organize assessment data to include the type of assessment given (Colby, 1999). This allows teachers to ensure that they adequately assess, and reassess, the most important learning goals using a variety of assessment types throughout the year. Lastly, SBG eases communication with parents and students about student achievement (Colby, 1999). Grading practices are comprehensible and reflect student mastery of explicitly stated learning goals, thus students and parents receive clear communication about grades and student progress.

Additional benefits to teaching practices are reported. O'Connor (2011) posits that once a school fully commits to SBG practices, other improvements and evolution of teachers' assessment practices will follow. Teachers' focus will shift from behavioral issues such as attendance, assignment tardiness, or academic dishonesty, to student mastery of standards (O'Connor, 2011). Responsibility for achieving mastery will shift as well. Traditional grading practices put teachers in control of assessments; SBG practices transfer this control to students. By empowering students to take ownership over their learning and assessments, teachers are able to focus their attention on other aspects of instruction.

Relevant research and scholarly articles on SBG practices echo common elements and the impact on student learning and behavior. Though there are many identified benefits of adopting an SBG system, traditional grading requirements dictated by district and school policy limit teachers' capacity to fully adopt an SBG system. As school and district leaders mandate elements of teachers' assessment practices, administrative support is needed for full SBG adoption.

Four identified standards-based grading practices. Due to school and district assessment policies, complete adoption of SBG practices is limited. Review of related literature revealed four SBG practices that are consistently included in research studies and articles related to SBG; these practices also meet district and school assessment practices for approved implementation in this action research study. These four SBG practices are:

1. All assessment items are explicitly aligned to unique content descriptors.
2. All content descriptors are assessed at three tiers of difficulty (*proficient, approaching mastery, mastery*).
3. Students are able to self-select the assessment difficulty level for each content descriptor.
4. Students are able to improve the score they receive on all assessments, choosing to take assessments at a higher difficulty level to reflect their growth in understanding.

Next steps. The adoption of systematic educational practices increases focus on standards. Grades communicate student mastery of these standards, and are used to determine grade level promotion, measure school performance, and evaluate student preparation for competitive academic programs. Thus, educators must implement valid and reliable assessment practices that communicate student progress against established learning criteria. Evidence suggests that SBG practices may accomplish this goal. Standards-based grading practices provide students authority over their grades by inviting participation and allowing choice in the assessment process. Further, SBG practices communicate students' current level of mastery exclusive to the learning criteria, rather than a cumulative report of student performance over time that include measures of student behavior. The benefit of implementing SBG practices to student understanding of assessment practices may align with the conditions needed to support the accuracy and quality of students' SA. This possibility establishes the theoretical framework motivating this research.

Summary

Supporting students' ability to accurately and critically self-assess their learning and performance has been identified as a valuable pedagogical goal, but there exists a need to determine specific strategies that develop these processes. Review of existing research and related literature reveals several recommendations. Teachers should model SA processes, provide opportunities for students to revise their work or reassess, and efforts should be made to shift instructional practices towards a student-centered classroom. Additionally, in order to self-assess, students must have a clear understanding of the assessment criteria and should be involved in the assessment process. Though some recommendations to support SA are intuitive and easy to implement, strategies that promote understanding of assessment criteria and practices that involve students in the assessment process require exploration and testing.

Standards-based grading is a series of assessment and reporting practices based on student demonstration of mastery of defined learning objectives. Fundamental SBG practices include clear alignment between assessment items and standards, the ability to demonstrate understanding of standards at varying levels of complexity, and the ability to demonstrate growth in understanding by revising or re-assessing. Evidence suggests that SBG grading practices provide students with a greater understanding of assessment criteria, increased motivation and feelings of self-efficacy, and a feeling of control over the assessment process.

Supporting SA requires teachers to develop strategies that involve students in the assessment process and strengthen student understanding of assessment criteria. Research suggests that incorporating SBG practices may accomplish these goals. Thus, this action research study hypothesizes that implementing the four identified SBG practices will support students' ability to accurately and critically self-assess their learning. Additionally, this will have a positive impact on student achievement. The following chapter outlines the methods used in this mixed method, action research study to test this hypothesis and answer the research questions.

Chapter 3:

Research Design and Methods

Overview of Study

Students who monitor their learning and self-assess their understanding are better equipped to construct their own knowledge (McMillan & Hearn, 2008) and to become proactive self-regulators (Zimmerman, 2013). Supporting students' ability to self-assess their performance and understanding has also been shown to increase student achievement, motivation, and self-efficacy (Andrade & Valtcheva, 2009; Brown & Harris, 2014; Kelberlau-Berks, 2006; Leise, 2010; Noonan & Duncan, 2005; Panadero et al., 2015; Shepard, 2000; Stallings & Tascione, 1996; Stiggins, 2002; Stiggins & Chappuis, 2010). In the fourth component for certification, the National Board for Professional Teaching Standards (NBPTS, 2016) endorsed the practice of self-assessment (SA) by asking National Board candidates to describe instructional strategies or activities that support SA (NBPTS, 2016). In my experience as a National Board candidate support provider, I found that teachers, myself included, had difficulty identifying specific, intentional, practices that support SA. However, we accepted the NBPTS' endorsement of SA as evidence of its positive impact on student learning. As a result, we began to brainstorm instructional strategies that could support SA.

The purpose of this mixed method, action research case study, was to determine whether four Standards-Based Grading (SBG) assessment practices support students' ability to critically and accurately self-assess their performance on formative assessments. This study introduced four SBG practices in a ninth grade, Secondary Math 1 Honors class, and measured their impact on students' ability to accurately and critically self-assess their performance on formative assessment items, as well as the impact that these practices and behaviors had on student achievement.

Research Questions

The following research questions frame this practical action research study:

1. What impact do standards-based grading practices have on Secondary Math 1 Honors students' ability to *accurately* self-assess their performance on assessments items correlated to unique content standards?
2. What impact do standards-based grading practices have on the *quality* of Secondary Math 1 Honors students' self-assessment of their understanding of the content standards?
3. What impact do standards-based grading practices have on student *achievement* in Secondary Math 1 Honors?

Research Design and Intervention

Described by a variety of terms and lacking a concise definition due to its expansive application; action research is a deliberate, cyclical, reflective research process in which the research is conducted by insiders, usually within the researcher's local context (Herr & Anderson, 2015; Merriam & Tisdell, 2016). Whereas traditional researchers typically conduct their research by imposing carefully constructed treatments within a controlled setting, action researchers engage as participants in the research setting. Focusing more on local application than global theory, action research is a process of inquiry, usually initiated to solve an immediate local problem (Herr and Anderson, 2015). Action research is a reflective cycle that provides teachers an opportunity to engage in a process of professional-inquiry, promoting professional learning and increasing the potential to improve student learning (Ryan et al., 2017).

This practical action research study was motivated by a desire to support the development of students' ability to self-assess their knowledge and performance, while also considering the impact that these practices have on student achievement. In this study, SA is the process that students take when comparing their performance against established learning criteria. Students self-assess for the purpose of improving future performance by identifying strengths and areas that need improvement. The "plan-act-observe-reflect" (Merriam & Tisdell, 2016, p.54) cycle of action research allowed this educator-researcher to measure the impact of new assessment practices as an insider in my immediate setting (Herr & Anderson, 2015). As the research questions stemmed from my personal practice, utilizing an action research methodology enabled me to engage in a systematic cycle of inquiry to determine the impact of new pedagogical strategies.

Throughout the research cycle I engaged as a reflective practitioner, conducting this research for the purpose of informing my personal practice by identifying strategies that support student learning. Though traditional research methods require the researcher to be removed from the research environment, action research is participatory (Herr & Anderson, 2005). As the main purpose of action research is to empower teachers and improve classroom practices, an action research model was appropriate for this study (Mertle, 2014).

This practical action research study took place in a ninth-grade mathematics classroom over the course of twelve instructional weeks. This research occurred in two cycles. The first cycle lasted four weeks, during which time students' baseline data was collected and analyzed. The treatment was implemented over the course of eight weeks. All assessment and instructional practices were maintained throughout the twelve weeks, with the exception of the treatment intervention.

This twelve-week research cycle takes place during the spring semester (terms three and four) of a year-long Secondary Math 1 Honors class. This class meets every other day for ninety minutes. As the research cycle occurs during the twelve weeks prior to administration of the end of course standardized assessment, the treatment was implemented as students reviewed standards that were taught previously during the school year. The Mathematics Vision Project (2016) curriculum for Secondary Math 1 Honors was used to motivate and facilitate student learning throughout terms one and two. Though instructional strategies varied from day to day, similar instructional methods, grounded in constructivist learning theory, were primarily used to develop student understanding of the content standards.

Twelve Common Core Mathematics standards (CCSM, 2010) (Appendix A) were selected from instructional units that span the first two terms of the academic year. Standards were numbered and randomly divided into two groups of four (used in baseline cycle) and eight (used in the intervention cycle) using a random number generator. The group of four standards was assessed during the first four weeks to establish students' baseline data; the second group of standards were assessed during the final eight weeks of treatment. Students completed one formative assessment correlated to a unique standard each week. In addition to completing the formative assessment, students also reflected on their performance on each assessment immediately after its completion by completing an assessment reflection. This practice was maintained throughout the twelve weeks.

While the review of standards process was held constant, four SBG practices were introduced in Weeks 5-12. The four SBG practices were 1) alignment of formative assessment items to content standards, 2) opportunity to be assessed at varying levels of difficulty, 3) student choice of assessment difficulty, and 4) opportunity to improve assessment results. These particular practices were selected based on their consistent appearance in research and articles related to SBG (Clayton & Shores, 2015; Colby, 1999; Iamarino, 2014; Marzano & Heflebower, 2011), as well as this researcher's ability to effectively implement these practices in my academic setting, which utilizes a traditional grading system.

As the research questions focus on both the impact that SBG practices have on SA, as well as student achievement, a concurrent triangulation mixed method research design is appropriate (Efron & Ravid, 2013). Qualitative data is used to understand how students' process their experience (Merriam & Tisdell, 2016), thus qualitative data was used to collect and analyze the quality of students' SA of their performance and position within the learning cycle. Quantitative methods were used to analyze and assess the accuracy of students' SA, as well as the overarching impact on student achievement.

When teachers engage in inquiry, they develop the inner curriculum, further developing authority over their practice and learning (Brubaker, 2004). As a member of the mathematics department at my school, I meet weekly with other teachers in curriculum teams to discuss instructional strategies, plan common assessments, and analyze student achievement data. This professional learning community (PLC) analyzed the impact of instructional strategies and activities on student understanding of content and development of the eight mathematical practice (CCSM, 2010). Although I am engaging in this action research independently, I continued to analyze evidence of student learning within my PLC. This experience enabled me to analyze and compare student achievement data to other classrooms where the SBG treatment was not implemented.

Participants

This intermountain urban school district serves a diverse student body of approximately 24,000 students. This district consists of twenty-six elementary schools, five middle schools, five high schools, three charter schools, three community learning centers, and one school that serves students in kindergarten through eighth grade. Statistics taken from their public website indicate that this district serves a diverse student population, consisting of 43.8% Caucasian, 36.9% Hispanic, 4.4% Asian, 4.4% Pacific Islander, 4.3% African American, 4.2% multi-ethnicities, and 1.1% Native American students. Approximately 59.95% of students are identified low income, serving a total of 13,592 students free or reduced lunch. Approximately 388 students are English language learners, with over 37 languages being spoken in the homes of students across our district.

My high school serves 1686 students in grades nine through twelve. Our student body consists of 57.7% Caucasian, 24.1% Hispanic, 5.2% multi-ethnicities, 4.5% Pacific Islander, 3.9% African American, 3.4% Asian, and 1.3% Native American students. Approximately 41.25% of students are identified low income, and 44.4% of students receive free or reduced lunch. There are 84 teachers, 5 counselors, 4 members of the administrative team, and 9 support staff that serve our students.

Participants for this study were selected using a convenience sampling method. Of the six classes and 172 students I teach, I selected one class of 34 Secondary Math 1 Honors students. This particular class was selected because of its larger size, range in student ability as evidenced by student performance during the first two terms of this academic year, as well as its diversity. This class originally consisted of 36 students; two students withdrew from school for medical reasons, and another student withdrew in order to take a lower level math course. A new student was added to this course one month prior to beginning this research. The parents of one student elected not to include their student in this study. This student, student 9, was removed from the data analysis.

This group of 33 research participants consists of 20 female and 13 male students; 24 students are Caucasian, 7 Hispanic, 1 African American, and 1 Pacific Islander. Five students speak Spanish as their first language and continue to speak primarily Spanish when outside of school; these five students are also proficient in English. Only one student currently receives formal language services at school as an English language learner (ELL). Seven students are enrolled at a sister charter school for the performing arts, only attending our school for their core academic courses. Students in this honors course have not been objectively identified as gifted and talented; rather their placement in this honors course is based on teacher recommendation and parent request.

The research participants display characteristics that may conflict with the research agenda. Academically motivated students are typically driven by achieving high marks; for many students in this class, developing concrete understanding of content is often perceived as a lesser priority. This characteristic could impose a challenge as participants may resist selecting tasks that align with their ability but result in a lower academic mark. However, it is important that tasks with varying levels of complexity result in different academic marks as participants may not challenge themselves to choose more difficult assessments if they have the opportunity to earn the same mark for completing an easier task. Conversely, participants are academically competitive with one another, often sharing and comparing assessment marks. Some participants may select tasks that are more challenging, inaccurately self-assessing their level of understanding in order to compete with their peers. Miller (2013) proposes that when we release students from the chore of completing assignments to protect their final grade, keeping the ownership and goal setting in the student's control, that students can be trusted to want to learn. Thus, it was important that I articulate the purpose of task selection and reflection, incorporating these conversations into our class norms and behavior discussions during the research cycle.

While conducting this action research I engaged as an insider collaborating with other insiders (Herr and Anderson, 2015). I identify with the academic characteristics of the research participants, sharing ethnic and economic similarities with most participants as well. Though I do not share ethnic characteristics with all students, I do share the experience of immigrating to the United States with some students. Further, though I speak English as my first language, my husband immigrated to this country from a Spanish speaking country and learned to speak English as a second language while enrolled as a student in public school. Though I do not personally share his cultural experiences, my relationship with his family and understanding of his history as an

English language learner has helped me develop a relationship with my ELL students through a shared understanding and communicated value of their cultural identity.

A researcher is unable to eliminate all bias and subjectivity from their research, but by reflecting on, challenging, and articulating their positionality they can increase the validity of their results (Derry, 2017). By considering my positionality and the characteristics of the participants, I identified potential conflicts and refined my research design and treatment implementation. It was necessary that I continued this practice throughout the research process, critically re-evaluating my positionality and learning from the research experience.

Data Collection Measures, Instruments, and Tools

Five types of instruments were created to facilitate data collection: two selected response assessments, traditional formative assessment items (Appendix B), tiered formative assessment items (Appendix C), and an assessment reflection instrument (Appendix D).

Selected Response Assessments

The selected response assessments were created using response items that were written and administered by my school district. These assessment items are uniquely correlated to CCSM (2010) standards and depth of knowledge (DoK) learning levels (Webb, 2002). After the twelve CCSM (2010) content standards were randomly divided into the baseline and treatment groups, assessment items were grouped by standard to create the pre- and post-assessments. The first pre- and post- assessment consists of 12 selected response items that are aligned to the four CCSM (2010) standards assessed during the baseline cycle. The second pre- and post- assessments consists of 24 selected response items that are aligned to the eight CCSM (2010) standards assessed during the treatment cycle. No changes were made to the pre- and post- assessments between administrations at the beginning and end of both the baseline and treatment cycles.

Formative Assessments

The traditional and tiered formative assessments were written by this educator-researcher for the purpose of this research, though these items are modeled after questions previously administered on other formative and summative assessments. The traditional formative assessments, which include approximately two to five items, required students to demonstrate mastery of the aligned standard by creating and analyzing multiple representations of mathematical relationships and patterns. These assessments were administered once per week during the first four weeks, the baseline cycle. These items are correlated to unique CCSM (2010) standards and are DoK level two or three (Webb, 2002).

The tiered formative assessment items were administered once per week during the final eight weeks during the treatment cycle. Like the traditional assessments, the tiered assessments consist of two to five items that required students to demonstrate mastery of identified content standards by analyzing and creating varying mathematical representations. These assessments are also correlated to unique CCSM (2010) standards, but the tiered design correlates tier one assessment items to a DoK level one, tier two assessment items to a DoK level two, and tier three assessment items to a DoK level three or four (Webb, 2002). In addition to their tiered levels of difficulty, the tiered formative assessment items communicate their correlated CCSM (2010) standard at the top of each assessment item. This standard is also communicated to students prior to the assessment administration to inform student selection of their assessment tier.

Assessment Reflection Instrument

Lastly, the assessment reflection (AR) instrument (Appendix D) was developed by this educator-researcher for use in this action research study. This instrument was completed by students once a week during the entire twelve-week research cycle, immediately following the completion of a formative assessment during both the baseline cycle and the treatment cycle. Questions one and two of the AR provide data on the accuracy of students' SA. These questions challenge students to self-assess their performance by rating the difficulty of the question on a Likert scale, and by estimating the number of points that they earned on the formative assessment item out of a given number of total points. Questions three and four provide data on the quality of students' SA and were piloted with other student groups and revised during the 2017-2018 and 2018-2019 academic years. Students responded to either question three or four, based on their assessment of their performance on the formative assessment. This instrument was revised during the research cycle based on this researcher's observations of student behaviors that revealed insight into student thinking about SA. Both the original AR instrument, and the revised instrument, are included in Appendix D.

Research Procedure

I received consent from the parents and students to be included in this study prior to conducting research. Both parents and students received a parental consent form

(Appendix E). To protect the rights and privacy of research participants, each participant was assigned a number so that no identifying information is shared. All personal identifying information was removed from all assessments and assessment reflections and replaced with the corresponding identification number prior to data storage. No data that was stored, analyzed, or reported, was associated with students' personal identifying information. The names of the research participants and their associated number is only accessible by the educator-researcher. A hard copy of the data was stored in the researcher's classroom, and an electronic copy was stored on the researcher's portable hard drive.

At the beginning of the twelve-week research cycle students were informed that they would complete one formative assessment item per week to review content learned in the previous semester, in preparation for their end of course assessment. The AR instrument was introduced and discussed with students in a small group and then whole group setting prior to administration. Students first reviewed the AR instrument in their small groups, and then asked clarifying questions in a whole group format. After discussing the AR instrument, students independently completed the first pre-assessment to establish their baseline understanding of the selected standards. Students did not receive any feedback on their performance on this pre-assessment.

During the first four-week baseline cycle, students completed one formative assessment item per week. The first formative assessment and AR was completed during the class period that immediately followed the AR introduction and pre-assessment administration. Students received written feedback on all formative assessments and were given five to ten minutes of class time to review their written feedback independently and ask clarifying questions. Written and verbal feedback was provided individually and was not discussed in a whole group setting. After reviewing their feedback, students returned their formative assessment items to the teacher-researcher; no formative assessment items left the classroom setting.

Students completed the first post-assessment during the final class meeting of the fourth week. Unlike traditional pre- and post-assessments, the pre-assessment was administered after the initial learning cycle for these standards was completed. Thus, any changes in achievement between pre- and post-assessment is evidence of the impact that the weekly formative assessment items, written teacher feedback on these assessment items, and engaging in the AR, had on student learning.

The treatment cycle began on the fifth week of the twelve-week research cycle. However, due to unanticipated changes to the class schedule, there was a one week break between the baseline cycle and the treatment cycle. Students completed the second pre-assessment during the first-class meeting of the fifth week of research, which began six weeks after the baseline cycle began. Students completed this pre-assessment independently and did not receive any feedback on their performance. After completing this pre-assessment, students were informed of the upcoming changes to the weekly formative assessment items.

Students were told that for the next eight weeks when they complete a formative assessment on a review standard, that they would choose the difficulty level of the formative assessment. Assessments were offered at tier 1, *proficient*, tier 2, *approaching mastery*, and tier 3, *mastery*, levels. If they selected a tier 1 level of difficulty, they would receive a score of at most 80% on this assessment. If they selected a tier 2 level of difficulty, they would receive a score of at most 90% on this assessment. If they selected a tier 3 level of difficulty, they were eligible to receive a score up to 100% on this assessment. Students were also informed that they could later choose to re-assess at a higher tier to improve their score. Opportunities to re-assess were offered twice a week after school, once per week at lunch, and once per week during weeks when this class met three times, which occurred every other week.

I did not initially intend to offer opportunities for students to retake tier 3 level assessments, however, there were two instances when students initially chose to complete a tier 3 assessment, and later wished to improve their score by reassessing again at a tier 3 level. In these cases I created alternative tier 3 assessments so that these students could reassess to demonstrate their improved understanding; I chose to deviate from my initial research plan to prevent students from strategizing over what tier to complete on their first attempt. During week six, a conversation with three high performing students revealed that these students were debating whether they should only complete Tier 2 assessments when first presented with the content descriptor, even if they felt capable of tackling a more challenging assessment. This conversation made me realize that students were prioritizing their academic marks on these assessments over their perceived understanding of the content descriptor; they viewed the tier 2 assessment as an opportunity to receive feedback before completing tier 3. This feedback, coupled with targeted studying, could ensure that they could earn a higher score when they take the tier 3 assessment at a later date. After consideration of their strategy, and review of Zimmerman's (2013) Cyclical phases model, I decided that the processes that students engage in when selecting their appropriate tier would be of greater benefit to their performance and self-assessment, if their strategizing related to the completion of the assessment, rather than their academic marks. For this reason, I decided to alleviate their concern by offering to create an alternate tier 3 assessment whenever requested. I made this opportunity available to all students prior to administration of formative assessment seven; students only took advantage of this opportunity twice (formative assessments five and seven).

Students completed one tiered formative assessment each week during the eight-week treatment cycle. Students were told the correlated standard for the tiered formative assessment item prior to administration and were given two minutes to silently consider which tier they would like to select. Students discretely indicated their selection by holding up 1, 2, or 3 fingers while the assessment was distributed to limit peer pressure to select a tier that did not align with their self-assessed competency. The verbiage for the correlated CCSM (2010) standard was also recorded at the top of each tiered assessment. Students responded to the AR immediately after completing their selected formative assessment.

Feedback was provided in the same manner during the treatment cycle as in the baseline cycle. Students were provided written feedback, which they read independently, and individually asked clarifying questions. To maintain assessment security, as students could later choose to reassess at a higher tier, no formative assessment items were discussed in a whole group setting, and no formative assessment items left the classroom. Students completed the final post-assessment during the last class period of the treatment cycle.

My initial research plan was to use the same assessment reflection instrument throughout the twelve-week research cycle. However, a series of experiences led me to make small changes to the writing prompts. During the week 6 formative assessment administration, I observed student 20 while they completed the assessment reflection. This student's behavior stood out to me, as I watched them flip their paper back and forth multiple times while responding to the question prompts. Their critical thinking about their performance was apparent as they estimated the number of points they earned and formulated their response. I recorded this observation in my researcher's journal and made a note to pay attention to students' behavior while completing the AR again the subsequent week.

During the week seven administration I noticed significant differences in several students' behavior while they completed the AR. Some students seemed to complete the AR very quickly, without pause. Others paused regularly or turned their page over to review their written responses to the assessment items. I interpreted the latter behavior as evidence that these students were thinking more critically about their performance than others and perceived their willingness to engage in this thinking as evidence that they valued the process of self-assessing their performance. Evidence of the varying levels of value that students have for engaging in self-assessing processes sparked my interest in the impact that this value may have on the accuracy and quality of their self-assessments.

In order to triangulate my perception of the value that students place on the SA experience, I asked students to complete an exit ticket following the week eight assessment administration. This exit ticket had two purposes. First, I wanted more evidence about the value that students place on self-assessment in order to better identify students who assigned low and high value. Second, I wanted to better understand the reasons why students valued SA differently in order to better understand how I could increase student buy-in.

The exit ticket (Appendix F) asked students to indicate how valuable they found the experience of self-assessing their performance, and to explain why they felt that way. In order to compare students' identified level of value to my perceived level of value, I required students to identify themselves when completing their exit ticket. I wrestled with this idea because I worried that some students may be less willing to answer honestly in this way. However, as part of my goal was to triangulate my observations, it was necessary that I could reliably identify students' responses. Students' responses to this exit ticket were compared with my observations in my research journal, and eight students were selected whose self-selected level of value seemed to align with my perceived level of value, which was based on my observation of student behaviors while completing AR six, seven, and eight. I chose these eight students to narrow the focus of my qualitative data analysis, providing two student representatives from each selected value level.

In addition to the exit ticket administration, I returned to my theoretical framework to determine strategies that I could implement to make the act of self-assessing their performance more valuable to students. Andrade and Valtcheva (2009) report that students value the process of self-assessing if they see it as an opportunity to improve future performance. Though students have the opportunity to improve their performance during the treatment cycle, this opportunity wasn't explicitly linked to the assessment reflection. For this reason, I decided to revise the writing prompts in questions three and four to highlight students' opportunity to reassess.

When analyzing the exit ticket responses, I found that the majority of students who selected a lower value of the self-assessment experience indicated that they believe that the experience may have value, but that they don't understand why they are asked to engage in these practices. Following Miller's (2013) recommendations, I attempted to articulate the purpose of the task selection and reflection routinely throughout the research cycle. However, students' feedback revealed that I hadn't adequately communicated this to students. In order to better communicate the purpose of the task selection and reflection, while also informing students of the changes to the writing prompt, I shared my reasons for asking them to engage in these reflective experiences prior to administration of the week nine formative assessment. I explained to students how the process of self-assessing relates to my beliefs about teaching and learning, and I highlighted how I hope the process of self-assessing helps them prepare for future assessments. I wondered aloud whether the self-assessment process could help them determine whether they intend to reassess, formulate a plan to reassess, or identify learning and problem-solving strategies that helped them during their performance today, so that they can remember them during a future performance. I made a similar, but more brief statement prior to the administration of the remaining formative assessment items in weeks ten, eleven, and twelve.

Treatment, Processing, and Analysis of Data

Data was collected and analyzed to measure the impact the imposed treatment had on three distinct constructs: the accuracy of students' SA, the quality of students' SA, and changes in student achievement. The first two questions on the AR (Appendix D) were used to collect data that addresses the first research question. This data was analyzed quantitatively using descriptive statistics. Question one measured students' perception of the question difficulty using a Likert scale. Question two measured students' accuracy, defined as the difference between the predicted and achieved score. The achieved score was determined by the researcher after marking students' work against a pre-established rubric. A negative measure of accuracy indicates that a student underestimated their learning, while a positive measure of accuracy indicates that a student over-estimated their learning. The closer the measure of accuracy is to zero, the more accurate the SA. The accuracy of students' SA was compared to their perception of the question difficulty, as indicated by their Likert scale rating in question one. These two measures were analyzed quantitatively to determine the impact of the treatment on the accuracy of students' SA.

Data measuring research question two was collected and analyzed using students' written responses to the narrative reflection questions three or four on the assessment reflection (Appendix D). Analysis of qualitative data was inductive and comparative, and occurred throughout the data collection period in order to inform qualitative design. At the beginning of the research cycle student responses were analyzed using an open coded method to identify emergent, tentative categories (Merriam & Tisdell, 2016). Throughout the twelve weeks, an axial coding method was also used to relate identified categories and develop overarching themes (Merriam & Tisdell, 2016).

The third research question was analyzed quantitatively using the pre- and post-assessments. Changes in student achievement between administrations of each pre- and post-assessment was analyzed to determine any impact that the treatment had on student achievement, as indicated by changes in their performance on these assessments.

Multiple data sources were used to increase the rigor and trustworthiness of the research results. A reflective journal was maintained by the researcher in order to maintain consistency, increase reflexivity, and reduce bias. Efforts were made to limit extraneous variables by maintaining consistent instruction and assessment practices throughout research cycle, introducing treatment during the final eight weeks. Multiple methods were used to collect quantitative data in order to support rigor between the quantitative measures and support trustworthiness of the research findings. (Merriam and Tisdell, 2016).

Summary

Improving students' ability to self-assess their performance has been shown to support multiple aspects of student learning. However, clear instructional strategies that develop self-assessing skills in students are unknown. The results of this practical action research study provide evidence whether four SBG practices support the accuracy and quality of ninth grade Secondary Math 1 Honors students' SA of their performance on formative assessment items. The overall impact that these practices have on student achievement was also measured.

Three types of instruments were created for the purpose of this study. Two pre- and post-assessments consisting of selected response questions were developed using assessments items created by my school district. An AR instrument was created to collect data that will be analyzed in response to research questions one and two. Lastly, a total of twenty-eight formative assessment items that are uniquely correlated to CCSM (2010) math standards and DoK learning levels (Webb, 2002) were created to assess student understanding of twelve CCSM (2010) math standards. During the baseline cycle students completed one formative assessment and one AR per week, correlated to four of the twelve standards. During the treatment cycle students self-selected one assessment item correlated to either a DoK (Webb, 2002) level one, two, or above, and each of the eight remaining CCSM (2010) standards. Students completed an assessment reflection, once per week, following the completion of these tiered assessment items as well. Changes were made to the assessment reflection during the ninth week based on student observation and research to support student buy-in of the self-assessment process.

Data collection focused on five forms of data that addressed the three research questions. Quantitative data was collected using the AR instrument and analyzed using descriptive statistics to address the first research question. The second research question required the analysis of qualitative data, which was also collected using the AR instrument. The qualitative data was coded first using an open coding method, and then an axial coding method once dominant themes and categories were established. The pre- and post-assessments administered at the beginning and end of the baseline cycle and treatment cycle were analyzed quantitatively in response to the third research question.

Although the purpose of this research is to inform my personal practice, the results of this study are applicable to secondary level mathematics teachers. Characteristics of the research participants should be considered when determining the transferability of the research results. Focus should be given to students who display characteristics similar to the research sample. Further research is needed to fully determine the impact of implementing these practices across different content areas

Chapter 4:

Presentation and Analysis of Data

Overview of Study

Research indicates numerous benefits of supporting student self-assessment (SA), including increased achievement, motivation, self-efficacy, and proactive self-regulated learning behaviors (Andrade & Valtcheva, 2009; Brown & Harris, 2014; Kelberlau-Berks, 2006; Leise, 2010; Noonan & Duncan, 2005; Panadero et al., 2015; Shepard, 2000; Stallings & Tascione, 1996; Stiggins, 2002; Stiggins & Chappuis, 2010; Zimmerman, 2013). Further, educators that subscribe to constructivist learning theory should support student SA, as students must be able to self-assess in order to determine when new knowledge has been constructed (McMillan & Hearn, 2008). Thus, it's not surprising that the National Board for Professional Teaching Standards (NBPTS, 2016) endorsed the practice of supporting student SA by asking National Board candidates to describe instructional strategies that support SA (NBPTS, 2016). In my experience as a National Board candidate support provider, I found that teachers, myself included, struggled to identify specific strategies and practices that support SA. Consequently, I began investigating instructional strategies and classroom practices that I could implement to support students' SA of their performance and understanding.

This study measured the impact that four standards-based grading (SBG) practices had on the accuracy and quality of students' SA in a Secondary Math 1 Honors class. The overall impact on student achievement was also analyzed. These four SBG practices were: all assessment items were explicitly aligned to unique content descriptors, all content descriptors were assessed at three tiers of difficulty (*proficient, approaching mastery, mastery*), students were able to self-select the assessment difficulty level for each content descriptor, and students were able to improve the score they received on all assessments by reassessing to a higher difficulty level to reflect their growth in understanding. Five types of instruments were created to facilitate data collection: two selected response pre- and post-assessments, four traditional assessment items, eight-tiered assessment items, and an assessment reflection (AR) instrument. The two pre- and post-assessments were administered at the beginning and end of each phase of the research cycle (baseline and treatment), and changes in student performance were analyzed to determine the impact that the treatment had on student achievement.

A convenience sampling method was used to select the 33 research participants who were enrolled in Secondary Math 1 Honors. I chose this particular class of students due to their larger class size, range in student ability, and diversity. This group of ninth grade students consisted of 20 female and 13 male students. Twenty-four of the research participants are Caucasian, 7 are Hispanic, 1 African American, and 1 Pacific Islander. All students are proficient in English, but five students speak Spanish as their first language. These five students primarily speak Spanish in their homes. Seven students are enrolled at a sister charter school for the performing arts, only attending our school for their core academic courses. Though these students are enrolled in an honors level course, they have not objectively been identified as gifted and talented.

Intervention/Strategy

The purpose of this mixed method, action research case study, was to determine whether four SBG practices support students' ability to critically and accurately self-assess their performance on formative assessments. Data collection occurred in two cycles. During the first four weeks of the research cycle, the baseline cycle, students completed one formative assessment per week and its accompanying AR. Traditional assessment practices were implemented during the four-week baseline cycle. During the last eight weeks of the research cycle, the treatment cycle, four SBG practices were introduced, and students continued to complete one formative assessment and AR per week. The accuracy of students' SA was analyzed quantitatively using the score they achieved on the assessment and their predicted score from the AR. The quality of students' SA was analyzed qualitatively using their response to either question 3 or 4 of the AR. Though student responses to questions 3 and 4 of the AR yielded qualitative data, changes in the frequency of qualitative codes over time was analyzed. A pre-and post-assessment was administered at the beginning and end of both the baseline and treatment cycles to measure changes in student achievement. Change in student achievement was analyzed quantitatively.

General Findings/Results

Five instruments were used to collect both qualitative and quantitative data, and each piece of data is uniquely aligned to one of the three research questions. Research question one addresses the accuracy of students' self-assessment, which was determined using their predicted score (AR question 2) and achieved score. The achieved score was determined by marking their assessment responses against a pre-established rubric. Research question two addresses the quality of students' self-assessment, which was analyzed using student responses to AR questions three and four. Lastly, research question three addresses changes in students' achievement, which was analyzed using student performance on pre- and post-assessment one and two. As there is no overlap between the data collected and the research question addressed by each data type, I will present and analyze the data in context of the corresponding research question, in sequence. Weekly analysis of the qualitative data and observation of student behaviors during the research cycle sequence heavily influenced my framework for analyzing both the qualitative and quantitative data. For this reason, I will first outline the qualitative data analysis and its impact on the research methodology, and then discuss influential experiences that occurred during the research cycle. I will then analyze the data in the context of the research questions by responding to research question two, followed by research questions one and three.

Qualitative Data Analysis

Using open coding and axial coding methods, this researcher-educator analyzed the qualitative data weekly throughout the 12-week research cycle until I reached data saturation (Merriam & Tisdell, 2016). Qualitative analysis revealed that student responses to questions three and four of the AR addressed one, or multiple, of three identified themes: level of understanding, strategies that support understanding, and plans to improve understanding.

Level of understanding. Student responses that addressed their level of understanding were divided into two categories; these responses either provided a reason for their level of understanding, or an assessment of their performance on the formative assessment. Within each category four codes were identified, as outlined in Table 4.1.

Table 4.1

Qualitative Coding Theme 1: Level of Understanding

Category 1: Reason for level of understanding		
Code	**Title**	**Example**
A	Gives generic reason to explain their level of understanding.	"I remember how to do it."
B	Gives specific reason to explain their level of understanding.	"It's easy to write an equation when the table is given to you."
C	Reason provides personal responsibility for learning.	"I like doing problems like these so I got good at them."
D	Reason implies external responsibility for learning.	"You haven't reviewed this with us in a long time."
Category 2: Assessment of performance		
Code	**Title**	**Example**
E	Identifies math specific competencies.	"I understand the pattern and writing explicit equations."
F	Identifies math specific challenges.	"I forget how to write a recursive equation."
G	States non-specific math challenges.	"I forget how to do this."
H	Identifies specific portion of task that is challenging but does not relate to the math learning goals.	"I think 1c is hard, but the rest was easy."

Strategies that support understanding. The second theme, strategies that support understanding, was divided into two categories: math-specific strategies, and non-math specific strategies. Within each category four codes were identified, as outlined in Table 4.2.

Table 4.2

Qualitative Coding Theme 2: Strategies that Support Understanding

Category 3: Math specific strategies that support understanding.		
Code	**Title**	**Example**
I	Pattern seeking.	"I understand equations better than pictures, but putting them together helped me see patterns better."
J	Analyzing multiple representations.	"Making a table helped me see the pattern."
K	Considers specific/simplified case.	"At first I thought I would have to just use trial and error, but by back-tracking to a simpler one I was able to make sense of this one."
L	Develops strategy by connecting to prior experience or analyzing the structure.	"This was confusing at first, but then I remembered the format of explicit equations and ones like this that I did already that were just a little different."
Category 4: Strategies that support understanding that are not math specific.		
Code	**Title**	**Example**
M	Memorization.	"It's easy for me to memorize a formula."
N	Practice or repeated exposure.	"We've done this a lot."
O	Classroom experiences (note taking/asking questions).	"I learned this through past lessons in this class."
P	Studying.	"I restudied this topic so it's fresh on my mind."

Plan to improve understanding. Lastly, the third theme, plan to improve understanding, encompasses student responses that identify a learning goal, and/or a method to improve their understanding. Student responses that addressed this theme were divided into two categories, which were further refined into two codes per category. These codes are outlined in Table 4.3.

Table 4.3

Qualitative Coding Theme 3: Plan to Improve Understanding

Category 5: Identifies a learning goal.		
Code	**Title**	**Example**
Q	States math specific learning goal that relates to the challenge of the task.	"I need to remember the format of explicit equations."
R	States generic learning goal that does not relate to the challenge of the task.	"I need to relearn this."
Category 6: States method to improve understanding.		
Code	**Title**	**Example**
S	States a specific plan that relates to the challenge of the task.	"Maybe if I could find a way to put explicit into words (like new = old + …) I could remember explicit equations better."
T	States generic plan that doesn't relate to the challenge of the task.	"I'll review my notes."

Student Engagement in AR Process and Impact

Each week I reviewed students' responses and analyzed the qualitative data, looking for trends in students' responses. Though the qualitative codes went through several iterations before reaching saturation, this preliminary analysis provided a framework for thinking about how different students self-assessed their performance, which influenced my focus while observing students as they engaged in this reflective practice. During the treatment cycle I began to notice differences in student behaviors while completing the AR. Initially I chose to pay attention to two particular students who seemed to have consistent, and different, SA responses. I later widened my observation to all students in an attempt to group students who maintained similar SA behaviors. These observations played a significant role in my data analysis and influenced changes in my methodology.

These observations revealed that students valued the opportunity to self-assess their performance differently. In order to triangulate my observations, I administered an exit ticket after the week eight formative assessment. This exit ticket asked students to self-select the level of value that they assign the experience of self-assessing using a 1 (no value) to 4 (high value) scale; I compared their self-assigned value level against the behaviors that I observed while students completed the AR. Eight students were selected, two from each self-assigned value level, whose indicated level of value aligned with my perceived level of value. These eight students comprise a subset of the participants and provided me an opportunity to investigate how the value that students attribute to engaging in SA interacts with the impact of the treatment. Though I continued to analyze all student data, I also analyzed the data specifically for this subset to identify trends across varying value levels.

Research Question 2

What impact do standards-based grading practices have on the quality of Secondary Math 1 Honors students' self-assessment of their understanding of the content standards?

In order to determine the impact that the treatment had on the quality of student SA, I identified characteristics of low- and high-quality SA. High quality SA included identification of specific reasons or strategies that support understanding, attribution of learning to internal factors, identification of math-specific competencies or challenges that relate to the learning objectives, and/or the creation of math-specific learning goals. Low quality SA provided generic reasons or strategies that support students' level of understanding, attributed learning to external factors, identified challenges that are not specific to the math learning goals, and/or created generic learning goals that did not relate to the math learning objectives. Twenty qualitative codes were identified (see Tables 1, 2 and 3), which were separated into ten codes that are evidence of a high-quality SA, and ten codes that are evidence of a low-quality SA.

Once the qualitative data was analyzed to determine differences in the quality of students' SA, changes in the frequency of high- and low-quality qualitative codes was analyzed. Figure 4.1 presents the combined frequencies of all high-quality SA codes (B, C, E, F, I, J, K, L, Q, S) during the 12-week research cycle. Figure 4.2 presents the combined frequencies of all low-quality SA codes (A, D. G, H, M, N, O, P, R, T) during the 12-week research cycle. The frequencies of high- and low-quality responses did not reveal significant changes from Week 1 to Week 12, as Figures 4.1 and 4.2 demonstrate.

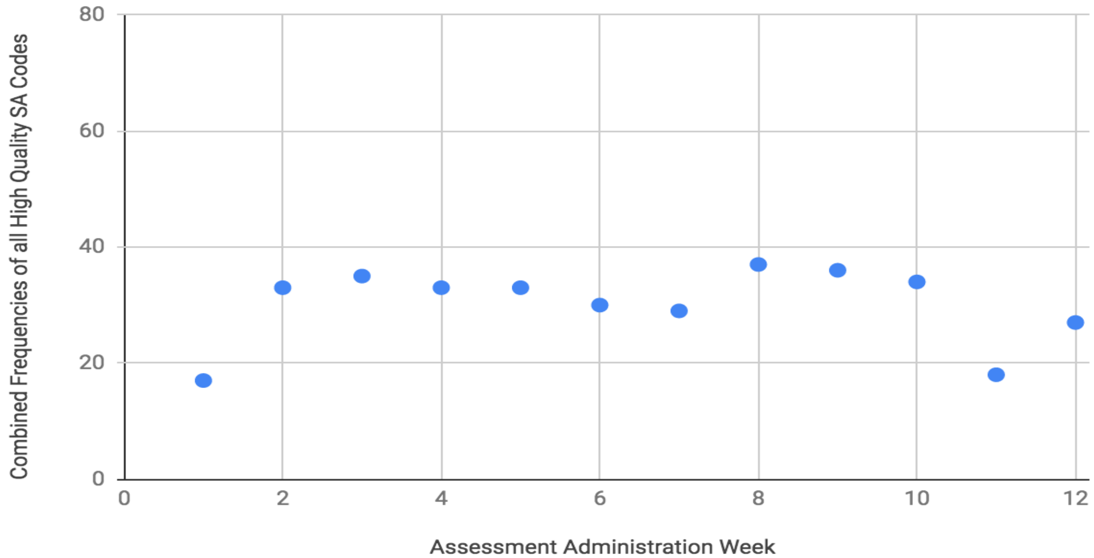


Figure 4.1. Combined frequencies of all high-quality SA codes.

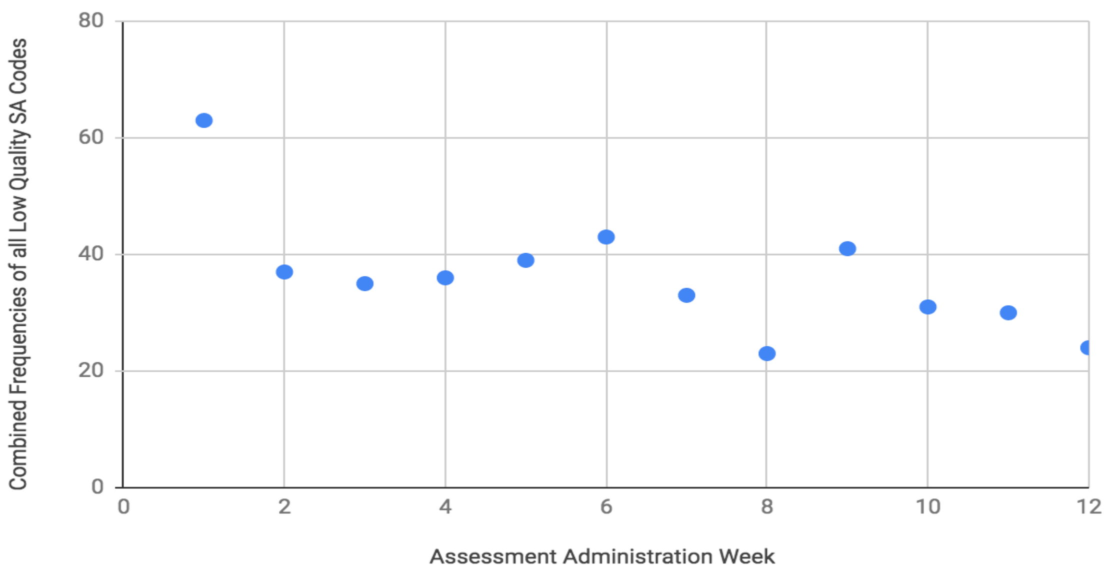


Figure 4.2. Combined frequencies of all low-quality SA codes.

The frequency of high- and low-quality SA codes reveals that there were no significant changes in the quality of students' SA between the baseline and treatment cycles. However, my experience observing students self-assess their performance, coupled with the identified divergent levels of value students have for the SA experience, caused me to question whether the degree to which students value the experience of self-assessing their performance influences the quality of the SA. This question is addressed later in this chapter and promotes this researcher's recommendations for future research.

Research Question 1

What impact do standards-based grading practices have on Secondary Math 1 Honors students' ability to accurately self-assess their performance on assessments items correlated to unique content standards?

The accuracy of students' self-assessment is defined as the difference between their predicted score and their achieved score. The achieved score was determined by marking students' work against a pre-established rubric. A negative measure of accuracy indicates that a student underestimated their performance, while a positive measure of accuracy indicates that a student over-estimated their performance. If an accuracy score is not reported, either that student did not take the particular assessment before the end of the research cycle, or they did not answer question two of the AR and do not have a predicted score. In cases when a student reassessed, their accuracy on multiple tiers was averaged and recorded. In cases when their predicted score was reported as a range of values, the midpoint of the range was assumed as their predicted score. For example, if a student predicted that they earned between 8 and 10 points, a predicted score of 9 was used to calculate their accuracy.

The accuracy of students' self-assessment was analyzed to determine if the treatment had any impact on whether or not students had a positive or negative measure of accuracy. Generally speaking, the majority of students overestimated their performance most frequently, and no trends in whether students over- or underestimated their performance were identified. Analysis of the data indicated that the treatment had no consistent impact on whether or not students under- or overestimated their performance.

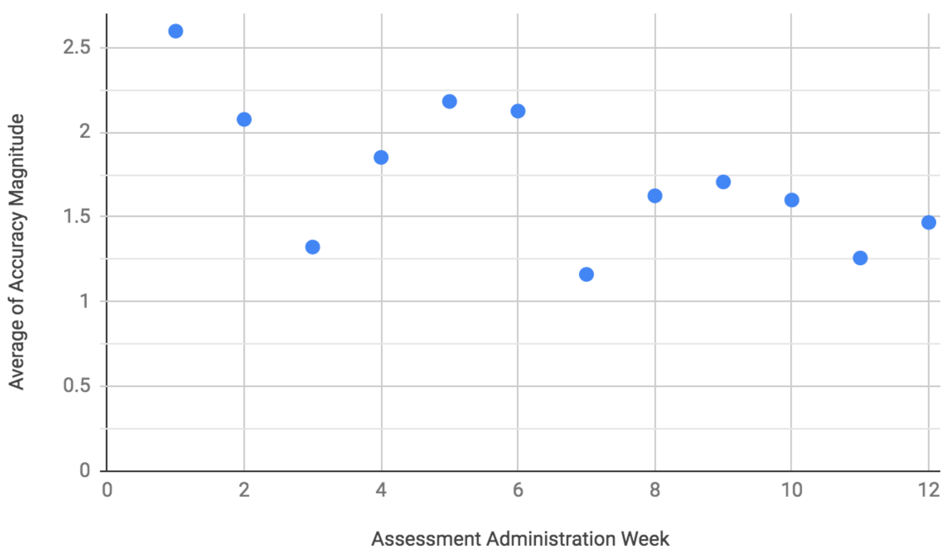


Figure 4.3. Average of students' SA accuracy.

Once I determined that the treatment did not have a noticeable impact on the direction of students' accuracy, I considered the magnitude of the accuracy to see if students became more accurate in their self-assessment over time by averaging all students' accuracy each week. For the remainder of this analysis, the accuracy of students'

SA is defined as the magnitude of the difference between the predicted and achieved score. The results are represented graphically in Figure 4.3.

Table 4.4

Average Accuracy for all Students' SA Scores

Assessment Administration Week	**Average Accuracy**
1	2.5968
2	2.0758
3	1.3226
4	1.8516
5	2.1818
6	2.1250
7	1.1613
8	1.6250
9	1.7069
10	1.6000
11	1.2581
12	1.4677

Table 4.5

SA Accuracy Summary Statistics

Research Cycle	**Mean**	**Standard Deviation**
Baseline	1.9617	0.5282
Treatment	1.6407	0.3664

During the baseline cycle, students self-assessed their performance within 1.96 points of their achieved score, on average. During the treatment cycle, students' average accuracy decreased to 1.64 points. This decrease indicates that students were able to self-assess their performance more accurately when the four SBG practices were implemented. This data is organized in Table 4.4 and summarized in Table 4.5.

A one-tailed t-test to compare the two means returned a p-value of 0.021 (t = 2.48), supporting the claim that the treatment had a significant impact on the accuracy of students' SA. Further, when the results from the treatment cycle are disaggregated between treatment weeks that used the original AR writing prompts (before the exit ticket), and treatment weeks that used the revised AR writing prompts (after the exit ticket), the accuracy of students' SA after the exit ticket administration also decreased significantly. During weeks five through eight, four weeks of the treatment cycle that used the original AR to prompt students' SA, returned an average accuracy of 1.77 points. During weeks nine through twelve however, the revised AR writing prompts were introduced, which yielded an average accuracy of 1.51 points. A one-tailed t-test to compare the difference in the mean accuracy between these two phases of the treatment cycle returned a p-value of p = 0.036 (t = 2.74), which suggests that making the purpose of self-assessing more explicit to students, and/or directly connecting the opportunity to reassess to the SA experience, has the potential to support the accuracy of students' SA.

Research Question 3

What impact do standards-based grading practices have on student achievement in Secondary Math 1 Honors?

Change in student achievement is defined as the difference between the number of questions that students answered correctly on the pre- and post-assessments. As the pre- and post- assessments contain different numbers of assessment items, the percent change between the pre- and post-assessment scores was analyzed. The pre- and post-assessment for the baseline cycle tested student mastery of four learning objectives, and the pre-and post-assessment for the treatment cycle tested student mastery of eight learning objectives; these tests contain twelve and twenty-four assessment items respectively.

Two students were excluded from this analysis. Student 24 was absent due to illness during the first week of the baseline cycle and did not complete the first pre-assessment until week three. Student 34 was absent when the second pre-assessment was completed in class and did not take this pre-assessment until week six. As these students had the opportunity to receive feedback on their weekly assessment items before they completed the pre-assessment, I chose to exclude them from this portion of the data analysis; their pre- and post-assessment scores were excluded for both the baseline and treatment cycles. Thus, the changes in achievement were analyzed for 31 of the 33 research participants.

Student achievement data for the baseline and treatment cycle is organized in Appendix G. Six students performed higher on the pre-assessment than the post-assessment during the baseline cycle, which yielded a negative change in achievement. During the treatment cycle two students experienced a negative percent change in achievement. The total number of questions answered correctly for each assessment is reported; as the baseline pre- and post-test contained 12 questions, whereas the treatment cycle pre- and post-test contained 24 questions, the percent change between administrations of each assessment is analyzed quantitatively.

During the baseline cycle the average percent change in student achievement was 7.53%, with a standard deviation of 14.49%. During the treatment cycle the mean percent change in student achievement increased to 15.49%, with a standard deviation of 11.21%. This change indicates that students experienced a 7.96% gain in achievement during the treatment cycle compared to the baseline cycle. Further, more students consistently experienced gains in achievement, as evidenced by the lower standard deviation between the two data sets. Using a two-tailed paired samples t-test, it was determined that changes in student achievement over the two research phases was statistically significant ($p = 0.009$, $t = 4.28$), indicating that the four SBG practices that were implemented during the treatment phase positively impacted student achievement.

Supplemental Analysis of Data

Qualitative data analysis yielded the defining characteristics of high- and low-quality SA, though the frequency of high- and low-quality codes did not significantly change between the baseline and treatment cycles. This quantitative analysis of the qualitative data revealed that the research treatment did not have significant impact on the quality of students' SA. However, observation of student behaviors while they self-assessed revealed that students thought differently about the process of self-assessing. Some students paused regularly while self-assessing before responding to the research questions, and others referred back to their responses on the assessment items, as evidenced by turning their page over repeatedly to consider their performance while completing the AR. On the other hand, some students complete all AR questions without pause, immediately after finishing the assessment times. The former behaviors were interpreted as evidence that these students valued the opportunity to self-assess their performance, whereas the latter behaviors were interpreted as evidence that those students did not value the opportunity. My perception of whether or not students valued the opportunity to self-assess their performance were compared to the value that students self-assigned on the exit ticket that followed the week assessment administration (Appendix F). Students who selected value level one indicated that the experience of self-assessing their performance has no value. Students who selected value level two indicated that they do not see the value of self-assessing, while value level three students indicated that their SA experience is somewhat valuable. Students who selected value level four indicated that the experience of self-assessing is valuable. Eight students were selected, two of each value level, whose self-assigned level of value aligned with my perceived level of value.

When I analyzed the frequency of high- and low-quality codes for the pairs of students separately across value levels, I found that students who valued the opportunity to self-assess their performance were much more likely to produce a higher quality SA. This trend continued for all four value levels. In order to further analyze this trend, I considered the combined relative frequency of low- and high-quality codes for each pair of students assigned to the four value levels. I noticed that students whose SA consisted of high-quality codes had fewer codes assigned to the SA; these students may have only engaged in reflective thinking that fell under one of the three themes. Students who produced a lower quality SA however, often included statements that aligned with multiple themes. Low-quality codes seemed to naturally appear more frequently, as students who produced a low-quality SA engaged in thinking about multiple aspects of their performance. To counter this effect, I chose to analyze the relative frequency of the codes assigned to students' SA. The graphic representation of the relative frequencies for high quality codes are provided for each pair of students in Figures 4.4, 4.5, 4.6, and 4.7.

All students included in this subset produced evidence of a high-quality SA at some point during the twelve-week research cycle. However, analysis of the relative frequency of high-quality codes indicated that students who value the opportunity to self-assess produced a higher quality self-assessment more consistently. This idea is reinforced by the average of the relative frequencies over the twelve-week period. Students with a value level of four, three, two, and one, had an average relative frequency of high-quality codes of 0.696, 0.466, 0.422, and 0.3677, respectively.

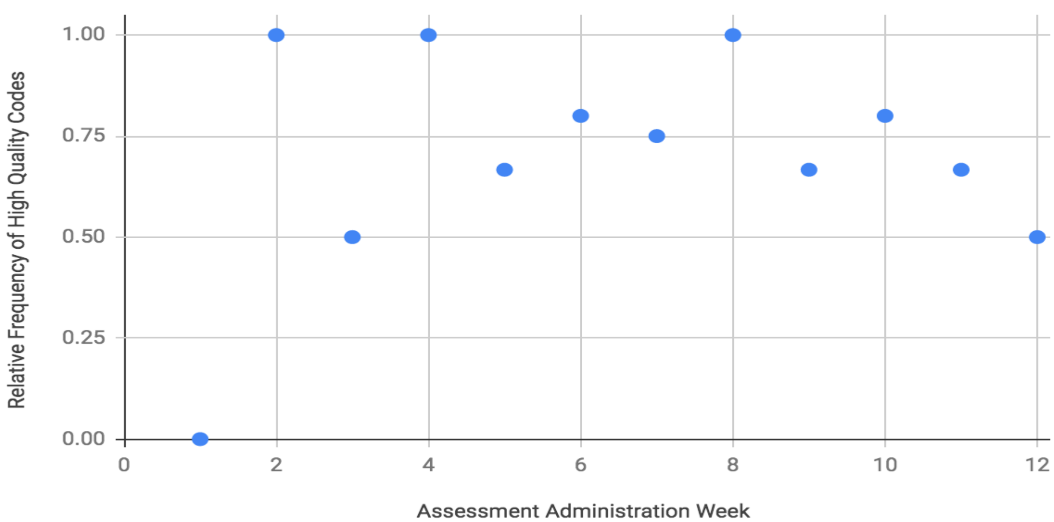


Figure 4.4. Relative frequency of high-quality codes for Value Level 4 students.

*Level 4 students believe that self-assessing their performance is valuable.

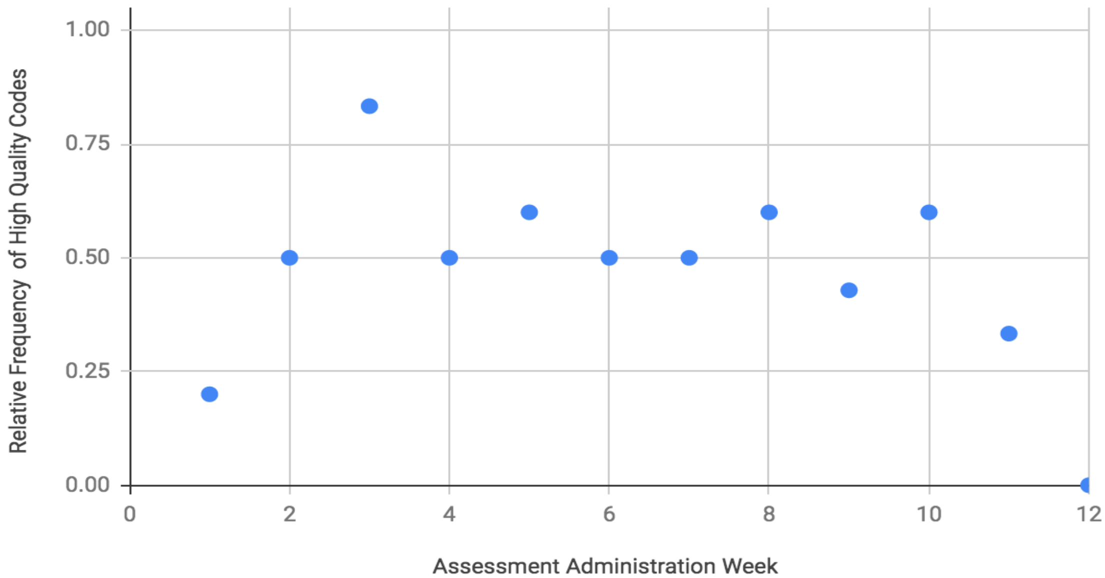


Figure 4.5. Relative frequency of high-quality codes for Value Level 3 students.

*Value level 3 students believe self-assessing their performance is somewhat valuable.

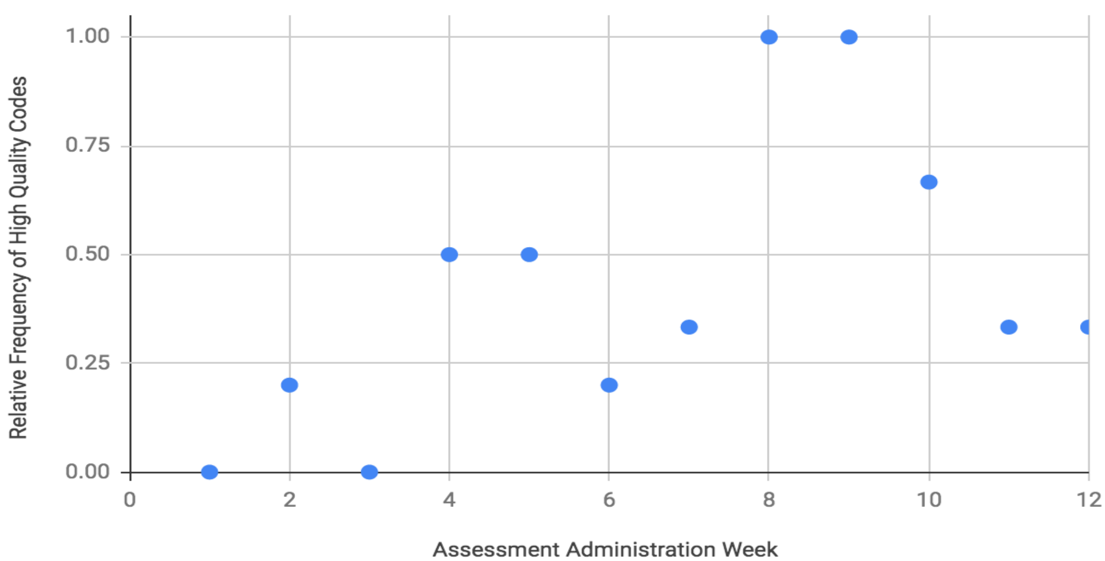


Figure 4.6. Relative frequency of high-quality codes for Value Level 2 students.

*Value level 2 students do not see the value in self-assessing their performance.

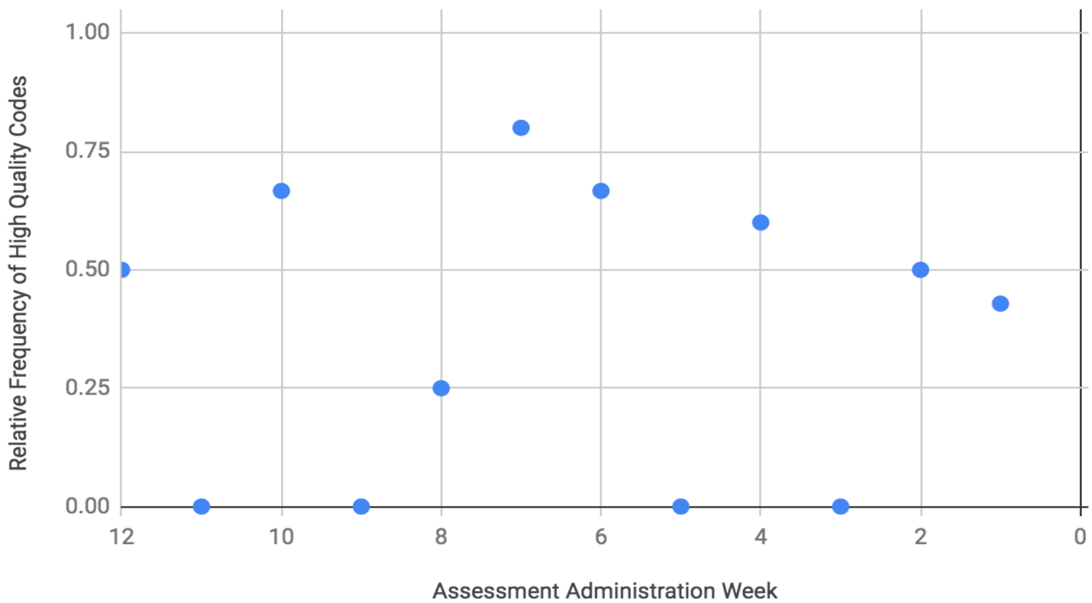


Figure 4.7. Relative frequency of high-quality codes for Value Level 1 students.

*Value level 1 students believe that self-assessing their performance has no value.

Summary

The purpose of this study was to determine whether four SBG practices support students' ability to critically and accurately self-assess their performance on formative assessments, and to determine the overall impact on student achievement. Quantitative data analysis provided evidence that implementation of the four SBG practices increased the accuracy of students' SA, and also had a positive impact on student achievement. However, no noticeable changes in the quality of students' SA were observed over the twelve-week research cycle.

Qualitative data analysis revealed trends in how students think about their performance. Student responses to the narrative reflection questions identified three themes in the way that students self-assess their performance, as students' reflective thinking on their performance assessed their level of understanding, indicated strategies that support their understanding, and/or created a plan to improve their understanding. Student responses were categorized by one, or multiple themes. Categories emerged within each theme that align with either high- or low-quality SA. Generally, a high-quality SA is characterized by assessment and strategies that imply a personal responsibility for learning and are linked to the math-specific learning goals of the assessment. Low quality SA rather, is characterized by assessment and strategies that are generic, or not linked to any math-specific learning goals. Low quality SA may also assign responsibility for their performance to external factors.

Analysis of students' SA and their behaviors while self-assessing revealed that students valued the opportunity to self-assess their performance very differently. I found that students who expressed a higher value level were more likely to produce a high-quality SA than their peers who expressed a low-quality SA. I expect that the quality of students' SA is tied to whether or not they value SA, and strategies to increase student engagement to the SA process should be explored.

Certain limitations should be considered when interpreting the results of this action research, as the small convenience sample limits the transferability of the data results. However, transferability was not prioritized in my methodology. I believe that supporting students' ability to critically and accurately self-assess their performance on formative assessments is a worthwhile endeavor, and I engaged in this research as a reflective practitioner seeking to improve my practice and support student learning. The implications that these findings have on my professional practice, as well as my recommendations for future research, will be further discussed in Chapter 5.

Chapter 5:

Summary, Conclusions, and Recommendations

Overview of Study

This mixed method, action research study aimed to identify specific classroom practices that support students' ability to critically and accurately self-assess their performance on formative assessment items, while also investigating the overall impact that these practices have on student achievement. Supporting students' ability to self-assess their performance and understanding positively impacts student motivation, self-efficacy, and achievement (Andrade & Valtcheva, 2009; Brown & Harris, 2014; Kelberlau-Berks, 2006; Leise, 2010; Noonan & Duncan, 2005; Panadero et al., 2015; Shepard, 2000; Stallings & Tascione, 1996; Stiggins, 2002; Stiggins & Chappuis, 2010), and students who self-assess are better prepared to construct their own knowledge (McMillan & Hearn, 2008) and become proactive self-regulators (Zimmerman, 2013). When working with other science and math teachers I found that we could not identify intentional classroom practices and strategies to support student self-assessment (SA), which motivated this research.

Review of existing research and related literature revealed recommendations to support student SA. Teachers should model SA processes (Brown & Harris, 2014), provide opportunities for students to self-assess their performance on formative assessments when there is an opportunity to revise their work or reassess (Brown et al., 2015; Bruce, 2001), and implement instructional practices that support a student-centered classroom (Butler & Lee, 2010) where students are involved in the assessment process (Stiggins & Chappuis, 2010). This research sought to translate these general support guidelines into specific classroom practices and policies by investigating the impact that four standards-based grading (SBG) practices have on students' ability to self-assess their performance.

Effective implementation of SBG practices includes communicated alignment between assessment items and standards (Colby, 1999; Iamarino, 2014), clear grading criteria that informs student understanding at varying levels of complexity, (Marzano & Heflebower, 2011), and student involvement in the assessment process by providing students choice and authority over their assessment schedule (Clayton & Shores, 2015; Colby, 1999; Marzano & Heflebower, 2011). These tenets of the SBG framework formed the foundation of the treatment imposed during this action research study, which included the following four SBG practices:

1. All assessment items are explicitly aligned to unique content descriptors.
2. All content descriptors are assessed at three tiers of difficulty (*proficient, approaching mastery, mastery*).
3. Students are able to self-select the assessment difficulty level for each content descriptor.
4. Students are able to improve the score they receive on all assessments, choosing to take assessments at a higher difficulty level to reflect their growth in understanding.

These four SBG practices were introduced during the final eight weeks of the research cycle in a ninth grade, Secondary Math 1 Honors class; their impact on students' ability to critically and accurately self-assess their performance on weekly formative assessment items was measured, as was the impact that these practices and behaviors have on student achievement. Quantitative and qualitative data analysis revealed that the treatment had a significant, positive impact on the accuracy of students' SA, as well as student achievement; the treatment did not have a significant impact on the quality of students' SA.

Research Design

By engaging in action research, I was able to participate in a cycle of inquiry that enriched my understanding of what students think about SA, and how they think when self-assessing. The cyclical, reflexive nature of action research allowed me to routinely reflect on what I observed in my classroom, and the agency to enact change that positively affected the overarching impact of the research cycle. I utilized the "plan-act-observe-reflect" (Merriam & Tisdell, 2016, p.54) cycle to make changes to the research plan when I observed differences in student behavior while completing the assessment reflection (AR). After reflecting on these differences, I modified the design plan by administering an exit ticket to triangulate my perception of whether students valued the experience of self-assessing their performance. This change allowed me to investigate reasons for the differences in value, and to consider the impact that these differences were having on the quality of students' SA. Ultimately this investigation strengthened my understanding of what students think about SA and challenged me to consider approaches to strengthen my research treatment and execution.

This twelve-week concurrent triangulation mixed methods action research case study took place in a ninth grade Secondary Math 1 Honors. The group of 33 research participants consisted of 20 females and 13 males. Though this is an honors level class, research participants have not been objectively identified as gifted and talented; placement in this honors level course is based on teacher recommendation and parent request.

Five types of instruments were used to collect qualitative and quantitative data; data was collected simultaneously, but each data addressed a unique research question. Data was collected during two phases of the research cycle: the baseline cycle, which occurred during the first four weeks, and the treatment cycle, which occurred during the last eight weeks of the research cycle.

Discussion and Reflections

Research Question 1

What impact do standards-based grading practices have on Secondary Math 1 Honors students' ability to *accurately* self-assess their performance on assessments items correlated to unique content standards?

Quantitative data analysis revealed that the treatment had a significant impact on the accuracy of students' SA. Between the baseline and treatment cycle, the mean accuracy of students' SA decreased from 1.9617 points to 1.6407 points, and the standard deviation decreased from 0.5282 points to 0.3664 points. The accuracy of students' SA is the difference between the predicted and the achieved score; this decrease in accuracy implies that the treatment had a positive impact on students' accuracy, as during the treatment cycle students' predicted score was an average 0.321 points closer to their achieved score than during the baseline cycle. A smaller standard deviation during the treatment cycle indicates that not only were students more accurate during the treatment cycle, but they were also more consistent.

Several factors could have contributed to students' improvement. Brown et al. (2015) found that students were less likely to self-assess their performance based on their ability when they did not have a clear understanding of the assessment criteria. It seems reasonable that if the difficulty of an assessment is far higher than a student's proficiency, that it would be difficult for that student to accurately self-assess. Offering assessments at varying levels of complexity could have supported the accuracy of students' SA by allowing them to engage with an assessment that was better aligned to their level of understanding, providing them a clearer understanding of the assessment criteria. Brown et al. (2015) also found that students who were stronger academically self-assessed more accurately; allowing students to complete assessments aligned to a lower depth of knowledge level (Webb, 2002) may have prompted a similar effect.

Alternatively, prompting students to self-select their assessment difficulty could have impacted the accuracy of their SA. Zimmerman's (2013) cyclical phases model of his self-regulated learning theory (SRL) outlines three phases: forethought, performance, and self-reflection. Processes during the forethought phase include task analysis, goal setting, and strategic planning; these processes inform student thinking during the performance and self-reflection phases (Zimmerman, 2013). I wonder if asking students to select their assessment complexity activated student thinking in the forethought phase. When students self-select the complexity of their assessment, they analyze their understanding of a given standard to establish a mastery goal for that standard. The thinking that occurs during the selection process could initiate thinking in the forethought phase, which subsequently impacts thinking during the final two phases of task completion.

Research Question 2

What impact do standards-based grading practices have on the *quality* of Secondary Math 1 Honors students' self-assessment of their understanding of the content standards?

Initial analysis of the frequency of the qualitative codes revealed that the treatment did not have a significant impact on the quality of students' SA, as there was not a significant change in the frequency of high- and low-quality SA codes during the baseline and treatment cycles. However, students who provided evidence of a low-quality SA consistently had more codes assigned to their SA than students who engaged in a high-quality SA. Students who provided a higher quality SA may have only reflected under one of the three qualitative coding themes, whereas students who provided a lower quality SA often reflected on multiple qualitative code themes.

The differences in the number of codes assigned to high- and low-quality SA prompted consideration of whether students' willingness to reflect on multiple aspects of their performance is more a measure of students' tenacity than their willingness to think thoroughly about their performance. Producing a high-quality SA requires students to think critically, and they may not be willing to think critically about as many aspects of their performance. A low-quality SA on the other hand, does not challenge students to think critically about their performance, and thus they may be willing to reflect on more aspects of their performance. For example, on the week eight AR, student fourteen responded to AR question three by stating, “I think I did the exponential function correctly, since 2^t is the correct function displayed in the graph.” This student continued his SA by writing down an ordered pair that he used to verify the accuracy of his work (x=0, y=1). This response was assigned codes E (provides a math-specific competency) and K (considers a specific or simpler case as a strategy to support understanding). On the other hand, on the week nine AR, student eight responded to question three by stating, “I definitely need to practice this more, but I think I mostly understand. I’m a little confused on question 3.” This response was assigned codes A (states generic for understanding), H (identifies specific portion of task that is challenging or manageable but doesn’t relate to the math-learning goals), R (identifies a generic learning goal), and T (states generic plan to improve understanding). As low-quality codes seemed to naturally occur more frequently, the relative frequency of high- and low-quality codes were analyzed. Though the treatment did not promote significant changes in the relative frequency of high- or low-quality codes throughout the research cycle, observation of student behaviors while self-assessing did suggest that the value that students attribute the experience of self-assessing their performance may influence the quality of their SA.

During the treatment cycle several steps were taken to determine the value that students attribute the opportunity to self-assess, and the results were triangulated to identify students who value the opportunity to self-assess across four value levels. My perception of whether students valued SA was compared to the level of value that students selected on an exit ticket (Appendix F). Students who selected value level one indicated that the experience of self-assessing their performance has no value. Students who selected value level two indicated they believed that SA could have some value, but that they do not see the value of self-assessing. Students who selected value level three indicated that the experience of self-assessing is somewhat valuable, and students who selected value level four indicated that the experience of self-assessing is valuable. Eight students total, two from each value level, whose self-selected value level aligned with my perceived value, were selected for closer analysis. Analysis of these students' SA across the four levels suggests that the value a student assigns the opportunity to self-assess their performance impacts the quality of their SA, as students who valued SA had a higher relative frequency of high-quality codes than students who did not value SA. Thus, in order to impact the quality of student SA, we should investigate strategies to increase student buy-in to the experience of self-assessing.

Andrade and Valtcheva (2009) advise that SA should be coupled with the opportunity to reassess, as students are more likely to value SA if they see it as an opportunity to improve. The assessment reflection (AR) was revised to highlight this opportunity to increase student buy-in. I also made attempts during whole group and small group discussions to make the purpose of self-assessing more explicit to students. However, it's possible that the overall design of the research process relied on practices and procedures that inherently limited the value that students attribute self-assessing their performance; recommended changes to the research design are provided later in this chapter.

Research Question 3

What impact do standards-based grading practices have on student *achievement* in Secondary Math 1 Honors?

Quantitative analysis of the change between students' pre- and post-assessment scores during the baseline and treatment cycle indicates that the treatment had a significant impact on student achievement. I believe that there are several factors that may have supported the growth in student achievement. Colby (1999) found that increasing student involvement in the assessment process shifts students' perception of grades as terminal measurements of learning, towards learning opportunities informed by feedback. The effect that Colby (1999) observed could have been imitated in this study by allowing students the opportunity to reassess. This practice increased the amount of time that some students spent grappling with content they found challenging, as additional time was needed to prepare for and engage in the reassessment. This additional time provided students with additional learning opportunities, which likely strengthened their learning and positively impacted their performance on the final post-assessment.

Prompting students to reflect on their level of understanding may also explain the significant growth in achievement. The results of this study are consistent with other researchers' findings that there is a positive association between supporting SA and student achievement (Andrade & Valtcheva, 2009; Brown & Harris, 2014; Kelberlau-Berks, 2006; Leise, 2010; Noonan & Duncan, 2005; Panadero et al., 2015; Shepard, 2000; Stiggins, 2002; Stiggins & Chappuis, 2010). The positive impact that the four SBG practices had on the ability of students to accurately self-assess their performance has already been established; thus, it's possible that the positive impact on achievement is a result of the increase in students' SA accuracy. Alternatively, asking students to select the difficulty level of their assessment prompted students to assess their level of understanding, create performance goals, and strategize; all of which are processes in the Zimmerman's (2013) forethought phase. As these processes inform student thinking during the performance and reflection phases, initiating this thinking may have contributed to the significant change in student achievement as well.

It is worth noting that this research occurred during the weeks prior to the end of year standardized assessment for this course. Though this standardized assessment does not impact students' academic mark in this class, many of the students in this class are highly motivated to achieve academic success. It is possible that some of the growth in achievement can be attributed to students' desire to be successful on their end of course assessment. Though I maintained consistent classroom practices throughout the baseline and treatment cycles, this study was unable to control for students' willingness to engage in review outside of the classroom.

Limitations of Study

The impact of whether or not students value the opportunity to self-assess their performance was not thoroughly considered before beginning this action research study which limited the impact of the research results. In light of the research results, I believe that whether or not students value the opportunity to self-assess plays a significant role in the quality of their SA. Further consideration should be given to identifying strategies that increase student value of SA, and the overall impact that this has on the quality of their SA. Once student buy-in to the process of self-assessing has been established, the impact that the four SBG practices have on the quality of their SA should be considered. Though this study included students in an honors level mathematics course, there were no observed patterns in the change to students' SA across ability level. Further, students in this class are not objectively identified as gifted and talented. Thus, the impact that the four SBG practices had on students' SA may translate to other ability level groupings. As the benefits of supporting students' SA and the treatment itself are not rooted in mathematics instruction, the results of this study may also translate to other content areas as well.

Practice Recommendations

Overall the treatment had a positive impact on students' SA; these four SBG practices are recommended to those interested in supporting the accuracy of students' SA and student achievement. The AR was a successful tool to prompt students to self-assess; the revised AR is recommended if using the AR in conjunction with the four SBG practices. Andrade and Valtcheva (2009) found that students will not value SA unless it leads to an opportunity to improve their performance; the revised AR explicitly links SA to the opportunity to improve.

Further consideration should be given to instructional strategies and practices that increase the value that students' attribute the opportunity to self-assess their performance, as this has the potential to positively impact the quality of their SA. Supporting students' ability to think critically about their performance and engage in high-quality SA is a worthwhile pedagogical goal, as improving the quality of SA supports the development of Zimmerman's (2013) proactive self-regulatory learning practices. Further, critical thinking about their performance may improve the quality of processes during the forethought phase of Zimmerman's (2013) learning cycle, which enhances learning. It is worth noting however, that whether students value the opportunity to self-assess was not associated with overall high achievement. Not all participants who valued SA were high achieving, and not all high achieving participants valued self-assessing their performance. Thus, supporting high quality SA is not a requirement to support the learning of all students, but may be a potential avenue to support the learning for some students.

I also recommend investing instructional time to model SA behaviors, and to discuss SA in whole group and small group settings. Students are more likely to take ownership of their learning if they are allowed the opportunity to communicate about their performance with others (Andrade & Valtcheva, 2009; Stiggins, 2002; Stiggins & Chappuis, 2010). Students may view SA as a worthwhile practice if attempts are made to make SA behaviors more intentional and more public. Thus, providing structured opportunities for students to discuss SA, share their process for self-assessing their performance, or to work in groups to assess collaborative work should be considered. I wonder if this will not only prepare students to engage in richer thinking while they self-assess but may also increase the value that they attribute to self-assessing.

Transferability

Action research aims to inform local, rather than global, practice. Unlike traditional research methods, transferability is not the goal of action research (Herr & Anderson, 2015; Merriam & Tisdell, 2016). For this reason, transferability was not prioritized in my research methodology; my research participants were chosen for their accessibility using a convenience sampling method. However, I have taken several steps to support transferability outside of my local setting. As I engaged in this research acting as an insider collaborating with other insiders (Herr & Anderson, 2015), it was important that I was mindful of my beliefs, culture, and relationship to the research participants in order to reduce bias and subjectivity from this research. I reflected thoroughly on my positionality, or my status within my network of research participants, before, during, and after the research cycle. Though a researcher is unable to eliminate all bias and subjectivity from their research, I attempted to confront my positionality in order to increase the validity, or trustworthiness, of the research results.

Though external validity is not a priority of action research, establishing internal validity is a universal goal (Merriam & Tisdell, 2016). I have attempted to be transparent with my research design, the decisions I made that altered the implementation plan, as well as the quantitative and qualitative data analysis. Through this transparency I hope to increase the transferability of the research results, but it is the reader's responsibility to determine whether the research results apply to their local setting. Consideration should be given to the methodology, changes in the implementation plan, and characteristics of the research participants and the researcher's local context before adapting treatment.

Recommendations for Future Research

By engaging in this action research study, I deepened my understanding of what students think about SA, and how they think when self-assessing. Throughout this cycle of inquiry I reflected on student work and behaviors to refine my research methods. Through this process I determined several changes that I would make to the research plan before conducting this research again. I also developed new questions that may motivate future research.

The first change would occur before beginning the twelve-week research cycle; I would incorporate more opportunities for students to self-assess their performance throughout the entire academic year. Brown and Harris (2014) recommend that students transition from simple, concrete SA techniques before complex techniques that include holistic judgements of their performance; numerous other researchers recommend that students receive explicit instruction on how to self-assess (Andrade & Valtcheva, 2009; Brown et al., 2015; Brown & Harris, 2014; Bruce, 2001; Collett, 2014; Hoernke, 2014; Mazloomi & Khabiri, 2018; Stiggins, 2002; Zimmerman, 1990). Prior to investing specific strategies that improve student SA, I think that it would be beneficial to develop basic SA practices by implementing strategies that shift from concrete to abstract SA techniques.

I would also make changes regarding the way that students receive feedback on their performance and SA. Researchers recommend that teachers support SA by providing opportunities for students to participate in the assessment process by inviting students to participate in record keeping by monitoring their progress, and to communicate with others about their performance (Andrade & Valtcheva, 2009; Stiggins, 2002; Stiggins & Chappuis, 2010). Changes must be made to the design plan to incorporate these strategies. In order to keep test items secure so that students could later reassess at a higher tier, students in this study were not permitted to discuss any aspects of their performance, strategy, or SA, with their peers. Instead, students received individual feedback from this teacher-researcher. It's possible that more students may have valued the opportunity to self-assess their performance if they were able to communicate with their peers about their assessment and SA and were given more ownership over their own record keeping. These practices could have also strengthened our student-centered classroom climate, a shift that is recommended to support student SA (Butler & Lee, 2010). Providing opportunities for students to engage in the assessment process in this way also may have challenged students to take more ownership over the assessment process, which has the potential to positively impact the value they attribute to SA.

Lastly, this experience revealed that the quality of a student's SA is related to whether or not the student values the opportunity to self-assess. Thus, research that investigates strategies to increase the level of value that students attribute to SA is needed. Research suggests that SA opportunities should be linked to an opportunity to revise or reassess (Andrade & Valtchava, 2009). Consideration may also be given to Bradbury-Bailey's (2011) orientations for interpreting grades: the mastery goal orientation (MGO) and performance goal orientation (PGO). Bradbury-Bailey found that students who maintained an MGO had a greater value for grades; strategies that support the shift from PGO to MGO should be considered as potential strategies to increase students' value for grades, as these strategies may also increase students' value of their SA of their grades.

Summary

In 2016 the National Board for Professional Teaching Standards (NBPTS) asked candidates how they support student SA to achieve learning objectives. This writing prompt highlighted an instructional need that was lacking in my personal practice.

Investigation of classroom strategies that support student SA lacked specificity in action, which motivated this mixed methods action research study. The impact that four SBG practices have on the accuracy and quality of student SA, as well as the overall impact on student learning, was determined. The four SBG practices had a significant, positive, impact on the accuracy of students' SA, as well as on student achievement. The SBG practices did not impact the quality of students' SA. However, this investigation and analysis of the quality of students' SA illuminated differences in the way students think while self-assessing, and what they believe about the experience of self-assessing. Ultimately this experience provided evidence of specific classroom practices and policies that support student SA, and also challenged me to consider alternate factors that influence the quality of students' SA.

This final chapter of this study provides an overview of the research design, the research conclusions, my reflections on how these conclusions relate to previous research and the theoretical framework that drives this research, as well as implications for my instructional practice and future research. Though this study began with the goal of supporting the accuracy and quality of student SA, it evolved to reveal another ancillary goal: to increase the value that students assign the process of self-assessing. I now believe that investigating strategies that increase student value of the SA process is a worthy goal that must be tackled in order to better support the quality of student SA. Investigation of the impact that the opportunity to revise or reassess has on how students value SA, as well as strategies that support shifts from Bradbury-Bailey's (2011) PGO to MGO, should be considered.

Implementation of the research treatment is recommended to support student SA, but other conditions should be considered to increase student engagement while self-assessing, including increased modeling of self-assessing behaviors, transition from simple, concrete SA techniques to more complex techniques, and increased student involvement in the assessment process by encouraging students to maintain a record of their assessments and supporting collaborative communication about assessment processes and results amongst students.

References

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice, 48,* 12–19.

Andrzejewski, C. E., Davis, H. A., Bruening, P. S., & Poirier, R. R. (2016). Can a self-regulated strategy intervention close the achievement gap? Exploring a classroom-based intervention in 9th grade earth science. *Learning and Individual Differences, 49,* 85–99.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86*(1), 13–22.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Educational Assessment: Principles, Policy and Practice, 5*(1), 7–74.

Blazar, D., & Pollard, C. (2017) Does test preparation mean low-quality instruction? *Educational Researcher, 46*(8), 420–433.

Bradbury-Bailey, M. (2011). *A preliminary investigation into the effects of standards-based grading on the academic performance of African-American students* (Doctoral dissertation). Retrieved from PQDT Open. (3511593)

Brookhart, S., Andolina, M., Zuza, M., & Furman, R. (2004). Minute math: An action research study of student self-assessment. *Educational Studies in Mathematics*, 57(2), 213–227.

Brown, G., Chen, F., & Andrade, H. (2015). Accuracy in student self-assessment: Directions and cautions for research. *Assessment in Education Principles Policy and Practice, 22*(4), 1–26. doi:10.1080/0969594X.2014.996523

Brown, G. T. L., & Harris, L. R. (2014). The future of self-assessment in classroom practice: Reframing self-assessment as a core competency. *Frontline Learning Research, 3*, 22–30. doi:10.14786/flr.v2i1.24

Brubaker, D. L. (2004). *Revitalizing curriculum leadership: Inspiring and empowering your school community*. Thousand Oaks, CA: Corwin Press.

Bruce, L. B. (2001). Student self-assessment: Making standards come live. *Classroom Leadership, 5*(1), 1–6. Retrieved from http://www.ascd.org/publications/classroom_leadership/sept2001/Student_Self-Assessment.aspx

Bruner, J. (1962). *On knowing: Essays for the left hand.* Cambridge, MA: The Belknap Press of Harvard University Press.

Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing, 27*(1), 5–31.

Bush, G. H. W. (1993). *America 2000: An education strategy sourcebook.* Washington, DC: Diane Publishing.

Clayton, M., & Shores, M. (2015). Standards-based grading in secondary mathematics: Review of literature. *Global Education Journal, 2015*(3), 157–174.

Colby, S. A. (1999). Grading in a standards-based system. *Educational Leadership, 56*(6), 52–55.

Collett, P. (2014). Researching self-regulated learning and foreign language learning. *Studies in Self-Access Learning Journal, 5*(4), 430–442.

Derry, M. (2017, January 9). Research writing: What's your positionality? *The Weingarten Blog.* Retrieved from https://weingartenlrc.wordpress.com/2017/01/09/research-writing-whats-your-positionality/

Desjarlais, M., & Smith, P. (2011). A comparative analysis of reflection and self-assessment. *International Journal of Process Education, 3*(1). Retrieved from www.processeducation.org/ijpe/2011/refl ectionh.pdf

Dewey, J. (1938). *Experience and education.* New York, NY: Simon & Schuster.

Fernando, S., & Marikar, F. (2017). Constructivist teaching/learning theory and participatory teaching methods. *Journal of Curriculum and Teaching, 6*(1), 110–122. https://doi.org/10.5430/jct.v6n1p110

Goodpasture, J., Reinhardt, L., & Thomas, M. (2007). A study of the self-regulated learning inventory on a HBCU student population in allied health. *The Internet Journal of Allied Health Sciences and Practice, 5*(4). Retrieved from http://ijahsp.nova.edu

Herr, K., & Anderson, F. (2015). *The action research dissertation: A guide for students and faculty*. Thousand Oaks, CA: SAGE.

Hoernke, M. A. (2014). *Perceived effects of standards-based grading on student assessments* (Doctoral dissertation). Retrieved from ProQuest Dissertations Publishing. (3623550)

Hochbein, C., & Pollio, M. (2016). Making grades meaningful. *Phi Delta Kappan, 98*(3), 49–54. https://doi.org/10.1177/0031721716677262

Iamarino, D. L. (2014). The benefits of standards-based grading: A critical evaluation of modern grading practices. *Current Issues in Education, 17*(2), 1–12.

Kelberlau-Berks, D. R. (2006, July). *The effects of self-assessment on student learning* (Master's dissertation). Retrieved from http://digitalcommons.unl.edu/mathmidactionresearch/8

Leise, C. (2010). Improving quality of reflecting on performance. *International Journal of Process Education, 2*(1), 65–74.

Machi, L., & McEvoy, B. (2016). *The literature review: Six steps to success* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

Marzano, R. J. (2009). *Formative assessment & standards-based grading*. Bloomington, IN: Marzano Research.

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

Mazloomi, S., & Khabiri, M. (2018). The impact of self-assessment on language learners' writing skills. *Innovations in Education and Teaching International, 55*(1), 91–100.

McMillan, J. H., & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *Educational Horizons, 87*(1), 40–49.

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation.* San Francisco, CA: Jossey-Bass.

Mertler, C. A. (2017). *Action research: Improving schools and empowering educators* (5th ed.). Thousand Oaks, CA: SAGE Publications.

Miller, J. J. (2013). A better grading system: Standards-based, student-centered assessment. *The English Journal, 103*(1), 111–118. Retrieved from http://www.jstor.org/stable/24484069

National Board for Professional Teaching Standards (NBPTS). (2016). *Early adolescence-adolescence and young adulthood/mathematics component 4: Effective and reflective practitioner portfolio instructions and scoring rubric*. Retrieved from http://boardcertifiedteachers.org/certificate-area

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics (CCSM).* Retrieved from http://www.corestandards.org/Math/

Noonan, B., & Duncan, R. (November 2005). Peer and self-assessment in high schools. *Practical Assessment Research & Evaluation, 10*(17). Retrieved from http://pareonline.net/getvn.asp?v=10&n=17

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

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*(422), 1–28. http://doi.org/10.3389/fpsyg.2017.00422

Panadero, E., Brown, G., & Strijbos, J.W. (2015). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review, 28,* 803–830. doi:10.1007/s10648-015-9350-2

Piaget, J. (1964). Part 1: Cognitive development in children: Piaget development and learning. *Journal of Research in Science Education, 2*, 176–186.

Polikoff, M. S., & Porter, A. C. (2014). Instructional alignment as a measure of teaching quality. *Educational Evaluation and Policy Analysis, 36*(4), 399–416.

Rosales, R. (2013). *The effects of standards-based grading on student performance in algebra 2* (Doctoral dissertation). Retrieved from TopSCHOLAR. (http://digitalcommons.wku.edu/diss/53)

Ryan, M., Taylor, M., Barone, A., Pesca, L.D., Durgana, S., Ostrowski, K., … Pikaard, K., (2017). Teacher as researcher, teacher as scholar, and teacher as leader. *The New Educator, 13*(2). 102–116. https://doi.org/10.1080/1547688X.2016.1144120

Sharikzadeh, N. (2003). Assessment in constructive learning environments. *Proceedings of the International Conference on Emerging Technologies,* Minneapolis, MN. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.57.9792

Shepard, L. A. (2000). The role of classroom assessment in teaching and learning. Washington, DC: Office of Educational Research and Improvement.

Stallings, V., & Tascione, C. (1996). Student self-assessment and self-evaluation. *Mathematics Teacher, 89*(7), 548–554.

Steinkruger, C. (2007). Do students progress if they self-assess? A study in small-group work (Master’s dissertation). Retrieved from http://digitalcommons.unl.edu/mathmidactionresearch/27

Stiggins, R. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan.* Retrieved from http://www.pdkintl.org/kappan/k0206sti.htm

Stiggins, R., & Chappuis, J. (2010). Using student-involved classroom assessment to close achievement gaps. *Theory Into Practice, 44*(1), 11–18.

Sung, K. (2006). Literature review on self-regulated learning. *Singapore Nursing Journal. 33*(2), 38–45.

U.S. Department of Education. (2002). *No child left behind: A desktop reference.* Retrieved from https://www2.ed.gov/admins/lead/account/nclbreference/reference.pdf

U.S. Department of Education. (2010). *Race to the top program guidance and frequently asked questions.* Retrieved from https://www2.ed.gov/programs/racetothetop/faq.pdf

Webb, N. (2002, March 28). *Depth-of-knowledge levels for four content areas.* Retrieved from http://facstaff.wcer.wisc.edu/normw/All%20content%20areas%20%20DOK%20levels%2032802.pdf

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3–17. doi:10.1207/s15326985ep2501_2

Zimmerman B. J. (2013). From cognitive modeling to self-regulation: a social cognitive career path. *Educational Psychologist, 48*, 135–147. doi:10.1080/00461520.2013.794676

Appendix A:

Selected Common Core Standards

Baseline Cycle:

1. F.IF.7: Graph functions expressed symbolically and show key features of the graph, by hand in simple cases and using technology for more complicated cases.
2. F.IF.2: Use function notation, evaluate functions for inputs in their domains, and interpret statements that use function notation in terms of a context.
3. A.REI.12: Graph the solutions to a linear inequality in two variables as a halfplane (excluding the boundary in the case of a strict inequality), and graph the solution set to a system of linear inequalities in two variables as the intersection of the corresponding half-planes.
4. F.BF.1: Write a function that describes a relationship between two quantities.

Treatment Cycle:

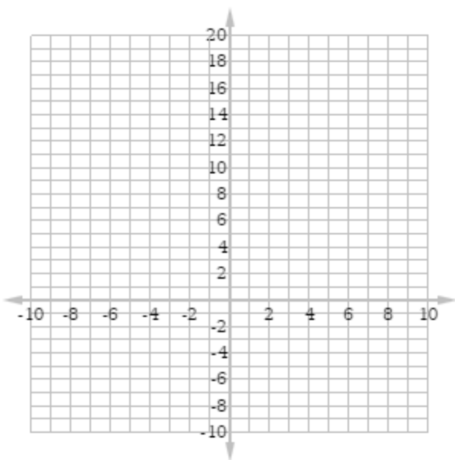
1. A.CED.1: Create equations and inequalities in one variable and use them to solve problems. Include equations arising from linear and quadratic functions, and simple rational and exponential functions.
2. A.CED.2: Create equations in two or more variables to represent relationships between quantities; graph equations on coordinate axes with labels and scales.
3. A.REI.3: Solve linear equations and inequalities in one variable, and give examples showing how extraneous solutions may arise.
4. A.REI.6: Solve systems of linear equations exactly and approximately (e.g., with graphs), focusing on pairs of linear equations in two variables.
5. F.BF.2: Write arithmetic and geometric sequences both recursively and with an explicit formula, use them to model situations, and translate between the two forms.
6. F.IF.4: For a function that models a relationship between two quantities, interpret key features of graphs and tables in terms of the quantities, and sketch graphs showing key features given a verbal description of the relationship. Key features include: intercepts; intervals where the function is increasing, decreasing, positive, or negative; relative maximums and minimums; symmetries; end behavior; and periodicity.
7. F.IF.6: Calculate and interpret the average rate of change of a function (presented symbolically or as a table) over a specified interval. Estimate the rate of change from a graph.
8. F.LE.1: Distinguish between situations that can be modeled it linear function and with exponential functions.

Appendix B:

Traditional Formative Assessments

Formative Assessment 1

1. Graph the function $f(x) = 3(2)^x$ on the axes below, and identify the requested key features. (6 points)



Key Features:

a) Identify any x- or y-intercepts:

b) State the range of $f(x)$:

c) State where $f(x)$ is positive using interval or inequality notation:

d) As x approaches $-\infty$, what does y approach?

2. State at least two differences and two similarities between the graphic representations of the functions given below. (4 points)

- $g(x) = 2(x - 1) + 3$
- $f(x) = 3 * 2^{x-1}$

Differences	**Similarities**

Formative Assessment 2

A driver travels along a straight road at a constant speed. The distance the driver travels, measured in miles, is modeled by the function, $f(t)$. The driver began driving at $t = 0$ minutes, and stops driving at $t = 10$ minutes. Indicate appropriate units when responding to each question.

$$f(t) = 2(t - 1) + 2$$

1. Evaluate $f(8)$. (2 points)

2. For what value of t does $f(t) = 12$? (3 points)

3. It is known that $f(1) = 2$. What does this information mean in the context of this situation? (2 points)

4. It is known that $f(10) - f(0) = 20$. What does this information mean in the context of this situation? (2 points)

Formative Assessment 3

Image 1

Image 2

Image 3

1. Analyze the pattern represented above to determine the number of Cheerios that would appear in the fourth image. (1 point)

2. Write an explicit equation, $f(n)$, that models the number of Cheerios that would appear in any image, n. (4 points)

3. Write a recursive equation, $f(n)$, that models the number of Cheerios that would appear in any image, $n \geq 2$. (4 points)

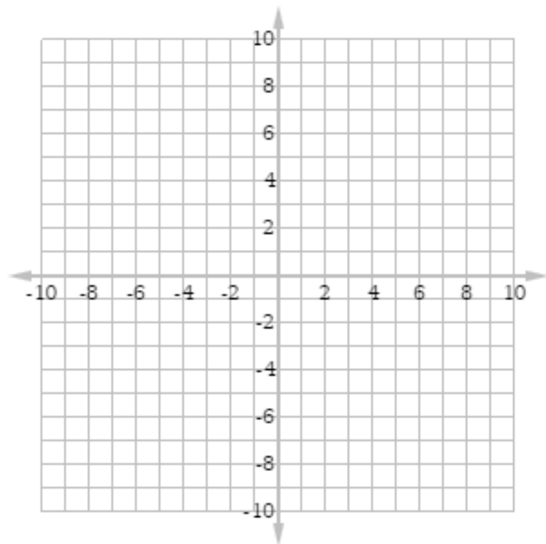
Image from VisualPatterns.org

Formative Assessment 4

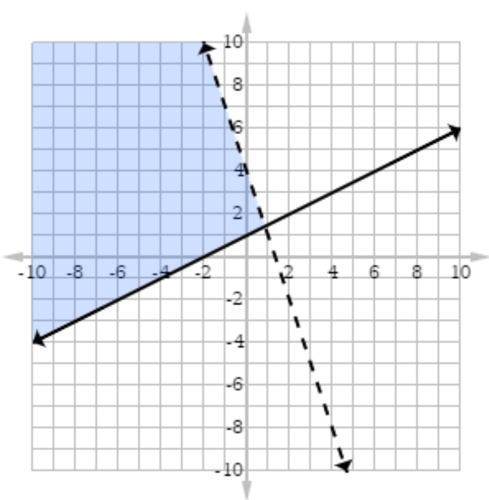
1. Graphically represent the solution set of the system of linear inequalities given below. (5 points)

$$y < 2x - 1$$

$$2 - 3y \leq x$$



2. Write the system of inequalities whose solution set is represented below. Then circle all of the given coordinates that are members of the solution set. (6 points)



1. Write the system of inequalities. (4 points)

2. Circle all ordered pairs that are members of the solution set of this system of linear inequalities. (2 points)

(-4, 4)

(-4, -2)

(4, 2)

(0, 4)

(-4, -1)

Appendix C:

Tiered Formative Assessments

Formative Assessment 5 Tier 1

A.CED.1: Create equations and inequalities in one variable and use them to solve problems. Include equations arising from linear and exponential functions.

1. Solve the equation given below. (3 points)

$$2(x-3)=4x+4$$

2. When half of a number is increased by 3, the result is 11.
   a. Write an equation that models this situation. (2 points)

   b. Write down the value of this unknown number. (1point)

3. Solve the inequality given below. (2 points)

$$2(3-x) < 10$$

Formative Assessment 5 Tier 2

A.CED.1: Create equations and inequalities in one variable and use them to solve problems. Include equations arising from linear and exponential functions.

1. The number of boxes in each image given below forms a sequence. Your challenge is to determine the number of the image that has more than 60 boxes.

Image 1 Image 2 Image 3 Image 4

   a. Write an inequality that models this situation. (3 points)

   b. Solve your inequality from part (a) to describe the image number of all possible images that have more than 60 boxes. (2 points)

2. The equation given below is true if $x = 4$. What is the value of k ? (2 points)

$$k * 3^x = 243$$

3. Karen finds a \$5 bill and a variety of coins in her wallet. She counts a total of 4 dimes, 9 nickels, 3 pennies, and some quarters. She determines that she has a total of \$8.13.

   a. Write an equation that models this situation. (2 points).

   b. Solve your equation from part (a) to determine the number of quarters Karen has in her wallet. (1 point)

Formative Assessment 5 Tier 3

A.CED.1: Create equations and inequalities in one variable and use them to solve problems. Include equations arising from linear and quadratic functions, and simple rational and exponential functions.

Equation #1	Equation #2
$2x+4=3x-1$	$kx-4=12$

1. The solution of Equation #2 is smaller than the solution of Equation #1. Determine one possible value of k. (3 points)

2. Create an equation or inequality that meets the given requirement by filling in the blanks with any integer.

   a. Create an equation that has infinitely many solutions. (3 points)

   _____$x + 3 =$ ___(___$x - 1) +$ _____

   b. Create an inequality whose solution is $x > 2$. (3 points)

   _____$\cdot 3^x -$ ______>5

Formative Assessment 6 Tier 1

A.REI.3: Solve equations and inequalities in one variable.

Solve each equation or inequality given. Please circle your final answer.

$2x + 3 = 15$ (2 points)	$2x - 1 = 3(x + 5)$ (3 points)
$3x - 4 < 5x - 6$ (3 points)	$5 < 3 - x \leq 9$ (3 points)

Formative Assessment 6 Tier 2

A.REI.3: Solve equations and inequalities in one variable.

Fill in the blank with any digits, 1–9, to create an equation or inequality that is true, and that also satisfies the given condition. Each question is worth 2 points.

1. The solution x = 2 satisfies the equation given below. What digit must be missing?

$$2x + \_\_\_\_ = 10$$

2. The equation given below has no solution. What digit must be missing?

$$3x - 5 = \_\_\_x + 1$$

3. The solution $x < 3$ satisfies the inequality given below. What digit must be missing?

$$4x - 1 < \_\_\_x + 5$$

4. The inequality given below has the solution [___, ∞). What digit is missing from this solution?

$$4 \geq 10 - 2x$$

5. Explain how you thought about question #2. How did you come up with your answer?

Formative Assessment 6 Tier 3

A.REI.3: Solve equations and inequalities in one variable.

Fill in the blanks with any digits, 1 – 9, to create an equation or inequality that satisfies the given condition. You can only use each digit once per question. Questions 1-3 are worth 3 points each; question 4 is worth 2 points.

1. The solution x = 3 satisfies the equation given below. Fill in the blanks with digits that would make this statement true.

$$\_\_\_(x + \_\_\_) - \_\_\_\_ = \_\_\_\_$$

6. The equation given below has infinitely many solutions. Fill in the blanks with digits that would make this statement true.

$$\_\_\_\_(\_\_\_\_x + \_\_\_\_) = \_\_\_\_x + \_\_\_\_$$

7. Fill in the blanks so that the inequality given below is only true if x is a negative number.

$$2x + \_\_\_ < \_\_\_\_$$

8. Explain your thought process for your response to question #3.

Formative Assessment 7 Tier 1

F.LE.1: Distinguish between situations that can be modeled with linear function and with exponential functions.

Identify whether each representation can be modeled by a linear equation, an exponential equation, or neither. Give a reason for your answer. (2 points each)

1. Type of function: _____________

X	Y
1	32
2	48
3	72
4	108

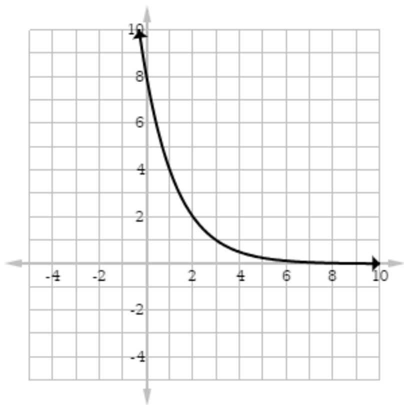
Reason:

2. Type of function: ________________

X	Y
3	7
4	4
6	-2
7	-5

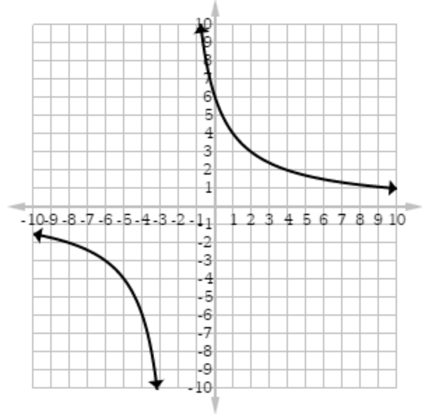
Reason:

2. Type of function: _____________



3. Type of function:

____



Reason:

Reason:

4. Each consecutive y-coordinate increases by 3 units.

Type of function:

Reason:

Formative Assessment 7 Tier 2

F.LE.1: Distinguish between situations that can be modeled with linear function and with exponential functions.

Identify whether each representation can be modeled by a linear equation, an exponential equation, or neither. Give a reason for your answer and predict the unknown term in the sequence when indicated. (#1 and 2 are worth 3 points, #3 and 4 are worth 2 points each)

1. 125, 25, ____, 1, 0.2, ...

Type of function: _________________

Reason:

What is the value of the third term?

2. 48, 36, 24, ____, 0, …

Type of function: _________________

Reason:

What is the value of the fourth term?

3. A driver travels 4 miles every 10 minutes. His distance along a straight road is a function of time.

Type of function: _________________

Reason:

4. The number of footballs in each image is a function of the image number.

Type of function: _________________

Reason:

Images from VisualPatterns.org

Formative Assessment 7 Tier 3

F.LE.1: Distinguish between situations that can be modeled with linear function and with exponential functions.

Each table of values given below models a pattern that can be modeled by a linear function, an exponential function, or neither. For each table of values given:

a) Determine if the pattern represented is linear, exponential, or neither (1 point).
b) Explain how you decided whether the pattern can be represented by a linear function, an exponential function, or neither (1 point).
c) Complete the table of values by filling in the missing components (2 points).

1.

Term #	1	2	3	5	
Term Value	3	6		15	192

a. Type of pattern:

b. Explanation:

c. Complete missing components in the table.

2.

Term #	1	2	3	5	
Term Value	3	6		48	192

a. Type of pattern:

b. Explanation:

c. Complete missing components in the table.

3.

Term #	1	2	3	4	
Term Value	6	9	14		54

a. Type of pattern:

b. Explanation:

c. Complete missing components in the table.

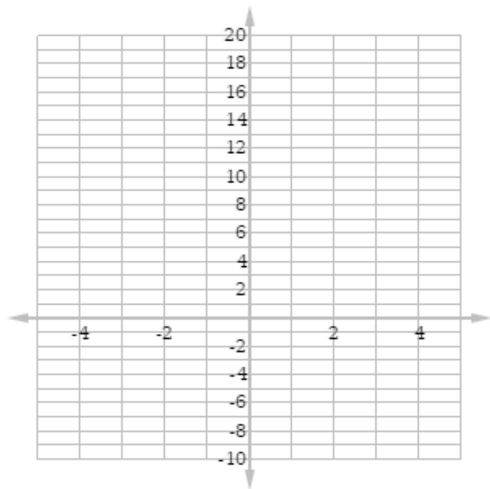
Formative Assessment 8 Tier 1

A.CED.2: Create equations in two or more variables to represent relationships between quantities; graph equations on coordinate axes with labels and scales.

First, write an explicit equation for the continuous function, $f(x)$, that models the data represented in each table (3 points). Then graph $f(x)$ on the axes provided (3 points).

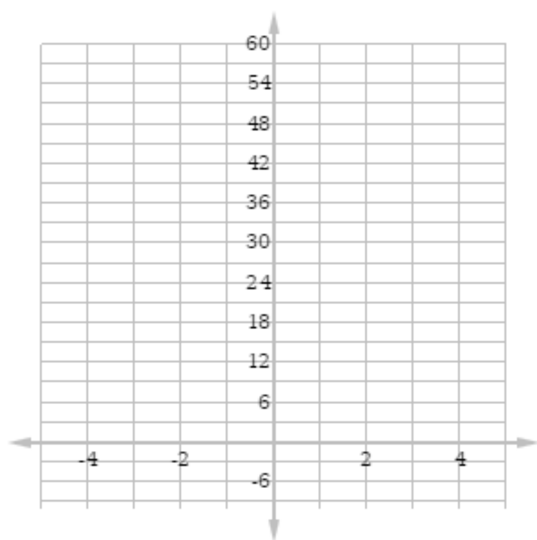
1.

x	$f(x)$
0	3
1	7
2	11
3	15



2.

x	$f(x)$
0	2
1	6
2	18
3	54



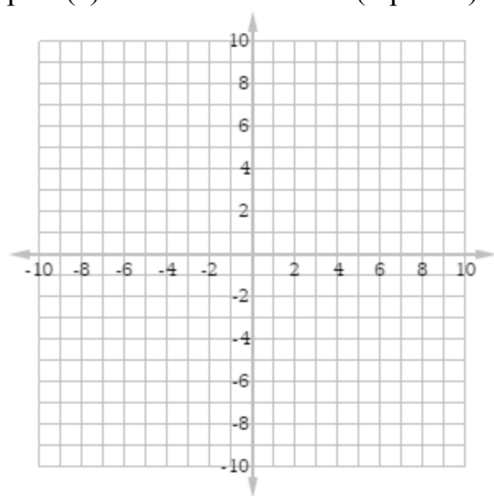
Formative Assessment 8 Tier 2

A.CED.2: Create equations in two or more variables to represent relationships between quantities; graph equations on coordinate axes with labels and scales.

1. A continuous line is modeled by the data given in the table below.

x	y
3	-2
4	2
5	6
6	10

b. Graph the new line that you created in part (a) on the axes below. (3 points)



a. Write down an equation of a new line that has a smaller slope, but a larger y-intercept. (3 points)

2. The continuous function $g(x)$ is represented ordered pairs provided in the table below.

x	1	2	3	4	
g(x)	162	54		6	23

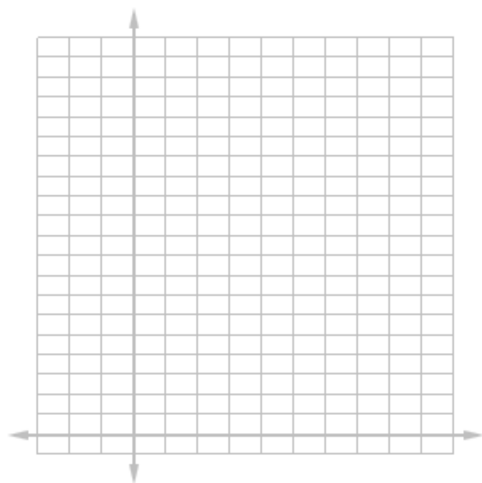
a. Fill in the missing values in the table.

(2 points)

b. Write an explicit equation that models $g(x)$. (3 points)

$g(x) =$

c. Graph $g(x)$ on the axes provided. Be sure to appropriately label and scale the axes. (3 points)



Formative Assessment 8 Tier 3

A.CED.2: Create equations in two or more variables to represent relationships between quantities; graph equations on coordinate axes with labels and scales.

Two strains of a zombie virus have been discovered.

Strain 1: Patient zero, the original zombie, infects an average of 12 new people every minute. However, zombies that contract the virus second hand (meaning they are infected from the original zombie), are unable to spread the infection. Only the first person to contract this virus, the original zombie, is able to spread it.

Strain 2: Patient zero, the original zombie, infects an average of 1 person every minute. Every person who is infected by this strain infects another person each minute.

Let t represent the number of minutes that have passed since either virus was first introduced.

Let $f(t)$ represent the number of people who are infected with Strain 1 of the virus.

Let $g(t)$ represent the number of people who are infected with Strain 2 of the virus.

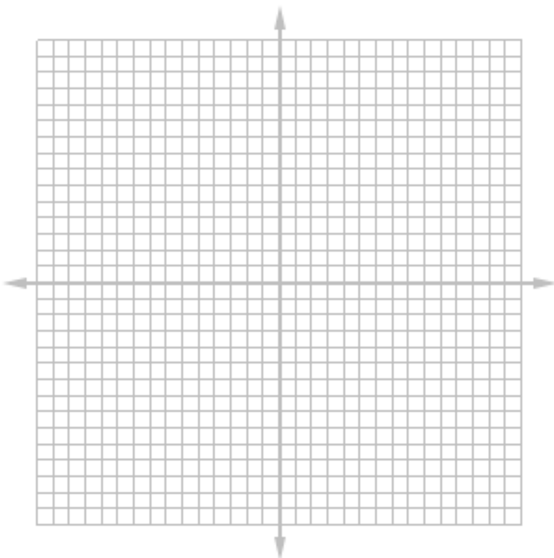
1. Write an explicit equation that represents the continuous function $f(t)$. (3 points)

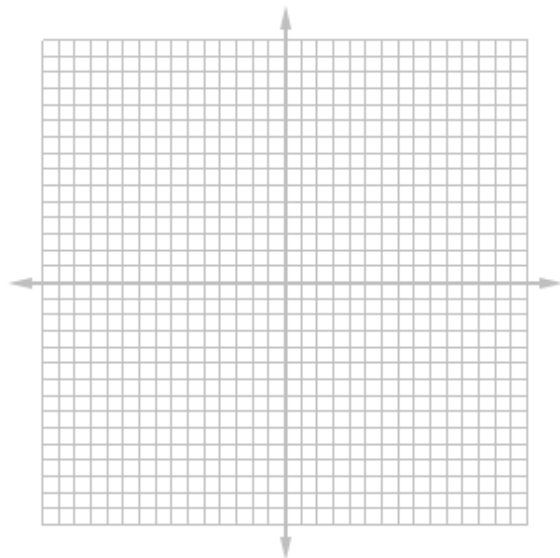
$$f(t) =$$

2. Write an explicit equation that represents the continuous function $g(t)$. (3 points)

$$g(t) =$$

3. Graph the continuous functions $f(t)$ and $g(t)$ on the axes below. Be sure to include all appropriate components of a well-done graphic representation. (6 points)

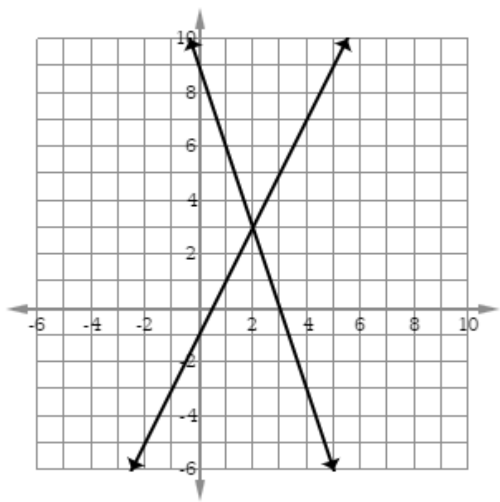




Formative Assessment 9 Tier 1

A.REI.6: Solve systems of linear equations exactly and approximately (e.g., with graphs), focusing on pairs of linear equations in two variables.

Solve each system given below. Write the solution the system in the space provided.

1.  Solution: _________ (2 points)	2. $8x-2y=22$ $6x+3y=21$ Solution: _________ (3 points)
3. $y=3$ $2x-3y=11$ Solution: _________ (3 points)	4. $2x-4y=8$ $y=12x-2$ Solution: _________ (3 points)

Formative Assessment 9 Tier 2

A.REI.6: Solve systems of linear equations exactly and approximately (e.g., with graphs), focusing on pairs of linear equations in two variables.

1. The two systems of equations given below have the same solution. First determine the solution of System 1. Then determine the value of ***A***, and write your answer in the space provided.

System 1	System 2
$2x-y=3$	$x+5y=7$
$3x-2y=4$	$y=\boldsymbol{A}x+3$

A= __________ (5 points)

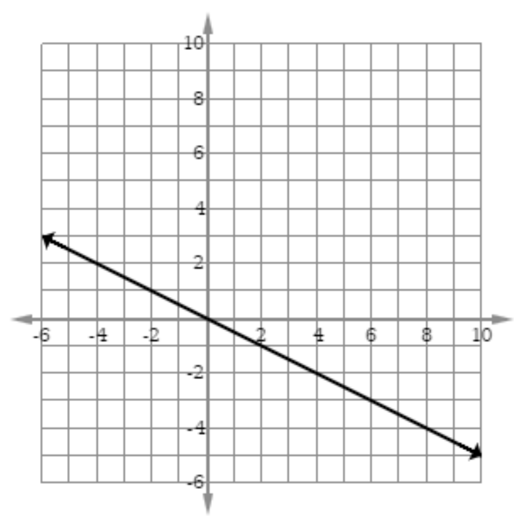
2. The system of equations given below has infinitely many solutions. Determine the value of B and write your answer in the space provided.

$6x-3y=9$

$y+B=2x$

B = _________ (3 points)

3. The solution to a system of equations is (4, -2). One of the lines from this system is graphed below. Graph a possible second line that completes this system (2 points).



Formative Assessment 9 Tier 3

A.REI.6: Solve systems of linear equations exactly and approximately (e.g., with graphs), focusing on pairs of linear equations in two variables.

1. The two systems of equations given below have the same solution. Fill in the blanks with any integer to complete the second system.

System 1	System 2
$5x+4y=-19$	$y=5+2x$
$4x+y=-13$	_____x +______y=______

(6 points)

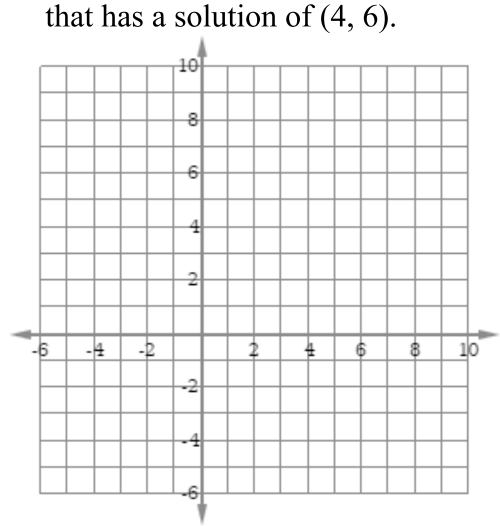
2. The system of equations given below has no solution. Fill in the blanks with any integer to complete the second system.

$8x - 2y = 20$

_____x + ______y = _______

(3 points)

3. Graph a system of linear equations that has a solution of (4, 6).



(2 points)

Formative Assessment 10 Tier 1

F.IF.4: For a function that models a relationship between two quantities, interpret key features of graphs and tables in terms of the quantities, and sketch graphs showing key features given a verbal description of the relationship. Key features include: intercepts; intervals where the function is increasing, decreasing, positive, or negative; relative maximums and minimums; symmetries; and end behavior.

Directions: Match each descriptor to the key feature it describes by writing the number of the descriptor next to the corresponding key feature in the far right column. (1 point each)

Descriptor	Key Features	Match
1. This feature describes all possible input values of a function.	A. Increasing	A. ______
2. This feature is a point that has an x-coordinate of zero.	B. Decreasing	B. ______
3. This feature occurs at a point where a function changes from increasing to decreasing.	C. Positive	C. ______
4. This feature describes the values the output approaches, as the input approaches either positive or negative infinity.	D. Negative	D. ______
5. A function is ________________ when it has positive slope.	E. Domain	E. ______
6. A function is ________________ when it is below the x-axis.	F. Range	F. ______
7. This feature is a point that has a y-coordinate of zero.	G. End Behavior	G. ______
8. This feature occurs at a point where a function changes from negative slope to positive slope.	H. X-Intercept	H. ______

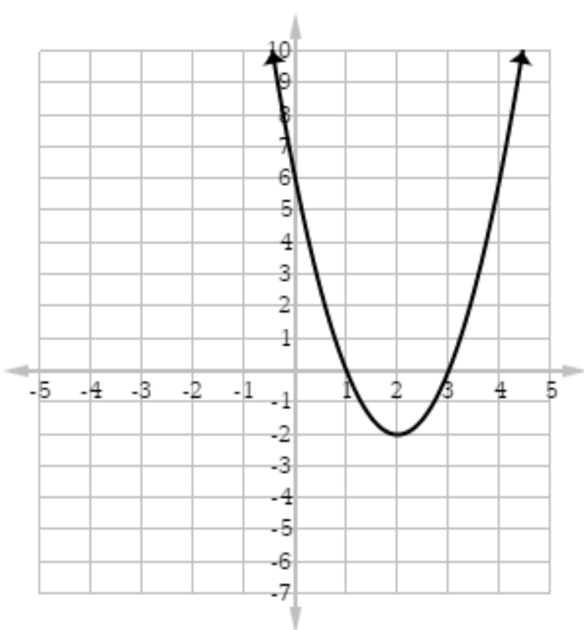
9. A function is ____________________ when it is above the x-axis.	I. Y-Intercept	I. ______
10. This feature describes all possible output values of a function.	J. Relative Maxi mum	J. ______
11. A function is ____________________ when has negative slope.	1. Relative Mini mum	K. ______

Formative Assessment 10 Tier 2

F.IF.4: For a function that models a relationship between two quantities, interpret key features of graphs and tables in terms of the quantities, and sketch graphs showing key features given a verbal description of the relationship. Key features include: intercepts; intervals where the function is increasing, decreasing, positive, or negative; relative maximums and minimums; symmetries; and end behavior.

Directions: Identify the indicated key features for each function. Write any intervals in either interval or inequality notation. (Example- the interval from a to b could be written as (a, b), [a, b], $a < x < b$, or $a \leq x \leq b$, depending on whether the end points a and b should be included in the interval. (1 point each)

1. The function $f(x)$ is graphed below. Answer questions A through D using the graph of $f(x)$



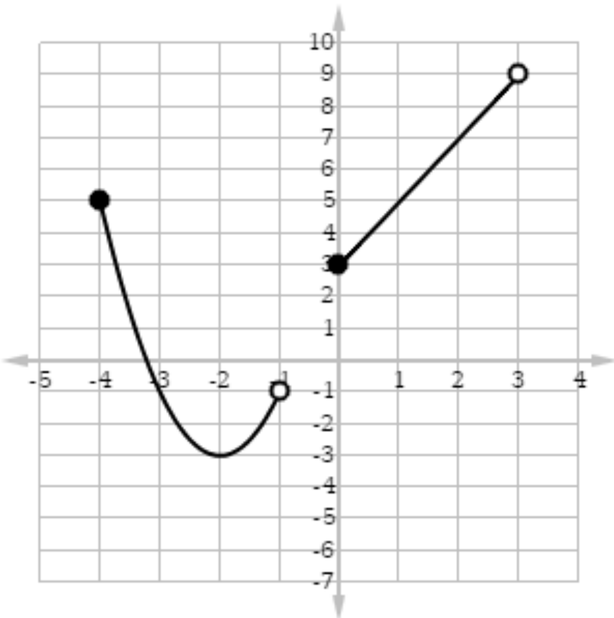
A. State all intervals where f(x) is negative.

B. State all intervals where f(x) is increasing.

C. State all intercepts of f(x). Record the intercepts as ordered pairs.

D. As x approaches ∞ what does f(x) approach?

2. The function g(x) is graphed below. Answer questions E through H using the graph of g(x)



E. State all intervals where g(x) is decreasing.

F. State the domain of g(x).

G. State the range of g(x).

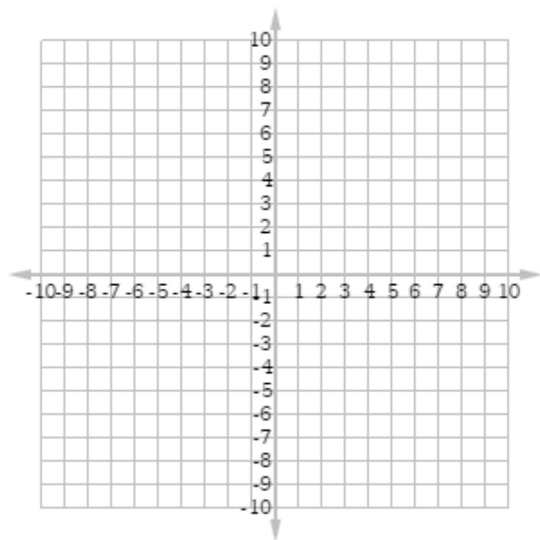
H. Does g(x) have any relative maximums or minimums? If so, indicated what type of relative extrema there is, and record both coordinates of the relative maximum or minimum below.

Formative Assessment 10 Tier 3

F.IF.4: For a function that models a relationship between two quantities, interpret key features of graphs and tables in terms of the quantities, and sketch graphs showing key features given a verbal description of the relationship. Key features include: intercepts; intervals where the function is increasing, decreasing, positive, or negative; relative maximums and minimums; symmetries; and end behavior.

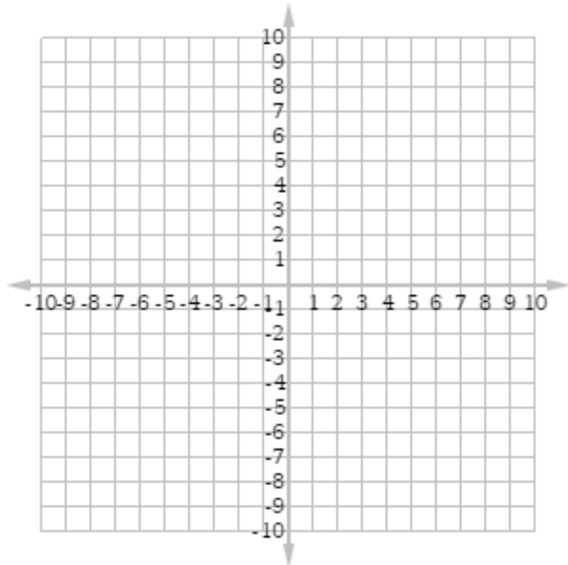
1. The key features given below describe the graph of the function f(x). Sketch a possible graph of f(x) that satisfies these key features. (5 points)

   - f(x) is positive when $x>2$
   - f(x) has one relative extrema
   - As x approaches negative infinity, f(x) approaches zero
   - f(x) is decreasing $-3<x<1$



2. The key features given below describe the graph of the function g(x). Sketch a possible graph of g(x) that satisfies these key features. (5 points)

- g(x) has exactly two x-intercepts.
- As x approaches ∞, g(x) approaches $-\infty$
- g(x) has a relative maximum at $x=-1$
- g(x) is positive when $x<-$



Formative Assessment 11 Tier 1

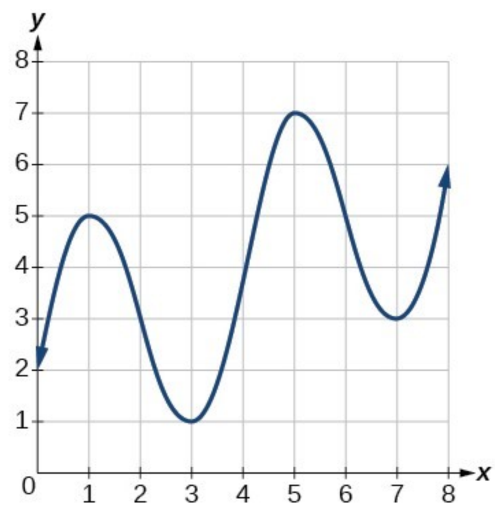
F.IF.6: Calculate and interpret the average rate of change of a function (presented symbolically or as a table) over a specified interval. Estimate the rate of change from a graph.

1. Write down the formula for the average rate of change of f(x), over the interval [a,b]. (2 points)

2. The continuous function f(x) is modeled by the data represented in the table below. Determine the average rate of change of f(x), over the interval [0, 3]. (2 points)

x	f(x)
0	3
1	6
2	12
3	24
4	

3. The function g(x) is represented graphically below. Determine the average rate of change of g(x) over the interval $1 \le x \le 5$. (2 points)



Formative Assessment 11 Tier 2

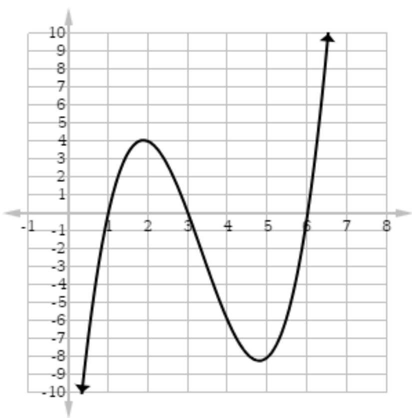
F.IF.6: Calculate and interpret the average rate of change of a function (presented symbolically or as a table) over a specified interval. Estimate the rate of change from a graph.

1. Below are two solutions to the function f(x). The average rate of change of f(x) over the interval $0 \leq x \leq 4$ is 3. Determine the value of k. (3 points)

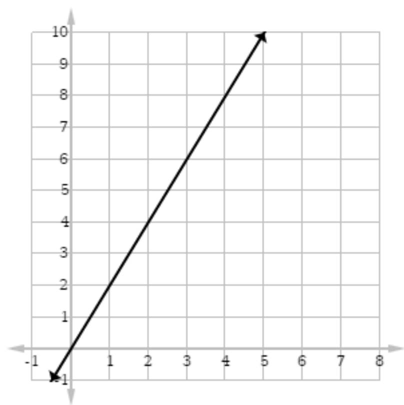
   (0, 3) , (4, k)

2. Four functions are represented below. Arrange these functions so that the value of their average rate of change over the interval [1,4] is organized in increasing order. Communicate this arrangement by writing their associated letter in the space provided. (4 points)
3.

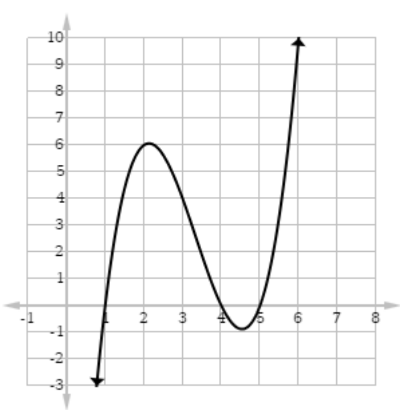
a.



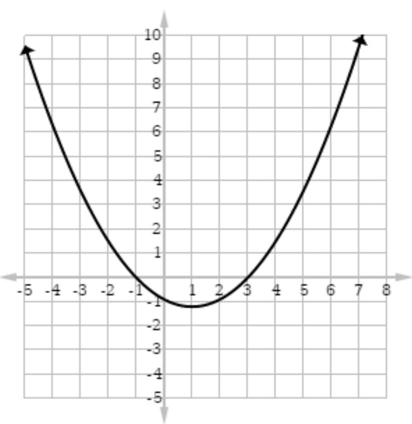
b.



c.



d.



Increasing order: ________ , ________ , ________ , ________

Formative Assessment 11 Tier 3

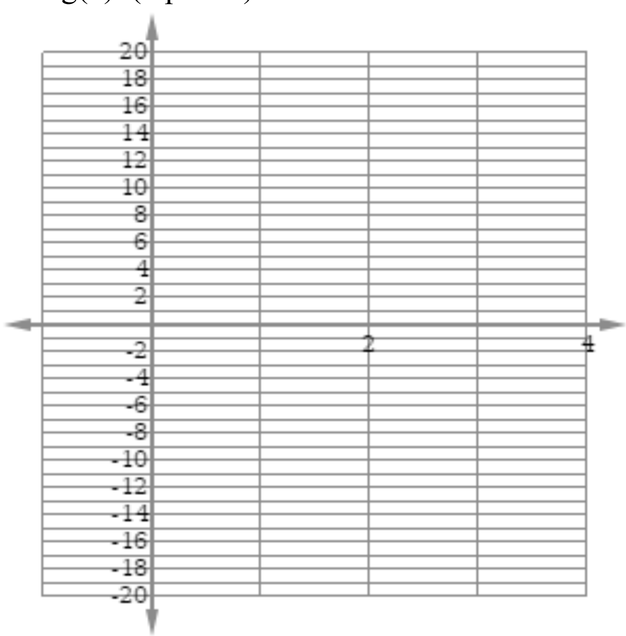
F.IF.6: Calculate and interpret the average rate of change of a function (presented symbolically or as a table) over a specified interval. Estimate the rate of change from a graph.

1. The average rate of change of f(x) over the interval [0, 3] is 42. Find a value of k. (4 points)

$$f(x) = 2 * k^x$$

2. The average rate of change of the continuous function f(x) is 3, over all intervals $-\infty \leq x \leq \infty$. What kind of function is f(x)? (1 point)

3. It is known that the average rate of change of the continuous function g(x) $0 \leq x \leq 3$, is larger than the average rate of change of f(x) $0 \leq x \leq 3$. If f(x) = 2x, graph a possible function g(x). (4 points)



Formative Assessment 12 Tier 1

F.BF.2: Write arithmetic and geometric sequences both recursively and with an explicit formula, use them to model situations, and translate between the two forms.

1. 9, 13, 17, 21, … is an arithmetic sequence. The index for the first term given is 1.

    a. Write an explicit equation that models this sequence. (3 points)

    b. Write a recursive equation that models this sequence. (3 points)

2. 2, 8, 32, 64, …is a geometric sequence. The index for the first term given is 1.

    a. Write an explicit equation that models this sequence. (3 points)

    b. Write a recursive equation that models this sequence. (3 points)

Formative Assessment 12 Tier 2

F.BF.2: Write arithmetic and geometric sequences both recursively and with an explicit formula, use them to model situations, and translate between the two forms.

1. A sequence has a recursive equation of $f(n) = f(n-1) + 3;\quad f(1) = 4$. Write an explicit equation that models this sequence. (3 points)

2. A sequence has a recursive equation of $f(n) = f(n-1) * 2;\ f(1\quad = 3$. Write an explicit equation that models this sequence. (3 points)

3. A sequence has an explicit equation of $f(n) = 12 - 4n$. Write a recursive equation that models this sequence. (3 points)

4. A sequence has an explicit equation of $f(n) = 2(5)^{n-1}$. Write a recursive equation that models this sequence. (3 points)

Formative Assessment 12 Tier 3

F.BF.2: Write arithmetic and geometric sequences both recursively and with an explicit formula, use them to model situations, and translate between the two forms.

Directions:

Use the digits 0-9 to complete the first three terms of an arithmetic and geometric sequences by filling in each box with only one digit (2 points). Then create an explicit and recursive equation that models each sequence (12 points total).

Arithmetic Sequence: ☐, ☐, ☐ ...

Geometric Sequence: ☐, ☐, ☐☐ ...

Explicit equation for the arithmetic sequence (3 points):	Explicit equation for the geometric sequence (3 points):
Recursive equation for the arithmetic sequence (3 points):	Recursive equation for the geometric sequence (3 points):

Appendix D:

Original Assessment Reflection (Weeks 1–8)

Consider your performance on this assessment while you answer questions 1 and 2.

1. Use the scale below to rate the difficulty of the question you just completed.

1	2	3	4	5
Very Easy	Somewhat Easy	Neutral	Somewhat Difficult	Very Difficult

2. This assessment item is worth ____ points. Approximate the number of points that you think you will earn.

Consider the content objectives that correlate to this assessment. With your performance in mind, respond to either question three or four.

3. If you believe that your performance on this assessment is **strong**, respond to the following prompt:

   Why do you think that you have successfully learned this content? What actions or experiences have helped you to develop a strong understanding of this content?

4. If you believe that your performance on this assessment **needs improvement**, respond to the following prompt:

   Why do you think that understanding content has been challenging? What is it about this content, or your learning experiences related to this content, that made it more difficult? What are your next steps to improve your understanding?

Revised Assessment Reflection (Weeks 9–12)

Consider your performance on this assessment while you answer questions 1 and 2.

1. Use the scale below to rate the difficulty of the question you just completed.

1	2	3	4	5
Very Easy	Somewhat Easy	Neutral	Somewhat Difficult	Very Difficult

2. This assessment item is worth ____ points. Approximate the number of points that you think you will earn.

Consider the content objectives that correlate to this assessment. With your performance in mind, respond to either question three or four.

3. If you **do not intend to retake** this assessment respond to the following prompt:

   Why do you think that you have successfully learned this content? What actions or experiences have helped you to develop a strong understanding of this content? Are there any questions that you think you answered incorrectly? If so, which ones, and what about these questions were difficult?

4. If you **intend to retake** this assessment, respond to the following prompt:

   Which questions did you have the most trouble with, and why do you think that is? What is it about this content, or your learning experiences related to this content, that made it more difficult? What are your next steps to improve your understanding?

Appendix E:

Parent Consent Form

Dear Parents/Guardians,

My name is Alison Espinosa, and I am your student's Secondary Math 1 Honors teacher. I am also a doctoral candidate in the Department of Instruction & Education at the University of South Carolina. I am conducting an action research study as part of the requirements of my degree in Curriculum and Instruction, and I would like to invite your child to participate.

I am studying the impact of standards-based grading practices on my students' ability to accurately and critically self-assess their performance on formative assessments. As supporting students' self-assessment of their performance and learning is known to increase student motivation, self-efficacy, and achievement, the results of this study will be used to help me improve my teaching practice.

By permitting your student to participate, you are allowing me to include my analysis of your students' work in my dissertation. The results of this study may be published or presented at professional meetings, but your student's identity will not be revealed.

Participation is completely voluntary and confidential. Participation, non-participation, or withdrawal will not affect your student's grade, or effect the academic expectations or requirements in this class. Study information will be kept in a secure location, and your student's name will not be reported or associated with their work. Participation is anonymous, which means that no one (not even members of my dissertation committee) will know your student's name. Your student's name will be removed from all of the research materials prior to data analysis and storage.

You may contact me (alison.espinosa@xxxschools.org) or my faculty advisor, Dr. Leigh D'Amico, Ed.D. (damico@mailbox.sc.edu) if you have questions about this study. If you have any questions about your student's rights as a research participant, you may contact the Office of Research Compliance at the University of South Carolina at 803-777-7095.

Please indicate your preference by completing and signing the respective section of the attached letter. You may keep this copy of the letter for your personal records.

With kind regards,

Alison Espinosa
alison.espinosa@xxxschools.org

By signing this consent form, I confirm that I have read the information in this parental permission form and have had the opportunity to ask questions regarding this study.

Consent: I voluntarily agree to allow my child to take part in this study.

Child's name ________________________________
Parent/Guardian's Name ________________________________
Parent/Guardian's Signature ________________________________

Dissent: I do not agree to allow my child to take part in this study.

Child's name ________________________________
Parent/Guardian's Name ________________________________
Parent/Guardian's Signature ________________________________

Appendix F:

Week 8 Exit Ticket

When we complete an independent warm-up, I ask you to self-assess your performance by estimating the number of points you think you will earn, while also reflecting on whether the topic is something you have a strong understanding of, or something that needs improvement. Now I want you to reflect on your experiences doing this.

1. Use the scale below to indicate how valuable you think this experience is. Do you think that self-assessing your performance and reflecting on your understanding is worthwhile?

1	2	3	4
These experiences have no value to me.	I don't see the value of these experiences.	These experiences are somewhat valuable to me.	These experiences are valuable to me.

2. Explain why you answered #1 in this way.

Appendix G:

Changes in Student Achievement During Baseline and Treatment Cycles

Student	Baseline Pre-Test	Baseline Post-Test	Percent Change	Treatment Pre-Test	Treatment Post-Test	Percent Change
1	5	4	-8.33	8	10	8.33
2	9	11	16.66	14	19	20.83
3	3	8	41.66	8	15	29.16
4	5	5	0	10	13	12.5
5	7	9	16.66	20	17	-12.5
6	8	5	-25	11	14	12.5
7	11	12	8.33	19	20	4.16
8	8	10	16.66	12	17	20.83
10	7	7	0	7	14	29.16
11	10	11	8.33	17	20	12.5
12	9	10	8.33	15	18	12.5
13	5	7	16.66	10	14	16.66
14	7	8	8.33	13	12	-4.16
15	6	9	25	11	18	29.16
16	7	7	0	11	16	20.83
17	10	11	8.33	18	21	12.5
18	5	9	33.33	14	14	0

19	11	10	-8.33	16	19	12.5
20	11	11	0	16	22	25
21	10	10	0	17	22	20.83
22	3	2	-8.33	6	6	0
23	9	11	16.66	15	24	37.5
25	12	12	0	18	22	16.66
26	11	10	-8.33	15	19	16.66
27	10	9	-8.33	17	21	16.66
28	5	8	25	11	17	25
29	3	6	25	8	11	12.5
30	8	11	25	9	16	29.16
31	10	10	0	14	16	8.33
32	11	11	0	14	20	25
34	9	9	0	14	20	25
Mean			7.53			15.49
Standard Deviation			14.49			11.21