12-2021

Examining Opportunities for Higher-Order Thinking in Secondary and Middle-Level Proficiency-Based Mathematics Classrooms

Quinton Donahue
quinton.donahue@maine.edu

EXAMINING OPPORTUNITIES FOR HIGHER-ORDER THINKING IN SECONDARY AND MIDDLE-LEVEL PROFICIENCY-BASED MATHEMATICS CLASSROOMS

By

Quinton S. Donahue

B.S. University of Maine, 2004

M.A. University of Maine, 2010

A DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

(in Educational Leadership)

The Graduate School

The University of Maine

December 2021

Advisory Committee:

Ian Mette, Associate Professor of Educational Leadership, Advisor

Catharine Biddle, Associate Professor of Educational Leadership

Justin Dimmel, Assistant Professor of Mathematics Education and Instructional Technology

Richard Ackerman, Professor of Educational Leadership

Betsy Webb, Libra Professor of Educational Leadership

EXAMINING OPPORTUNITIES FOR HIGHER-ORDER THINKING IN SECONDARY AND MIDDLE-LEVEL PROFICIENCY-BASED MATHEMATICS CLASSROOMS

By

Quinton S. Donahue

Dissertation Advisor: Dr. Ian Mette

An Abstract of the Dissertation Presented
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
(in Educational Leadership)
December 2021

In 2011, the state of Maine introduced a statute that required all diplomas to be awarded on the basis of proficiency by 2017. During this period of time, school districts moved to incorporate Proficiency-based Education (PBE) systems of instruction and learning. This prompted educators to bring clarity to standards, assessment, and grading practices. In many schools, they also began to focus on the idea that mastery drives movement. Districts were encouraged to integrate higher-order thinking opportunities for students when making these changes to their learning system. Despite this guidance, it is not clear as to the extent that higher-order thinking opportunities are available for students in proficiency-based classrooms. The mastery-drives movement approach may indicate that these opportunities are only available for students who have attained proficiency. There is also a concern that students may not pursue these opportunities if they are not required of them.

This research explored the frequency of higher-order thinking opportunities available for mathematics students in middle and secondary schools that made the move to a proficiency-based system. Additionally, the research aimed to discover how students are engaged in higher-

order thinking and what factors may enhance or impede the opportunities. The research used an explanatory sequential mixed-methods design. Two rounds of 50 observations were conducted within five different schools (2 middle schools and 3 high schools) using the Instructional Practices Inventory (IPI) tool. In a sixth school, one round of 50 observations took place. In the end there were 550 total data points. The research also included 11 total focus groups with mathematics teachers.

The results suggest that high school math students in a proficiency-based structure tend to receive more opportunities for higher-order thinking within their classrooms. The research also pointed to individual pace learning structures impacting student ability to access higher-order thinking opportunities and suggested that the advanced students receive the bulk of the opportunities in PBE classrooms.

DEDICATION

This dissertation is dedicated to my unbelievable wife, Jessica Donahue. She has supported me from the moment I took on this commitment. She never complained when weekends or evenings were spent with reading and writing. Instead she told me, “You’ve got this!” Thank you for believing in me when I was not sure I believed in myself. It is safe to say that I could not have done this without you!

ACKNOWLEDGEMENTS

I am greatly thankful for the mathematics teachers and administrators at the six participating schools. They allowed me bounce in and out of their classrooms on two full school days while conducting the observations. They also spent an a couple hours of their time outside of school hours to talk about math. Their passion for the content area and supporting students is highly evident. I would also like to thank the district and school-based administrators at each of the districts I worked with. Without their consent and support this research would have not been possible.

I would also like to thank the staff and students at Gorham Middle School and Mount View Middle School. They have been supportive of me, as I have worked to complete this research while working as a full-time middle school principal. I am very thankful for GMS Assistant Principal, Tom Smith. He has been so supportive as I have completed this work. Additionally, this would have not been possible without the support of my current superintendent, Dr. Heather Perry, and my former superintendent, Dr. Paul Austin. They have encouraged me in this endeavor every step of the way.

Thank you to my amazing dissertation committee chair, Dr. Ian Mette, and the rest of my committee: Dr. Catharine Biddle, Dr. Justin Dimmel, Dr. Richard Ackerman, and Dr. Betsy Webb. Their feedback, guidance and support have been essential.

I was very lucky to make this journey with an outstanding cohort. I cannot thank Katie Thompson, Barbara Moody, Pam Doyen, Theresa Gillis, and Kathy Harris-Smedburg enough. Your support, ideas, laughter, and encouragement helped to make this possible. Thank you! A special thank you to Katie for teaming up with me in the final stretch. I would also like to thank my good friend and fellow principal, Josh Tripp. You have always been there for support and advice.

Finally, I want to thank my family. Thank you Mom, Dad, Sasha, and Grammie. You have all been so positive and you have pushed me to achieve my goals and follow my passion. I could not be more thankful!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In 2011, Maine introduced Statute §4722-A, which mandated that all schools in the state award diplomas based on student proficiency by January 1, 2017. In addition, Maine directed schools to provide students with the opportunity to attain the necessary proficiency through multiple pathways (Maine Revised Statutes, 2015). This proficiency-based philosophy is based on the belief that, to learn effectively, students must demonstrate competency in foundational knowledge prior to moving on to the next phase of learning. If students do not understand the current lesson, additional support must be provided (Maine Department of Education, 2015). On June 5, 2019, Maine Governor Mills signed Public Law 2019, Chapter 202 into law, repealing the requirement for proficiency-based diplomas (Diploma Requirements, 2020). During the eight years that Statute §4722-A was in place, it pushed schools in Maine to reassess their curriculum and means of classroom instruction through the lens of differentiation and personalization. At the national level, proficiency-based education has received more recent attention. The U.S. Department of Education referred to Maine as one of five states that were advanced in their implementation of the proficiency-based model (Brodersen et al, 2017), while research funded by the Bill and Melinda Gates Foundation mentioned five states that are advanced in the process (Sturgis, 2016).

The implementation of the proficiency-based model in Maine has differed from district to district and school to school (Silvernail & Stump, 2014). To aid implementation, many schools have turned to organizations such as the Great Schools Partnership (GSP) for coaching and consultation or joined cohorts such as the Maine Cohort of Customized Learning (MCCL) to access training and resources. Other districts have taken part in regional collaboratives such as the Northern Maine Educational Collaborative (NMEC). Finally, some districts have developed their own frameworks and followed internal interpretations of Statute §4722-A to drive reform (Sturgis, 2016).

In addition to its focus on content area standards, Maine introduced "guiding principles" that each student should possess upon leaving school. For example, schools must ensure that students are creative and practical problem solvers and integrative and informed thinkers (Maine Department of Education, 2015). For educators, this focus has underscored the importance of higher-order thinking in the classroom. Researchers have specifically suggested that one way that students can demonstrate mastery of the guiding principles is by completing assignments aligned with higher-order thinking taxonomies (Stump & Silvernail, 2015). This is also true at the national and international levels, where higher-order thinking taxonomies are often used to assess students' ability to solve problems (Yuan & Le, 2014). However, the pursuit of higher-order thinking is not new. Over the past 60 years, educators have used tools such as Bloom's taxonomy or Webb's depth of knowledge framework to better understand the levels of thinking required for certain classroom tasks. The types of problem solving or integrative thinking tasks referenced by the Maine Department of Education tend to fall on the high end of the frameworks mentioned, where students are asked to extend their thinking or evaluate information. The skills required for such tasks are often referred to as higher-order thinking skills (Bloom, 1956; Krathwohl, 2002; Webb, 2007; Marzano & Kendall, 2007).

The explicit teaching of higher-order thinking strategies within the proficiency-based classroom is supported and referenced in most proficiency-based education circles in Maine (Great Schools Partnership, 2015; Maine Cohort for Customized Learning, 2015; Re-inventing Schools Coalition, 2015; Sturgis, 2016). Some of the clearest guidance regarding the implementation of the state's guiding principles suggests that students demonstrate their attainment through assignments such as portfolios, exhibitions, or capstone projects. Additionally, these principles should be taught in Grades 5–8, as they can help to create equity in the educational experience for all students. Many students do not have an opportunity to work on skills such as complex communication or problem solving at home. By focusing on these areas at an early age, such students can begin to build an educational foundation that is comparable to that of their peers (Great Schools Partnership, 2015).

It is clear that those who pioneered proficiency-based instruction understood that the teaching of higher-order thinking skills (HOTS) is an important component of the new educational landscape. However, it is unclear whether educators have focused on including HOTS in their curriculum design and classroom instruction. In addition, it is unclear whether all students are given the opportunity to participate in tasks that require the use of HOTS from an early age. Since schools in Maine do not have to demonstrate the completion of guiding principles until students graduate, this could easily become an area that elementary and middle school educators overlook (Maine Department of Education, 2015).

Additionally, many schools seem to only offer opportunities for higher-order thinking to students who have attained proficiency before their peers. Once most students reach proficiency, instruction shifts to a different objective. Thus, most students who engage in higher-order thinking already excel at school, which creates an equity issue; lower-performing students can rarely engage in higher-order thinking. On paper, districts in Maine have attempted to address the need for students to engage in a variety of learning structures. For example, schools that are members of the MCCL utilize a taxonomy known as the Marzano complex reasoning framework, which lists four levels of thinking: retrieval, comprehension, analysis, and knowledge utilization. At MCCL schools, each learning target is linked to a level of thinking in Marzano's taxonomy, often the retrieval or comprehension stage. Thus, students are only given the opportunity to tackle higher-order thinking tasks once they have achieved proficiency.

Consequently, lower-achieving students may never be able to participate in higher-order thinking opportunities, or these opportunities may be too infrequent for them (MCCL, 2015; Marzano & Kendall, 2007). This is not a new phenomenon; its prevalence was underscored in a study on the achievement gap and the rigor gap (Barton & Coley, 2009; Torff, 2014). However, considerable research has demonstrated that lower-achieving students greatly benefit from inclusion in higher-order thinking tasks (Schraw & Robinson, 2011; Ritchhart et al., 2011; Silver & Stein, 1996).

Moreover, studies have found that educators frequently subscribe to the belief that there is a hierarchical order to higher-order thinking tasks (Torff, 2008). Thus, students must complete lower-order cognitive tasks prior to pursuing higher-order thinking. For example, a mathematics student might be asked to demonstrate repeated proficiency in operations with fractions before being allowed to utilize fractions in a real-world task such as tripling the serving size of a recipe. Such beliefs create a systematic bias towards low-advantage students (Torff, 2014). However, educators have organized many proficiency-based standards in line with this thinking (MCCL, 2015). This is a significant issue, as the inclusion of HOTS tasks in the classroom has been demonstrated to increase the engagement levels of all students; moreover, students are more successful in their learning when they are allowed to examine topics in greater depth (Matsumura et al., 2008; Paige et al., 2013; Gine & Kruse, 2007; Levenia et al., 2010).

Background for the Study

Initial research in Maine demonstrated that there is a lack of consistency in the implementation of proficiency-based education (PBE) across the state (Silvernail et al., 2014). As a result, districts were given the option to extend the implementation deadline, which would require schools to implement PBE by January 1, 2017 (Rier, 2014). While there seems to be incongruence across districts, schools that have fully introduced proficiency-based models share many of the same educational tenets (MDOE, 2015; Great Schools Partnership, 2015). The latter include the differentiation and use of formative assessments, which help to paint a picture that describes PBE instruction. It is important to consider the tenets of PBE to better understand how higher-order thinking tasks can be incorporated into a classroom structure that must already include these components.

More recent research has demonstrated that schools in Maine are making progress in their implementation (Stump et al., 2016). There is now a greater focus on developing "habits of work" among students to increase accountability. Additionally, districts have made great efforts to better educate the community on the proficiency-based system. However, challenges remain with connecting the tenets of proficiency-based learning to local practices and policies; in addition, consistency remains an issue (Stump et al., 2016).

According to the Maine Department of Education's (2015) definition of PBE, students must complete one set of skills before being allowed to move on to the next. Johnston (2011) defined PBE as including a flexible time component, stating that students can work at their own pace until the concepts are mastered. From this perspective, education and instruction are highly student-centered and driven by the individual attainment of proficiency. The importance of ongoing assessment is also emphasized. The key difference between these two views of proficiency is that Maine considers proficiency to be determined through time spent on a topic, whereas Johnston emphasizes the pace set by individual needs. Regardless of how this is framed, an important theme in PBE systems is a less rigid pace.

The Great Schools Partnership (2015) developed a model called "Proficiency-Based Learning Simplified" to create a clear structure and definition for standards-based systems. It is grounded in 10 principles that have been linked to research as strong educational practices. These include transparency of learning targets, a common curriculum, standards that determine success, the use of formative and summative assessments, clearly delineated academic achievements and habits of work, grades that indicate both learning progress and places to grow, the ability for students to retest if they do not meet a standard, the availability of differentiated instruction and multiple pathways, and valuing student voice and choice. These practices are supported by research and help support student learning (Brookhart et al., 2011; Hess & Gong, 2014; Dean et al., 2012).

Another term that tends to be used interchangeably with PBE is competency-based education (CBE). The Reinventing Schools Coalition (2015), a division of Marzano Research, refers to CBE as "an educational system where students are placed in developmentally appropriate levels, and receive instruction on the competencies required to move to the next level." This definition indicates an approach to PBE that is similar to the one taken in Maine; however, it examines PBE from a group perspective rather than at the individual level.

A report that summarizes the state of CBE in New England used the term as a catchall and included Maine's efforts under this label. According to the report, CBE includes eight components: student advancement after demonstration of mastery, explicit and transparent learning targets, timely and differentiated support, assessments within a learning cycle, application of a broad set of skills (including critical thinking) among students, nurturing of a growth mindset, development of intrinsic motivation, and district-wide accountability (Sturgis, 2016). On the other hand, Johnstone and Soares (2014) referenced students' ability to learn at their own pace in a CBE education system but refer to the system as a whole as a disruptive innovation.

In summary, the term 'proficiency-based education' or 'PBE' is typically linked to several key educational components that focus on the individualization of education. As districts continue to develop their own proficiency-based systems, these are the areas where structural guidance is most frequently focused (Great Schools Partnership, 2015; MDOE, 2015; MCCL, 2015; Sturgis, 2016). Therefore, this research considers these areas to examine the perceptions of teachers in a proficiency-based classroom. However, it remains unclear how schools are integrating higher-order thinking into proficiency-based classrooms. Are proficiency-based educators setting rigorous standards for all students who require HOTS, or are they differentiating instruction by providing higher-order thinking opportunities for students who have met minimum proficiency standards? Is student choice a factor in students receiving higher-order thinking opportunities? It is essential to explore such questions to better understand how higher-order thinking fits into the proficiency-based classroom.

Defining Higher-Order Thinking

When describing the cognitive levels that students must call upon to complete classroom tasks, one may hear educators refer to rigor, critical thinking, depth of knowledge, complex reasoning, cognitive rigor, higher-order thinking, and cognitive complexity (Blackburn, 2013; Marzano et al., 1993; Webb, 2002; Hess et al., 2009). For the purposes of this research, I usually refer to these cognitive levels as "higher-order thinking skills." In fact, my working definition of

HOTS consists of “skills that enhance the construction of deeper, conceptually driven understanding” (Schraw & Robinson, 2011, p. 2).

Problem Statement

At the onset of this study, the research on PBE in Maine had focused on implementation (Stump et al., 2016; Silvernail et al., 2014). Additionally, there was considerable information available on the benefits of a proficiency-based system from those with a vested interest (MDOE, 2015; GSP, 2016; MCCL, 2015; Sturgis, 2016). Given Maine’s proficiency-based requirements and the variability of structure from school to school, it is essential to critically question whether all students benefit from the current structure. There is a concern that organizing tasks from least to most complex, as many schools have done, will lead to a rigor gap (Torff, 2014; Torff, 2008). This concern is underscored by recent research in which educators reported lower levels of rigor in proficiency-based classrooms (Johnson & Stump, 2018).Therefore, it is important to better understand whether all students have the same opportunities to engage in higher-order thinking tasks in a PBE classroom.

It is clear that HOTS tasks are an important component of student success and engagement from an early age (Matsumura et al., 2008; Paige et al., 2013; Gine & Kruse, 2007). This is especially true in mathematics classes, where the principle of depth over breadth has led to greater student achievement (LaVenia, 2010). While current PBE cohorts and collaborative efforts in Maine have highlighted the importance of HOTS opportunities, it is unclear how frequently these opportunities are given to students in PBE classes and whether they are available to all students. Early research on PBE implementation in Maine addressed districts' focus on habits of work but did not mention HOTS (Stump et al., 2016). However, a belief among educators that rigor is lacking in PBE classrooms emphasizes the importance of conducting research in this area (Johnson & Stump, 2018).

Most districts in Maine have made considerable efforts to implement a PBE system prior to the graduation requirement deadline. They were tasked with developing rigorous standards tied to skills that students must master to be successful in college and later in life (MDOE, 2015). To accomplish this, many districts relied on external support and sought guidance with clear tenets and structures. However, not all middle and secondary schools have set the proficiency level for their standards at an appropriate level of rigor that would require the use of HOTS, which is evidenced by students falling far short in standardized assessments that require greater rigor (Croft et al., 2014).

Instead, HOTS are often formally addressed through projects prior to graduation or as extension activities for students who have already completed tasks that demonstrate proficiency (Great Schools Partnership, 2015; Marzano & Kendall, 2007). As educators, particularly mathematics educators, begin to build their own frameworks for proficiency-based models at their schools, research indicates that there are important shifts in classroom practices that must take place. For example, connecting learning to prior knowledge, focusing on explanation and understanding, and engaging students in a productive struggle are recommended components of a proficiency-based math classroom (McGatha & Bay-Williams, 2013). If classroom structures continue to focus on practicing routines and algorithms while disregarding these components, students may never get a chance to pursue topics in greater depth. In such an approach, educators only focus on learning targets, not on depth of knowledge tasks. This challenge can differ from school to school depending on whether higher-order thinking was considered when choosing standards. Research has shown that building a school culture that promotes higher-order thinking is essential (Schraw & Robinson, 2011).

Specifically with regard to mathematics, the shift to Common Core Standards in the Maine Learning Results in 2012 created additional challenges for teachers and students. Research has demonstrated that schools that made the move to Common Core for mathematics needed professional time and support to change their instructional style, teaching resources and other opportunities to collaborate with their peers. If these elements are not in place, mathematics teachers and students alike may feel a level of frustration (Walters et al., 2014)

Another concern is the potential lack of HOTS tasks assigned to historically underachieving students. In an analysis of the achievement gap, Barton and Coley (2009) referenced a watered-down curriculum for low-achieving students. This appears to be a valid concern when it comes to differentiated curriculum for all learners. Some students may never get an opportunity to use HOTS. According to Schraw and Robinson (2011), considerable research exists shows that low-advantage students benefit from participation in critical thinking opportunities.

Justification for the Study

At the national level, many states have devoted considerable effort to reviewing or revising policies related to CBE. These policy changes have focused on credit flexibility, progression flexibility, and individual learning options (Brodersen et al., 2016). Both Maine and Vermont implemented future graduation requirements that required demonstration of proficiency (MDOE, 2016; VAOE, 2015). In a 2017 report, the U.S. Department of Education classified the following five states as advanced CBE states: Iowa, Kentucky, Maine, New Hampshire, and Oregon. All these states, along with Colorado, have utilized and supported competency-based collaboratives at the state level (Brodersen et al., 2016). However, how higher-order thinking tasks fit into these structures is not as clear. Nevertheless, Maine has attempted to address HOTS in its guiding principles (MDOE, 2015). It is imperative to better understand how schools integrate HOTS opportunities into their classrooms before students enter high school. This research provides clear insights on the levels of higher-order thinking taking place in proficiency-based classrooms. In addition, the results could support districts as they work on refining their own PBE systems. This is important because the move to proficiency has helped educators place a greater emphasis on instructing the individual student, although it remains unclear whether proficiency-based structures have considered the value of depth of understanding in learning targets. A focus on attainment, the documentation of proficiency, differing learning paces, differentiated instruction, and student choice may hinder opportunities for all students to use HOTS. This is important not only for Maine but also throughout the United States, where proficiency-based models are gaining traction.

At the national level, the National Council of Teachers of Mathematics (NCTM) has identified the importance of conducting research on how different curriculum approaches can support or impede students' mathematical proficiency (Arbaugh et al., 2008). This research examines the integration of higher-order thinking in proficiency-based mathematics classrooms. While it does not explicitly address this question, it assesses how educators navigate structures associated with proficiency-based learning and introduce tasks that have a positive impact on students' mathematical learning (LaVenia, 2010).

The inclusion of HOTS in classroom tasks can lead to greater student engagement and higher participation rates (Matsumura et al., 2008; Paige et al., 2013). Such activities can also help students to improve their real-world problem solving skills. These students have greater academic success in general, as evidenced by their standardized test scores (Cain, 2002). Thus, it is imperative for activities that require the use of HOTS to be included in all classrooms. Moreover, specifically regarding mathematics, research has shown that opportunities to explore a topic in greater depth and with fewer learning targets leads to high levels of student success (Gine & Kruse, 2007; LeVenia et al., 2010). Mathematics teachers in a proficiency-based system must consider the pressures of ensuring that all students attain proficiency while exploring ways to foster higher-order thinking. Therefore, it is crucial to understand how and to what extent HOTS tasks are currently integrated into middle and secondary school PBE instruction.

Purpose

The purpose of this study is to examine proficiency-based classrooms in middle and secondary school, with a focus on mathematics classrooms. The research considers opportunities to pursue authentic learning opportunities that require the use of HOTS for students in proficiency-based classrooms. In addition, it examines the extent to which students have access to opportunities to use HOTS. Finally, it examines whether educators believe that students' levels of engagement and achievement are impacted by the integration of HOTS activities in the classroom.

Research Questions

The research questions are as follows:

1. How frequently are mathematics students engaged in high-order thinking in proficiency-based classrooms?
    a. Are there significant differences across professional support styles in the frequency of higher-order thinking opportunities?
    b. Are there significant differences across grade spans (6–8, 9–12) in the

frequency of higher-order thinking opportunities?

2. How and in what ways are mathematics students engaged in higher-order thinking when taught in a proficiency-based education system?

3. What factors enhance or impede educators' ability to implement higher-order thinking tasks in proficiency-based mathematics classrooms?

Methods

This research uses an explanatory sequential methodology. This research design is used to assess a trend in the data and explain the reasons behind it (Creswell & Plano Clark, 2011). For this research, data was collected in two rounds using walkthrough classroom observations. The data was then shown to participants in teacher focus groups, who were asked to help explain the results. The first research question was addressed through 550 middle and secondary school classroom observations using the Instructional Practices Inventory (IPI) observation tool (Valentine, 2007). This tool was used to assess the level of students engaged in higher-order thinking in the classroom. Then, the data was examined using both inferential (*t*-test and ANOVA) and descriptive methods (e.g., frequency, mean, mode, standard deviation, and variance).

The second and third research questions were addressed through two rounds of focus groups consisting of six to eight math teachers. Transcripts from the focus groups were coded using open, axial, and selective coding. The qualitative results were used to help describe the data gathered from the classroom observations.

CHAPTER 2

REVIEW OF LITERATURE

This literature review begins with a background on PBE in Maine and at the national level. It reviews the history of PBE implementation and its various support structures, underscoring the most common components of a proficiency-based system. Next, the chapter presents an overview of pedagogical practices in the United States to underscore the dichotomy between the behaviorist and constructivist views on knowledge acquisition. One might see this tug-of-war within the PBE system with an emphasis on both common standards for all and individualization of instruction. Next, research related to the use of HOTS in the classroom and HOTS frameworks is reviewed to emphasize the importance of higher-order thinking and demonstrate how educators assess its level of integration. A special focus is given to structures of higher-order thinking in the mathematics classroom, as this is the current study's area of focus. Next, the current state of higher-order thinking in middle and secondary school is examined. This is important, as the research took place in both middle- and secondary-level mathematics classrooms. Finally, a conceptual framework for the research is presented.

Background on Proficiency-Based Education

In 2011, Maine introduced a statute that would require all students to demonstrate proficiency to obtain a diploma by January 1, 2017. Proficiency attainment was linked to a student's educational experiences in core content areas in all years of high school, demonstration of proficiency in the school's learning system, demonstration of proficiency in Maine's Guiding Principles (a component of the Maine Learning Results), and the fulfillment of requirements set by individual schools. The statute was amended to allow for a phase-in approach, beginning with the 2020–2021 graduating class. Along with these guidelines, schools were told that students must be allowed to demonstrate proficiency through multiple pathways and using various types of evidence (Maine Revised Statutes, 2015). The requirement for proficiency-based diplomas was repealed in June 2019 (Diploma Requirements, 2020).

The Maine Department of Education defined PBE as “any system of academic instruction, assessment, grading and reporting that is based on students demonstrating mastery of knowledge and skills they are expected to learn before they progress to the next lesson, get promoted to the next grade level or receive a diploma.” It also stated that the goal of a PBE system was for students to “acquire the knowledge and skills that are deemed essential to success” (MDOE, 2015). This definition clearly puts the onus on mastery prior to movement. While this is the basis for proficiency-based learning throughout the state, PBE schools have instituted multiple instructional practices to support this mission.

This idea of students moving at their own pace appears to have gained considerable momentum in Maine following the publication of a book titled *Inevitable: Mass Customized Learning in the Age of Empowerment* by Schwahn and McGarvey (2011). According to the authors, the current educational system is outdated and industrialized. They stated that “customizing, individualizing, and personalizing education to meet the learning needs of every learner is inevitable” (p. 19). The book emphasizes the importance of students developing their own learning pathways based on individual interests and explains that time on task should be determined by a student’s individual pace. Many people believe that PBE is somewhat interchangeable with mass customized learning due to the book’s widespread dissemination in Maine educational circles after its publication. Through statute (2015), Maine clearly drew a link by stating that students can attain proficiency through “multiple pathways, using various types of evidence.” Although it has pockets of support, mass customized learning has been criticized due to the risk of students not being exposed to content that is currently uninteresting to them. This allows students to self-select out of areas in which practice could have led to the future development of interests, growth, or success (Kellinger, 2012).

It is clear that Maine's structure requires that documentation of student learning move well beyond the level of recording a student's grade at the end of a quarter. Instead, this statute has forced districts to directly confront the idea of a proficiency-based education system. Many districts, such as those in the MCCL have undertaken a curriculum overhaul by breaking down core subject courses into standards, measurement topics, and learning targets. Learning targets are accompanied by a very specific rubric that allows teachers, students, and parents alike to know which skills must be demonstrated for a student to be designated as proficient. As a result, the cohort schools strongly advocate for transparency of learning (MCCL, 2015).

However, this is only one facet of what Maine districts have been tasked with. The second area in which assessment and documentation is required is the state's guiding principles, which stipulate that students must leave school as clear and effective communicators, self-directed and lifelong learners, creative and practical problem solvers, responsible and involved citizens, and integrative and informed thinkers (Maine Department of Education, 2015). However, it is difficult to second-guess the importance of each of these guiding principles in the classroom.

As Maine schools have ventured into proficiency-based education, clear challenges have arisen. One of the most glaring is the lack of consistency from district to district in terms of their understanding of what a proficiency-based system is (Silvernail et al., 2014; Shakman et al., 2018; Evans et al., 2020). Thus, there may be a very different picture of PBE from one school to the next throughout the state or even schools in the same district. Different definitions of proficiency and interpretations of standards have been barriers to implementation for many educators (Evans et al., 2020). This seems to be one of the reasons why the state has given districts the option of extending the deadline for the integration of a PBE system (Rier, 2014). To help schools wade through the murky waters of proficiency, educational organizations have provided frameworks for schools. Upon examination of the latter, some common themes can be found among the suggested educational practices.

Great Schools Partnership

In Maine, the GSP is one of the leaders in unpacking the components of a proficiency-based system. The organization outlined 10 components that are designed to help schools "establish a philosophical and pedagogical" foundation for their work (see Table 2.1).

Table 2.1 GSP 10 Principles of Proficiency-Based Learning

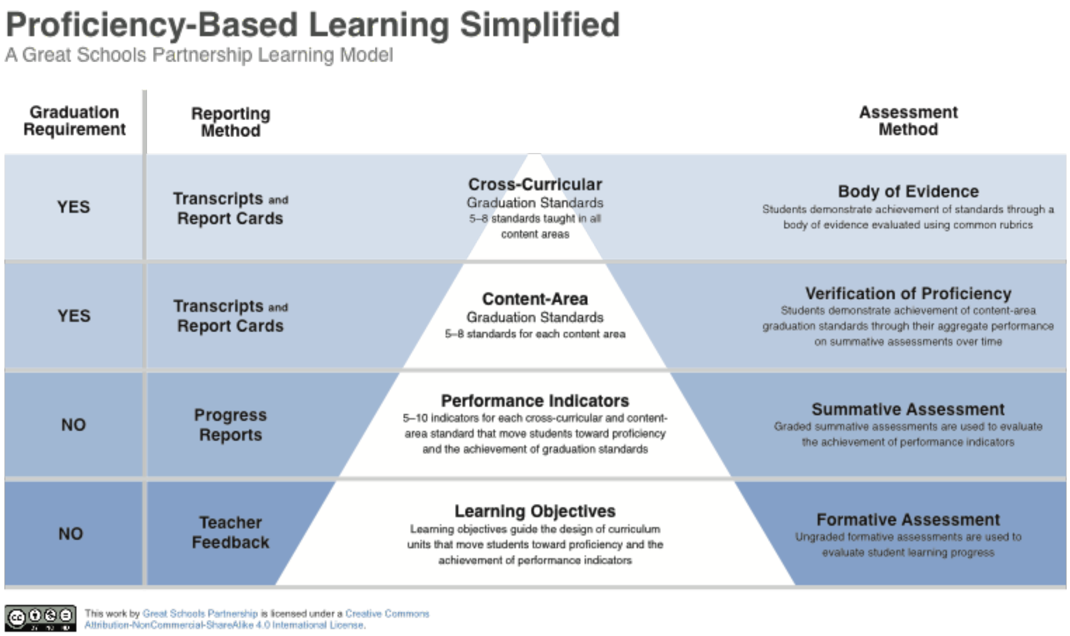
PBE principle (GSP language)	**Summary**
1. All learning expectations are clearly and consistently communicated to students and families, including long-term expectations (such as graduation requirements and graduation standards), short-term expectations (such as the specific learning objectives for a course or other learning experience), and general expectations (such as the performance levels used in the school's grading and reporting system).	Learning expectations Clearly communicated
2. Student achievement is evaluated against common learning standards and performance expectations that are consistently applied to all students regardless of whether they are enrolled in traditional courses or pursuing alternative learning pathways.	Common standards
3. All forms of assessment are standards-based and criterion-referenced, and success is defined by the achievement of expected standards, not relative measures of performance or student-to-student comparisons.	Standards-based assessments
4. Formative assessments measure learning progress during the instructional process, and formative-assessment results are used to inform instructional adjustments, teaching practices, and academic support.	Formative assessments used

Table 2.1 Continued	
5. Summative assessments evaluate learning achievement, and summative-assessment results record a student's level of proficiency at a specific point in time.	Summative assessments used
6. Academic progress and achievement are monitored and reported separately from work habits, character traits, and behaviors such as attendance and class participation, which are also monitored and reported.	Academic progress and work habits split
7. Academic grades communicate learning progress and achievement to students and families, and grades are used to facilitate and improve the learning process.	Grades communicate progress
8. Students are given multiple opportunities to improve their work when they fail to meet expected standards.	Retaking of assignments/assessments/ mastery drives movement
9. Students can demonstrate learning progress and achievement in multiple ways through differentiated assessments, personalized-learning options, or alternative learning pathways.	Differentiation and multiple pathways
10. Students are given opportunities to make important decisions about their learning, which includes contributing to the design of learning experiences and learning pathways.	Student voice and choice

Note. Principles retrieved from http://www.greatschoolspartnership.org/proficiency/

Additionally, GSP developed a learning model that defines key standards, discusses how they should be reported, and explains their assessment method (see Figure 2.1).

Figure 2.1 The GSP Proficiency-Based Learning Simplified Model



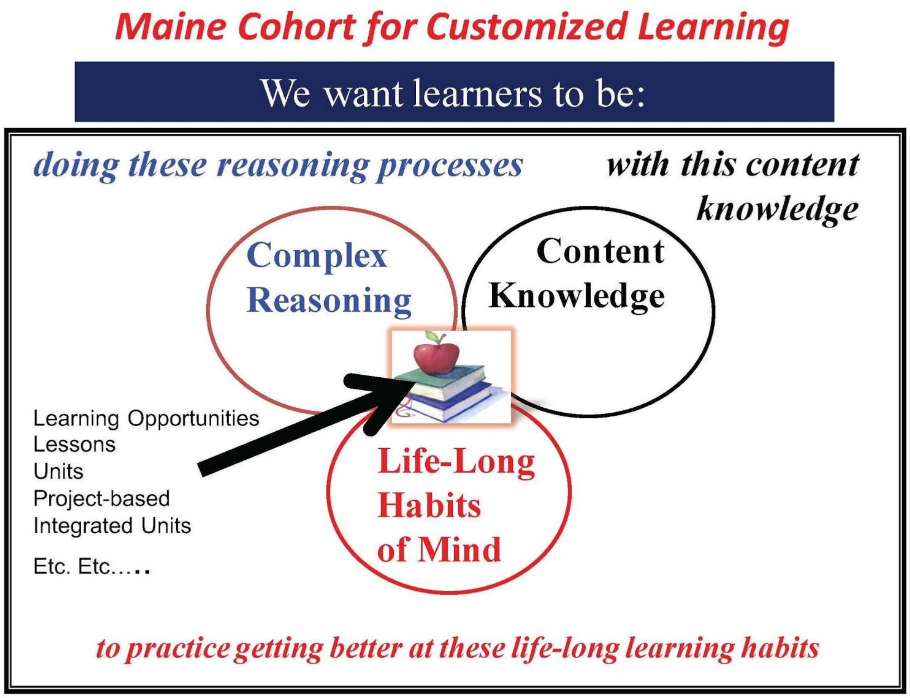
Note. Retrieved from http://www.greatschoolspartnership.org/proficiency/

In particular, the cross-curricular graduation standards, which align with Maine's guiding principles, should be assessed with a body of evidence. Additionally, it is recommended that they are taught in content area classes in Grades 5–8.

Maine Cohort for Customized Learning

The MCCL, a statewide cohort of schools that focus on the implementation of PBE, frames education as a convergence of content knowledge, complex reasoning, and lifelong habits of mind (see Figure 2.2).

Figure 2.2 The MCCL Diagram of Learning



Note. Retrieved from http://mainecustomizedlearning.org/wordpress/wp-content/uploads/2015/12/MCCL-Curriculum-Slide-1.jpg

The MCCL consisted of 13 school district members and had a relationship with the Northern Maine Education Collaborative, which included 17 members. Cohort schools use the same curriculum and grading software. They also have access to various professional development opportunities. The MCCL states that a proficiency-based model should include the seven components listed in Table 2.2.

Table 2.2 The MCCL's Components of a Proficiency-Based Model

PBE component (MCCL language)	Summary
1. Clear definition of what learners need to know, be able to do, and be like	Learning expectations clearly communicated
2. A system of student goal setting and monitoring of their progress to proficiency	Student goal setting/formative assessment use
3. A system that facilitates fluid movement of students among multiple learning opportunities	Multiple pathways/mastery drives movement/differentiation
4. A comprehensive, user-friendly, transparent recording and reporting system	Strong recording system
5. A model of successful integration of technology	Utilizing technology
6. A system of shared leadership, including students, staff, parents, and the community	Shared leadership/student voice
7. Continuous improvement practices embedded in the system	Continuous improvement

Note. Retrieved from http://mainecustomizedlearning.org/goals/

The MCCL addressed HOTS under the term "complex reasoning." Cohort school educators were trained to use the complex reasoning taxonomy developed by Marzano and Kendall (2007). This framework is discussed in greater detail later in this review (MCCL, 2015).

Competency Works Research

In a report based on research funded by the Carnegie Corporation, the Nellie Mae Education Foundation, and the Bill and Melinda Gates Foundation, Sturgis (2016) summarized the state of CBE in New England. The report defines competency education according to five elements, which are listed in Table 2.3.

Table 2.3 Sturgis's Five Elements of Competency-Based Education

Element (Sturgis's language)	Summary
1. Students advance upon demonstrated mastery.	Mastery before movement
2. Competencies include explicit, measurable, transferable learning objectives that empower students.	Clear communication of learning expectations
3. Assessment is meaningful and a positive learning experience for students.	Use of formative assessment
4. Students receive timely, differentiated support based on their individual learning needs.	Differentiated instruction
5. Learning outcomes emphasize competencies that include the application and creation of knowledge, along with the development of important skills and dispositions.	Inclusion of HOTS and habits of work

Note. Retrieved from (Sturgis, 2016, p. 11).

Again, HOTS and habits of work are mentioned as important components of a competency- or proficiency-based system of education. Researchers at CompetencyWorks have highlighted Maine as a leader in the field of PBE. They have listed Maine as a state that is advanced in their integration of PBE. Additionally, CompetencyWorks shares resources distributed by the Maine Department of Education (Competency Works, 2012). They believe that Maine is at the forefront of PBE integration due to cohort support and Department of Education and legislative requirements (Sturgis, 2016).

Summarizing the Components of PBE

In reviewing various approaches to PBE, common characteristics can be seen across frameworks. The following components emerged in several definitions:

1. Clear learning expectations (GSP, 2015; MCCL, 2015; Sturgis, 2016)
2. Mastery drives movement (MDOE, 2015; GSP, 2015; MCCL, 2015; Sturgis, 2016)
3. Formative assessments (GSP, 2015; MCCL, 2015; Sturgis, 2016)
4. Student voice (GSP, 2015; MCCL, 2015)
5. Differentiation of instruction (GSP, 2015; MCCL, 2015; Sturgis, 2016)

Local and National PBE Research

In 2014, researchers from the University of Southern Maine concluded the first phases of an important mixed-methods study consisting of survey and case studies that examined the implementation of PBE in Maine. Silvernail et al. (2014) found that schools and districts encountered the following challenges in PBE implementation: common definitions, insufficient time, stakeholder support, and an effective management system. Benefits of the new system included increased student engagement, greater levels of collaboration, and more focused interventions. Their conceptual model of a working proficiency-based system emphasized the importance of an effective learning management system and strong community and family support (Stump & Silvernail, 2014).

During the second phase of their research, Silvernail et al. (2014) examined PBE implementation at the district level. They studied eight school districts from across Maine at different stages of implementation. Again, they summarized the benefits and challenges of implementation. The former were very similar to those in the first phase of the research (including student engagement), which the researchers said was a direct result of clear standards and expectations. An additional benefit mentioned was an increase in transparency regarding the tracking of student progress.

During the 2016 phase of the research, six schools that had participated in all previous phases were examined as case studies. A culture of learning and strong professional collaboration were believed to be necessary for a PBE system to be successful. The schools' areas of focus included habits of work, interventions for students who did not meet standards, and public relations (Stump et al., 2016).

In May 2018, a survey was sent to 1,500 educators in K-12 schools throughout Maine. It asked their perceptions on the implementation and impact of proficiency-based diploma requirements. The results clearly indicated that implementation was slow and inconsistent. Only 32% of middle school educators and 42% of high school educators stated that systems were fully in place to track student proficiency. Moreover, 69% of middle school educators and 65% of high school educators reported that the standards that students were expected to meet had been identified. Student movement at an individual pace was slow to be adopted. Only 24% of middle teachers and 29% of high school teachers said that students progress based on proficiency attainment (Johnson & Stump, 2018).

In the same study, educators were asked about their perceptions on the impact of proficiency-based graduation requirement on academic rigor. Only 23% of educators agreed or strongly agreed with the following statement: "Proficiency-based graduation requirements increase academic rigor." In fact, only 6% of respondents within this group strongly agreed with the statement. The researchers believed that the many inconsistencies in the definition of proficiency led to challenges in maintaining a certain level of rigor (Johnson & Stump, 2018).

Once the decision was made to pull back Maine's proficiency-based diploma requirements, superintendents in the state were also surveyed. Despite some participants' decision to opt out of proficiency-based requirements, they reported that they would still implement many previously adopted practices. By contrast, superintendents that decided to opt into the requirements said that they would selectively discontinue some practices. Many respondents cited decreased rigor as a concern that led to the discontinuation of some practices in their district (Johnson, 2019).

Inconsistency seemed to be a major disadvantage in the implementation of PBE in Maine schools. It manifested not only from district to district or school to school but also within schools themselves. When researchers examined level of exposure to proficiency-based practices in 10 districts in Maine, they observed the same patterns in each district: (1) minimal, (2) low-medium, and (3) medium. They found no evidence of a high level of implementation from classroom to classroom at any school (Shakman et al., 2018).

At the national level, the PBE model has garnered mixed reviews. According to a national report, five states were advanced in their implementation of PBE (Brodersen et al., 2016). However, 10 states actively engaged in capacity building for CBE in their school systems (Sturgis, 2016). For example, Vermont added a flexible pathways initiative to its Education Quality Standards in 2013 (Vermont Agency of Education, 2015). The initiative calls for students to only advance once they demonstrate attainment of skills. In addition, New Hampshire passed legislation that stipulated that all schools should to be competency-based by 2017, and Connecticut passed mastery-based learning guidelines in 2015 (Sturgis, 2016). In Oregon, a house bill introduced a proficiency-based structure, only to be removed by policymakers through an emergency clause midway through the school year. While it is still has many supporters at the state level, it is left up to schools to decide how they will proceed (Hammond, 2014).

Throughout the Northeastern United States, principals at schools that have implemented a CBE or PBE structure explained that there are four major barriers to implementation: stakeholder buy-in, collaboration time for teachers, changing existing structures and policies, and professional development for teachers (Evans & DeMitchell, 2018). This list was later expanded in a literature review of CBE and PBE that focused on factors that can support or impede implementation. Obstacles included different definitions of proficiency and interpretations of standards. In addition, tensions between competency-based grading and meeting grade-level expectations were noted, while transparency and consistency were identified as facilitators (Evans et al., 2020).

Regarding the student experience in a proficiency-based mathematics class, research has shown that anxiety level seems to decrease for each individual assessment. However, this changes as students approach the end of the grading period. This is expected, as test retakes lower the need for students to attain proficiency on their first attempt. If assessments are not completed after multiple attempts, the danger of falling behind increases students' anxiety levels (Lewis, 2020).

It would appear that PBE or CBE is more than an educational trend. Although most of the initial policy implementation has been in New England, schools throughout the country have implemented PBE systems (Sturgis, 2016). Despite a lack of fidelity in implementation, the research to date suggests that this has not had an impact on students' transition from high school to postsecondary education (Guskey et al., 2020).

HOTS in PBE Frameworks

While higher-order thinking or complex reasoning are components that are included in several proficiency-based frameworks (GSP, 2015; MCCL, 2015; Sturgis, 2016), HOTS tasks in the classroom may not be implemented in a consistent manner. This could be due to several factors, such as lack of clarity of resources or the diversion of educators' focus to implementing other components of the PBE system.

A HOTS framework that is currently being used by PBE schools in Maine is Marzano's complex reasoning taxonomy (Marzano & Pickering, 1993; Marzano & Kendall, 2007), which is shown in Figure 2.3. Schools that are part of the MCCL use the terms "retrieval," “comprehension,” “analysis,” and “utilization,” which are derived from headings used in the taxonomy to determine the “rigor” of learning tasks. Examples of rubrics from an MCCL mathematics standard are shown in Figure 2.4. The rigor levels represented in these rubrics correspond to the four heading of the complex reasoning taxonomy illustrated in Figure 2.3.

Figure 2.3 Marzano’s Complex Reasoning Taxonomy

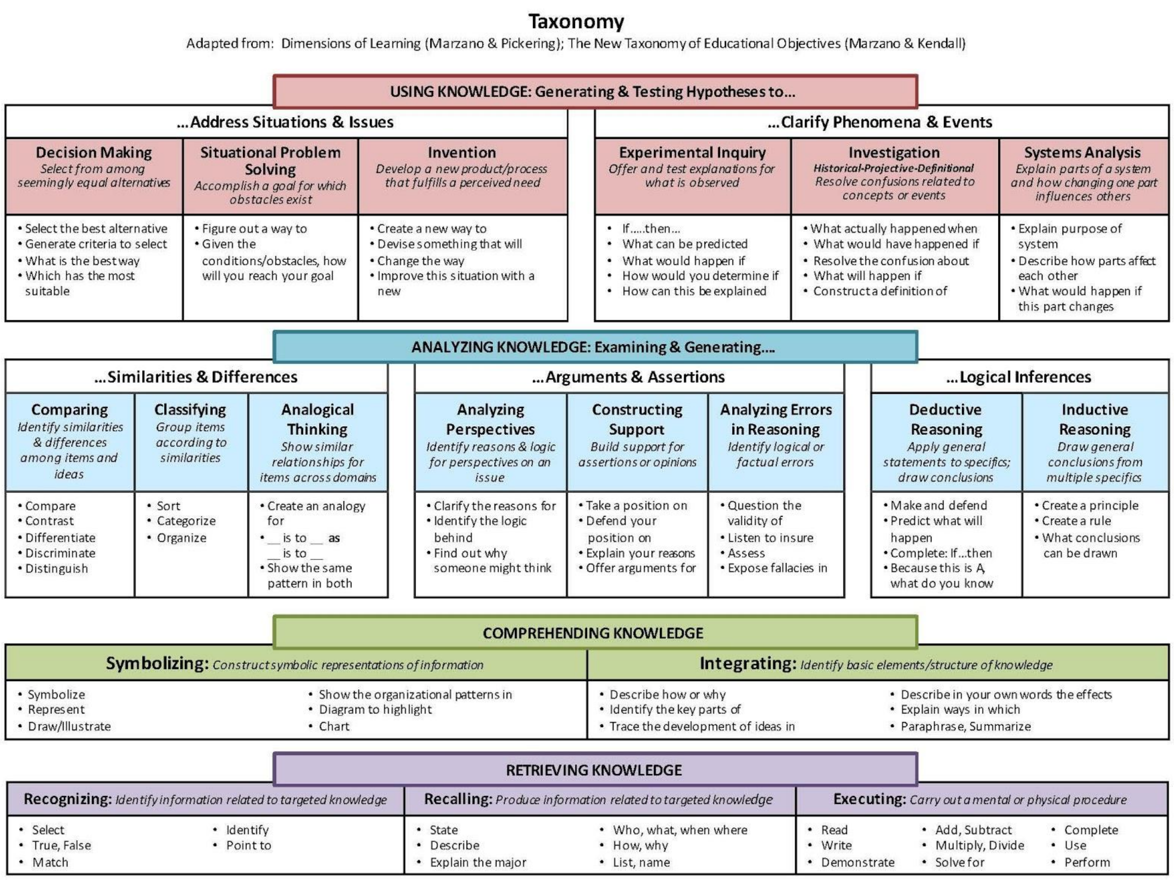
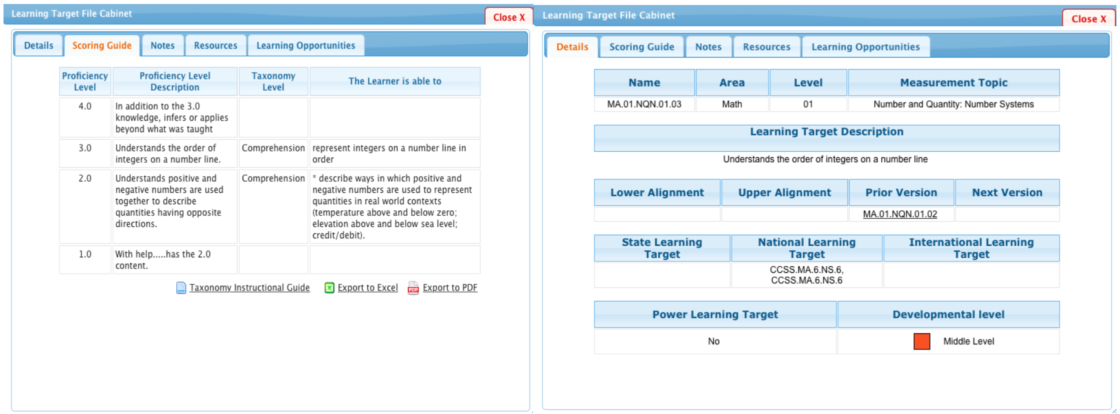


Figure 2.4 Example of an MCCL Mathematics Standard



Note. Retrieved from http://mainecustomizedlearning.org/resources/

While a system is in place to facilitate the integration of HOTS tasks in the classroom, most rubrics do not explicitly explain what constitutes a higher-order task for a given standard. For instance, the example provided in Figure 2.4 tells students how to attain a 2 (partial proficiency) and a 3 (proficiency). However, to attain a 4 (exceeding proficiency), it simply states that they must apply "beyond what was taught." In addition, proficiency in this standard corresponds to the level of comprehension in Marzano's taxonomy. Educators must design their own tasks for students to use HOTS, and only the students who have received a score of 3 can move on to that task if it is available. Thus, this approach to HOTS puts pressure on the teacher to develop resources for students who have achieved proficiency, and it is often only proficient students who have opportunities to engage in higher-order tasks.

Conversely, GSP includes HOTS as a subcomponent of what it refers to as the "cross-curricular graduation standards." In this case, HOTS are grouped with habits of work. The organization recommends the teaching of these standards in all classes from Grades 5 to 12. Upon graduation, students demonstrate their mastery of the standards by presenting a body of evidence, which can include portfolios, exhibitions, and capstone projects (GSP, 2015).

While research has shown that schools are beginning to focus on habits of work at an early age (Stump et al., 2016), it remains unclear whether HOTS is an area of focus for schools that implement PBE, particularly at the elementary school and middle school levels, where there is currently no system of accountability statewide. Despite the challenges associated with the implementation of HOTS in a PBE framework, it is clear that it can be accomplished. For example, New Haven Academy in New Haven, Connecticut is a magnet school for Grades 9–12. Its mission is to teach the value of depth of understanding and critical thinking to students. When the school developed a new grading system, it used formative assessments and core assessments, the latter of which were connected to higher-order tasks required for college or the workforce (Sturgis, 2016).

Importance of Higher-Order Thinking Skills

Higher-order thinking skills play an important role in better understanding the world (Schraw & Robinson, 2011), which community leaders have not overlooked. Recently, concerns that schools are not adequately preparing students for life after high school have been raised (ACT Inc, 2012; The Education Trust-West, 2011; Wagner, 2008). Research has outlined seven areas in which students need more practice and preparation to successfully pursue a career: critical thinking and problem solving, collaboration and leadership, agility and adaptability, initiative and entrepreneurialism, effective oral and written communication, accessing and analyzing information, and curiosity and imagination (Wagner, 2008). The link between many of these items and higher-order tasks is apparent. In fact, research has shown that the ability to apply critical thinking and problem solving are among the most important skills for pursuing a new career (Hess & Gong, 2014; Larson et al., 1998).

The challenge of regularly integrating higher-order thinking in the classroom is not new. However, PBE offers a new structure for the delivery of instruction and assessments. Therefore, it is imperative for educators to reflect on how these activities fit into their instructional planning. Community leaders have underscored the fact that skills connected to higher-order thinking are crucial for career success, which raises the question of whether all schools that are implementing a PBE system are prioritizing these skills accordingly.

In a study that used Webb's depth of knowledge framework to determine the rigor level of student work in the classroom, Paige et al. (2013) found that higher levels of cognitive rigor indicated a greater level of student engagement. The authors expanded on Hess's work and used the term "cognitive rigor" to describe the extent to which teachers ask students to use their critical thinking skills. Another notable finding from the study was that, as students became less engaged as they approached the end of a class. Moreover, students who worked on activities at a lower depth of knowledge tended to demonstrate a greater lack of engagement towards the end of a class. In addition, students were more willing to engage in mathematical sensemaking if their own ideas were made public within the classroom. This was also true if their ideas were extended upon by their teacher or classmates and if there was encouragement for explanations or justifications (Mueller et al., 2014).

Using HOTS in the classroom not only impacts students and teachers' perceptions of engagement, but it also has a visible effect on participation rates. In a study conducted in sixth and seventh grade classrooms, teachers' questioning techniques were analyzed. As teachers pushed students to a greater level of depth with their question choices, student participation increased (Matsumura et al., 2008) in terms of both the frequency and quality of participation. Researchers have surmised that overall classroom quality has an impact on student learning

(Miller, 2003; Matsumura et al., 2008; Yi & Lee, 2016). One of the key factors that determines classroom quality is the level of rigor demonstrated in the teacher's questions (Brookhart, 2016; Miller, 2003; Matsumura et al., 2008).

When middle school students were taught with the Connected Math curriculum, a standards-based and problem-centered approach, they demonstrated great gains in performance on a state basic skills test. In addition, teachers shared a belief that the curriculum fostered better problem-solving abilities among students (Cain, 2002). This is an example of a rigorous curriculum tied to real-world problem solving that impacts how students approach complex problems beyond the scope of the curriculum itself.

A similar finding was identified in the landmark Quantitative Understanding: Amplifying Student Achievement and Reasoning (QUASAR) study. This research project focused on the urban seventh grade student population and was based on the idea that urban students lack high-quality mathematics instruction, not ability. Students were asked to frame and interpret mathematics problems and regulate their own thinking processes; they also were asked to justify their problem-solving strategies and conclusions. As a result of the change in instructional practices, QUASAR students performed considerably better than their peers on the National Assessment of Educational Progress, and this difference was even greater when they were asked to justify their answers. In addition, participation in the QUASAR program led to an increase in the number of students who took freshman algebra classes (Silver & Stein, 1996).

Thus, the inclusion of HOTS in regular classroom activities can lead to greater student engagement in activities and lessons and higher participation rates (Paige et al., 2013; Matsumura et al., 2008; Mueller et al., 2014). Students also improved as real-world problem solvers and demonstrated greater academic success in general, as evidenced by standardized test scores (Cain, 2002; Silver & Stein, 1998). Additionally, HOTS are exactly the type of skills sought by employers (Hess & Gong, 2014; Larson et al., 1998). Therefore, it is imperative for activities that require the use of HOTS to be included in all classrooms.

Higher-Order Thinking Frameworks

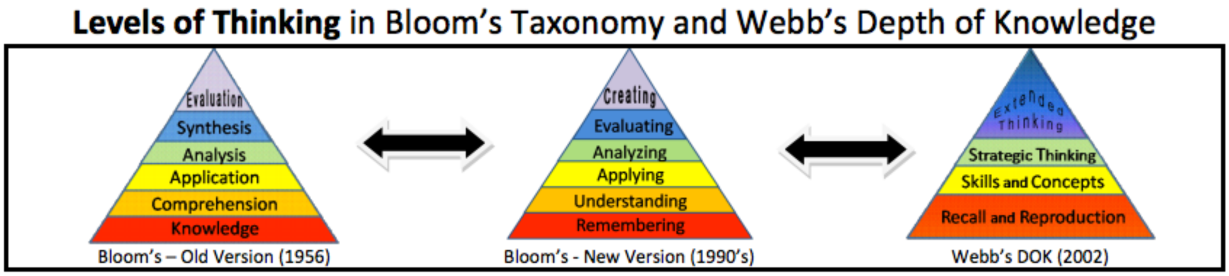
Over the years, educators who are interested in incorporating higher-order thinking tasks in the classroom have familiarized themselves with various frameworks or taxonomies that can help them improve the cognitive level of student work. Starting with Bloom's framework, these models have organized the cognitive demand associated with tasks from the recall level to the knowledge utilization or extended thinking level. Many of these frameworks are applicable across multiple content areas and grade levels, while others are more specific to content areas. A common theme is the transition from the mere retrieval or recall of information to activities that require the creation or synthesis of information. This review covers Bloom’s taxonomy, the new Bloom’s taxonomy, Webb’s depth of knowledge, Marzano’s complex reasoning taxonomy, Hess’s cognitive rigor matrix, and Daggett’s rigor/relevance framework.

In education, perhaps the best-known categorization of levels of reasoning required for specific classroom activities is Bloom’s taxonomy, which Bloom introduced in 1956. In fact, the cognitive taxonomy was only one of the three educational domains that Bloom addressed; the others were the affective and psychomotor domains. However, educators are most likely referring to the cognitive domain when citing Bloom’s taxonomy. The cognitive taxonomy consists of six levels: knowledge (consisting mainly of recall), comprehension (tied to the understanding of meaning), application (the use of the learned material), analysis (breaking up the material into smaller parts to understand the whole), synthesis (using individual components to develop a new whole), and evaluation (developing a conclusion based on judgment of material; Bloom, 1956).

In 2002, Krathwohl published a revision of Bloom's taxonomy. His first major change was to separate the knowledge dimension from the cognitive taxonomy. He also defined the types of knowledge that can be learned, which are factual, conceptual, procedural, and metacognitive knowledge. Krathwohl's cognitive process dimension is often referred to as the new Bloom's taxonomy; it is very similar to Bloom's taxonomy. The main differences are that knowledge was switched to "remember" and that comprehension was renamed "understanding." Perhaps the most significant change was that synthesis was renamed "creation," and its position was exchanged with evaluation on the final rung of the taxonomy (Krathwohl, 2002). These changes are illustrated in Figure 2.5.

Around the same time as the new version of Bloom's taxonomy was introduced, Webb (2007) presented the depth of knowledge (DOK) scale to aid in his research on curriculum and assessment alignment. This DOK scale consists of four levels. The first is recall and reproduction (directly recalling facts), the second is skill and concept (typically involving multiple steps), the third is strategic thinking (requires the development of a plan or sequence of steps and may have more than one answer), and the final level is extended thinking (requires longer periods of time, planning, synthesizing, and evaluating). Webb's DOK has been used to assess the rigor levels of questions on multiple standardized assessments, most recently the Smarter Balanced Assessment. Often, the levels of thinking addressed in Webb's DOK are cross-walked to Bloom's taxonomy (see Figure 2.5 for an example).

Figure 2.5 Comparison of Bloom's Taxonomy and Webb's DOK



Note. Retrieved from http://www.doe.in.gov/sites/default/files/turnaround-principles/bloom-vs-webb-chart.pdf

Another complex reasoning scale is the complex thinking standards introduced by Marzano, Pickering, and McTighe in 1993. Again, many comparisons can be made to Bloom's and Webb's DOK. Knowledge is grouped at four levels: retrieval, comprehension, analyzing, and using. When educators use this rubric in the classroom, each of the levels can be broken down into sublevels. For instance, retrieving consists of recognizing, recalling, and executing. Proficiency-based schools that implement this scale have connected proficiency to one of the four knowledge levels (MCCL, 2015).

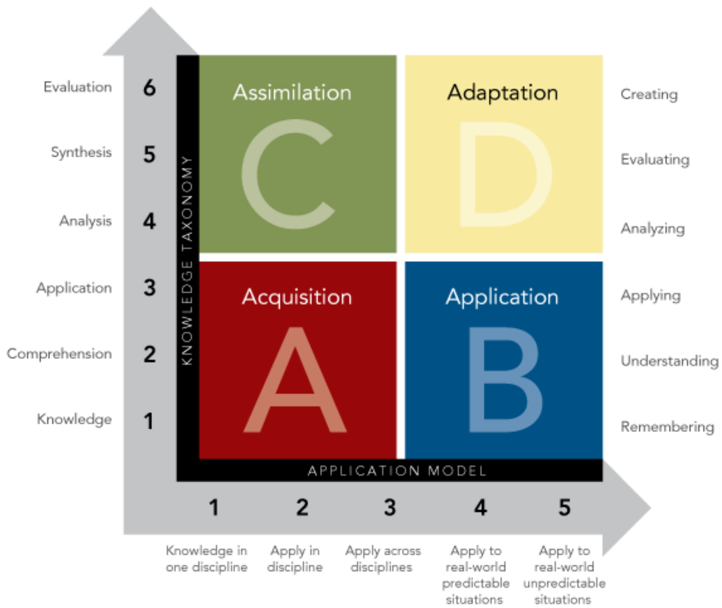
In 2009, Hess and colleagues introduced the term "cognitive rigor" when they developed a matrix that enables educators to characterize rigor using both Bloom's updated taxonomy and Webb's DOK. The matrix lists the six taxonomic levels introduced by Krathwohl on the vertical axis and Webb's four DOK levels on the horizontal axis. Evaluators can then choose the cell that reflects the level of rigor observed in the classroom at the time of evaluation. An interesting finding from Hess et al.'s research was that mathematics assignments tended to rank on the lower end of the matrix compared to English Language Arts assignments, which the researchers believed could have implications for future research. They stated that the desired class would have results that ranged across the matrix but were not focused in one area. A summary of the abovementioned frameworks can be found in Table 2.4.

Table 2.4 Summary of HOTS Frameworks

HOTS framework	HOTS levels
Bloom's taxonomy (1956)	1.Knowledge 2. Comprehension 3. Application 4. Analysis 5. Synthesis 6. Evaluation
New Bloom's taxonomy by Krathwohl (2002)	1. Remembering 2. Understanding 3. Applying 4. Analyzing 5. Evaluating 6. Creating
Webb's depth of knowledge (2007)	1. Recall and reproduction 2. Skills and concepts 3. Strategic thinking 4. Extended thinking
Marzano's complex reasoning standards (1993)	1. Retrieval 2. Comprehension 3. Analyzing 4. Using
Hess's cognitive rigor matrix (2009)	Six levels from new Bloom's taxonomy and four levels from Webb's DOK
Daggett's rigor/relevance framework (2017)	A. Acquisition B. Application C. Assimilation D. Adaptation

Through his work at the International Center for Leadership in Education, Daggett developed the rigor/relevance framework to examine curriculum instruction and assessment. The framework consists of a thinking continuum on the y-axis and an action continuum on the x-axis (see Figure 2.6). The goal is for educators to create opportunities for students to reach quadrant D, also known as the adaptation quadrant. In this quadrant, students use higher-order thinking strategies and apply them in a real-world context (Daggett, 2017).

Figure 2.6 Daggett's Rigor/Relevance Framework



Note. Retrieved from http://www.leadered.com/pdf/Rigor%20Relevance%20Framework%20White%20Paper%202016.pdf

Increasing HOTS or Rigor in the Classroom

The term "rigor" has a specific meaning in educational circles. A task's level of rigor is often synonymous with the level of higher-order thinking required to accomplish it. According to Paige, Smith, and Sizemore (2015), rigor is on a scale of cognitive complexity. They emphasized that rigor relates to critical thinking, problem solving, and helping students understand complex content.

Research has placed considerable emphasis on the teacher to increase the level of cognitive rigor in the classroom. In her book *Rigor is NOT a Four-Letter Word*, Blackburn (2013) outlined five influences: raising the level of content, increasing complexity, providing appropriate support and guidance, opening the focus, and raising expectations. Blackburn's message is targeted towards teachers and focuses on what they can do within their own classrooms to raise rigor levels. Blackburn advocates for a greater focus on depth rather than breadth of knowledge. She asks teachers to pay close attention to students' LEXILE level of and allow them "limited choice" within their own LEXILE range. Blackburn is also a proponent of what she calls "problem-based learning" rather than project-based learning. This type of learning is described as an "inquiry process" and requires students to find a resolution.

Clearly, a more in-depth approach targeted at individual needs seems to be a great start. However, students currently may not have the skills or ability necessary to complete rigorous tasks. Thus, schools may need to shift their focus to skills rather than tasks. Recently, greater emphasis has been placed on the introduction of thinking skills programs at schools. They focus on areas such as habits of mind and offer instruction on how to think when confronted with tasks that require skillful thinking (Costa & Kallick, 2015; Akınoğlu & Karsantık, 2016).

Nevertheless, there seems to be considerable confusion about what rigor entails or even what a rigorous task looks like. As a result, there are certain myths and misconceptions about rigor in the classroom, which are often perpetuated by teachers (Blackburn, 2013; Moon, 2009; Wilen, 2004). The first misconception is that rigor levels will increase if more work is assigned. This is true for both homework assignments and classwork in general (Blackburn, 2013). Another misconception is that standards must be lowered for all students to ensure success for all students. However, this strategy disregards the importance of ensuring that each student's needs are met within the classroom (Torff, 2014; Burris et al., 2008; Blackburn, 2013). In addition, researchers have made reference to the folk belief among educators that learning is a hierarchical process and that students cannot complete a higher-order task until they have demonstrated success on basic tasks (Torff, 2014). Finally, educators often believe that providing support to students means lowering levels of rigor. This is not the case, as students receive support to enable them to learn at higher levels (Burris et al., 2008; Blackburn, 2013).

Recent research indicates that school-chosen textbooks have a major influence on teachers' perceptions of higher-order thinking in the classroom. The curriculum influences teachers' curricular aims and objectives and their understanding of students' readiness to participate in problem solving in the classroom. Thus, it is crucial for districts to consider these factors when they choose resources for their mathematics curriculum (Davis et al., 2019).

Although teachers can certainly take measures to increase rigor in their own classrooms, the evidence shows that schools can significantly improve student learning if they take thinking skills seriously and design thinking skills programs as a component of instruction. Ritchhart et al. (2011) refers to this approach as the creation of a culture of critical thinking. His research showed that schools must prioritize this type of thinking and have students regularly reflect on their own thinking. Additionally, Ritchhart advocates for critical thinking routines as a regular part of classroom instruction.

HOTS in the Mathematics Classroom

On the surface, mathematics seems like a logical fit when integrating HOTS into classroom instruction. However, mathematics teachers tend to lecture more than teachers in other subject areas, which provides fewer opportunities for the integration of activities that require the use of HOTS (Marshall & Horton, 2011). In addition, mathematics teachers have had difficulty creating assessment items at each level of Bloom's taxonomy when asked to do so. This is true even when the teachers were relatively familiar with Bloom's taxonomy. More success has been observed when teachers use much simpler thinking skills frameworks (Thompson, 2008). These mathematics-specific frameworks also have a similar progression from retrieval to utilization. However, a slight difference is that mathematics focus first on understanding and practicing the procedure (or algorithm) before moving on to more comprehension-based activities, such as proofs or real-world connection problems.

HOTS in Middle and Secondary School Classrooms

At the national level, it is recommended that all mathematics students have the opportunity to achieve high levels of mathematical learning (NCTM, 2016). This is underscored by the language of the Common Core State Standards for mathematics. In both the middle and high school standards, students are asked to describe relationships and compare, interpret, and construct (Math Standards, 2016). In addition, the next-generation assessments currently being used by states were found to be good to excellent at assessing the cognitive demand required in the Common Core Standards (Doorey & Polikoff, 2016).

Regarding higher-order thinking in middle school, the Association for Middle Level Education (AMLE) advocates for introductory lessons to include higher-order thinking principles. Educators are encouraged to incorporate practices such as asking students to develop a rule, present a paradox to their peers, and incorporate problem solving (Kelleher, 2016). These strategies are recommended for educators at both the middle and secondary school levels. In addition, research demonstrates that inclusion in higher-order thinking tasks is beneficial for high and low achievers alike (Zohar & Dori, 2003; Torff, 2008; Torff, 2014; Silver & Stein, 1996).

The move to Common Core mathematics standards appears to have had an impact on middle school mathematics education at the national level. In a recent survey, six to eight teachers reported that they paid less attention to the memorization of basic mathematics facts and now focused more on application. Moreover, students were being taught multiple ways to solve a problem. However, the respondents reported that changing educational practices had created stress for students (Bay-Williams et al., 2016).

At the high school level, Advanced Placement (AP) courses enable students to examine a content area in greater depth using higher-order thinking strategies. Describing their AP Calculus (AB) course, the College Board emphasized that students would learn how to solve problems, experiment, interpret results, and support conclusions. Such learning would fall on the high end of Bloom's taxonomy (College Board, 2016). While enrollment in an AP Calculus course increases opportunities for higher-order thinking, students may still attempt to use memorization strategies. Research demonstrates that the students who succeed at AP Calculus courses tend to use visual learning strategies (Haciomeroglu & Chicken, 2012). This indicates that a greater conceptual understanding of the content exists and that higher-order thinking is taking place.

Overall, more high school students across the United States can access AP courses than in the past. In 2017, a total of 2,741,426 students registered for AP exams, which represents more than 1.2 million more students than in 2007. In addition, the mean score on the AP exams in 2017 and 2007 were similar (2.86 to 2.89, respectively). Thus, the rate of success in AP classes has not decreased with the influx of AP students (College Board, 2017). In 2017, there were approximately 15.1 million students in the United States, 7.5 million of whom were juniors and seniors. Excluding students who took multiple AP exams, approximately 35% of the junior and senior population were engaged in courses that required them to regularly use HOTS. With

AP course overlap, the number could be safely dropped to 25% or so. Therefore, it is important to understand how frequently the remaining 75% of students, along with freshmen and sophomores, are being afforded opportunities to use HOTS in the classroom.

Statistics show that more high school students are participating in programs such as AP courses. It is also clear that Common Core State Standards require educators to include more instructional practices that integrate higher-order thinking strategies. In addition, these practices are supported by national organizations such as the NCTM and AMLE. This raises the question of whether there has been an overall increase in higher-order thinking in classrooms in recent years. In 2005, Valentine shared the typical percentages of instruction that falls into each category of the IPI, a tool that he created to assess levels of thinking associated with classroom tasks based on walkthrough observations. He found that, in both middle and secondary school, students were engaged in higher-order tasks 18–25% of the time. Teacher-led instruction was 35–45% at the middle school level, whereas it was 30–40% at the high school level. Another notable finding was that there was an approximately 15% increase in higher-order tasks in noncore courses compared to the four major core content areas. In addition, more effective schools were observed to have 10% more higher-order thinking than less effective schools (Valentine, 2005).

Classifying HOTS Using Mathematics-Specific Frameworks

In 1976, Skemp published an article that introduced two forms of understanding in mathematics, which he named relational and instrumental understanding. He described instrumental understanding in mathematics as a process in which students are taught to memorize and repeat a procedure. Relational understanding, on the other hand, focuses more on the big picture and the role played by the current topic in this context. Skemp went on to describe four advantages of relational mathematics: it is more adaptable, easier to remember, can be a goal in itself, and has an organic quality. Thus, he advocated for teachers to approach mathematics with instructional methods that are more tied to relational than instrumental understanding. However, Skemp recognized that this can be difficult for teachers: "within its own context, instrumental mathematics is usually easier to understand; sometimes much easier" (p. 22).

In a 2002 work about measuring the content of instruction, Porter introduced a content matrix that enables educators to categorize the cognitive difficulty of various mathematical topics. He listed the following categories: memorize, perform procedures, communicate understanding, solve nonroutine problems, and conjecture/generalize/prove. Porter recognized the challenges associated with teacher self-reporting on levels of activities that took place in their classrooms using his system. He recommended professional development in this area.

Doyle (1982) divided mathematical tasks into two categories: procedural tasks and comprehension tasks. Procedural tasks typically require students to follow an algorithm or a step-by-step process. Comprehension tasks require students to first develop a cognitive representation of the ideas, which is a much more challenging process.

Smith and Stein (1998) developed four types of tasks that require varying levels of cognitive demand in mathematics classrooms: ones that require memorization, procedures without connection to concept or meaning, procedures that make connections to concepts of meaning, and doing mathematics. Tasks in the final category require complex thinking and "considerable cognitive effort."

The process of understanding and making the mathematical task sensible can be referred to as sensemaking (Fitzgerald & Palincsar, 2019). Kilpatrick (2001) introduced five strands in the process of mathematical sensemaking: (1) conceptual understanding, (2) procedural fluency, (3) strategic competence, (4) adaptive reasoning, and (5) productive disposition. It is important to note that Kilpatrick emphasized the need for a productive disposition in the process of making sense of mathematics, which includes students feeling that the subject is worthwhile and useful. A summary of the abovementioned mathematics-specific frameworks can be found in Table 2.5.

Table 2.5 Overview of Mathematics-Specific HOTS Frameworks

HOTS framework	HOTS levels
Skemp's framework for understanding (1976)	1. Instrumental understanding 2. Relational understanding
Porter's cognitive demand matrix (2002)	1. Memorize 2. Perform procedures 3. Communicate understanding 4. Solve nonroutine problems 5. Conjecture/generalize/prove
Doyle's academic tasks (1982)	1. Procedural tasks 2. Comprehension tasks
Smith and Stein (1998)	1. Memorization 2. Procedures without connection to concepts and meaning 3. Procedures with connection to concepts and meaning 4. Doing mathematics
Kilpatrick's sensemaking strands (2001)	1. Conceptual understanding 2. Procedural fluency 3. Strategic competence 4. Adaptive reasoning 5. Productive disposition

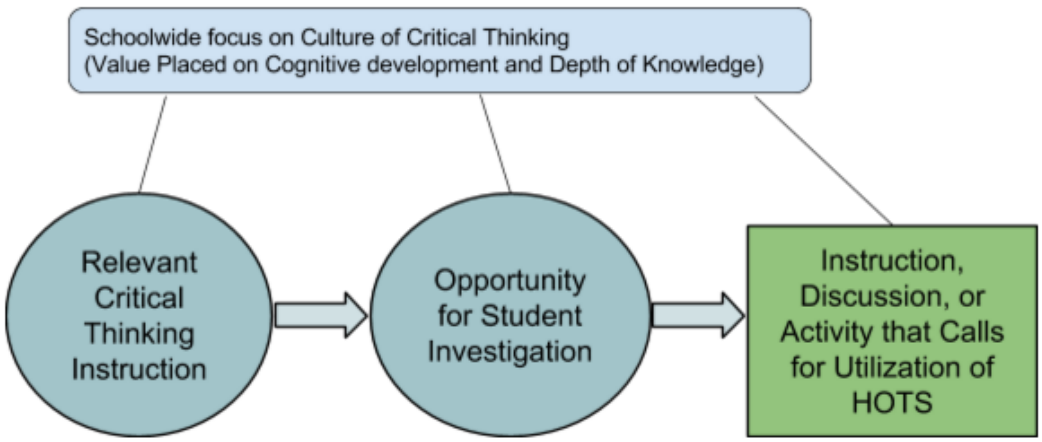
A common misconception in the mathematics classroom is that completing a higher-level course means achieving a deeper understanding of the material. In fact, the opposite seems to be true. Often, when students take higher-level courses in mathematics or pursue more challenging material, they do not gain a greater understanding of the content. Students who have taken advanced courses in high school, such as calculus, have not necessarily performed better on standardized tests (Schneider, 2009). This is a paradox that Schneider (2009) referred to as “the delusion of rigor” (p. 2). It is perhaps this innate understanding that has led some mathematics educators to examine their curriculum from a “less is more” approach. In this approach, the curriculum contains fewer targets, but all students are given the opportunity to tackle them on a deeper level. In some cases, this has led to considerable student success (Gine & Kruse, 2007; LeVenia et al., 2010).

So how do mathematics teachers begin to integrate higher-order thinking concepts into their classrooms? Research has found that student exploration is a key component and enables students to develop a deeper understanding. When teachers allow students to explore concepts prior to an explanation (either by a teacher or a student), they think more deeply about the content. In addition, when more exploration time was given to students, their resulting work was found to correspond to greater cognitive levels on Bloom’s taxonomy (Marshall & Horton, 2011).

Along with more exploration time, taking the time to specifically instruct students on the problem-solving process has had a positive effect on students’ ability to perform challenging cognitive tasks (Krawec et al., 2012). When students are provided with steps or a procedure that explains how to use critical thinking in mathematics, they are also much more likely to become invested in the process and able to write about their thinking (Kjos & Long, 1994).

Thus, research shows that, prior to formal instruction or work that requires students to use HOTS in the mathematics classroom, teachers should provide clear and relevant critical thinking skills instruction, then allow students to investigate the topic on their own. This progression is depicted in Figure 2.7.

Figure 2.7 Steps for the Successful Integration of Higher-Order Thinking



Integrating HOTS in the Mathematics Classroom

Figure 2.7 illustrates that a school-wide culture that supports the use of critical thinking must be present to ensure the success of higher-order thinking tasks. This means that students should be familiar with critical thinking routines and have time to reflect on their thinking (Ritchhart et al., 2011). In addition, explicit instruction in critical thinking or higher-order thinking strategies, along with built-in exploration time for students, are essential for completing higher-order tasks in the mathematics classroom. Research has demonstrated that these components lead to higher-performing students who are more likely to complete complex tasks (Krawec et al., 2012; Marshall & Horton, 2011).

When reviewing the QUASAR research with urban seventh-grade students, Silver and Smith (1996) emphasized the importance of building an atmosphere of trust and mutual respect in the classroom prior to engaging in HOTS tasks. Classroom norms should be well-established, and students should feel free to question each other respectfully. In a successful classroom, a comfort of community is built first, then the level of mathematics is increased. Before undertaking a HOTS task, it is also crucial for students understand the work that they are being asked to complete to make substantive progress towards a solution.

Integrating HOTS in Middle and Secondary School Classrooms

Recently, considerable research has been conducted in Indonesia on the integration of higher-order thinking tasks at the middle and secondary school levels. The Indonesian government emphasized the inclusion of mathematical HOTS in education to set apart their learners from others in the midst of the Industrial Revolution 4.0 (Tambunan, 2019).

An approach to mathematical problem solving examined by Indonesian researchers divides the process into four stages. The first step is to understand the problem. The second is to plan a solution. The third is to complete the model, and the fourth is to check the problem again (Polya, 2004). This approach to problem solving, which creates a heuristic (or more practical method), was found to improve students' ability to understand concepts and reason, create, and communicate. Additionally, it led to an increase in students' level of learning independence (Tambunan, 2018; Tambunan, 2019).

In terms of opportunities for teachers to incorporate HOTS in the mathematics classroom, combining blended learning strategies with projects can make a greater impact on student HOTS abilities. Blended learning is a mix of classroom-based learning practices and online tools or resources. Through a collaborative model that incorporates technology and the completion of higher-order tasks, students have a better chance to develop their abilities (Eliyasni et al., 2019).

Despite the use of terms associated with Bloom's taxonomy (e.g., analyze or evaluate), some confusion remains about what constitutes a higher-order task in the classroom. Students or educators may believe that higher-order work is taking place when it is simply memorization of an algorithm. Thus, it is important for educators to pay close attention to the problem-solving approaches or strategies used by students when seeking opportunities to incorporate HOTS in the classroom (Tanujaya & Mumu, 2020).

At the middle school level, students begin to familiarize themselves with critical thinking. Critical thinking can be divided into six subcategories: (1) interpretation, (2) analysis, (3) inference, (4) evaluation, (5) explanation, and (6) self-regulation. Overall, middle school students score low on these subcategories, particularly evaluation, analysis, and self-regulation. Therefore, it is crucial to teach middle school students specific critical thinking routines and strategies rather than assume that they will be ready to use them (Basari et al., 2019).

When examining HOTS among high school mathematics students, challenges with transformation and process skills tended to be most prevalent. Transformation refers to the process of reviewing information and developing an appropriate mathematical model to better understand the problem. Process skills are the fundamental skills that students should possess at their level of mathematics education. However, they most frequently impeded the completion of higher-order tasks. As a result, the lack of these foundational skills may give teachers pause when they must decide whether to incorporate higher-order tasks in their classroom instruction (Hadi et al., 2018).

Concerns Related to HOTS in a Proficiency-Based Classroom

Proficiency-based classrooms focus on students' attainment of proficiency in all grade level standards. Schools that follow a student-driven pace create a heterogeneous environment in which students focus on different material at the same time. To instruct students where they are, teachers use ability groupings. However, research on mathematics classrooms has shown that ability groupings have not produced a noticeable increase in student achievement (Linchevski &

Kutscher, 1998). The achievement gap between seventh grade high end students in homogeneous groups and heterogeneous groups was nearly nonexistent. In addition, average and lower-achieving students performed significantly better in heterogeneous groupings. Thus, a structure that favors ability groups could impact the achievement of these students.

A significant concern for teachers in proficiency-based classrooms is the risk of focusing on minimum competencies connected to the standards to the detriment of developing a plan for students to delve deeper into the content (Schraw & Robinson, 2011). For example, an algebra teacher may focus on teaching students how to graph a line in $y = mx + b$ form and stop there. When students can demonstrate this task, they are considered to have attained proficiency. By contrast, a true depth of knowledge task could demonstrate to students why linear equations are important through data collection and the modeling of a real-world situation. Another approach could be to examine the impact of various transformations on the equation itself. Such tasks require students to develop and analyze a hypothesis and synthesize various pieces of information.

A focus on minimum competencies could be compared to a behaviorist approach to education in which all students are asked to accomplish the same tasks and perform them in a sequential order. This is true when educators focus too closely on a sequence of targets rather than the differentiation component of PBE. A proficiency-based curriculum may also unintentionally create barriers for traditionally underachieving students. If these students are too concerned with keeping pace with minimum competency tasks, they may never have an opportunity to pursue authentic learning opportunities connected to HOTS tasks. This aligns with Barton and Coley's (2009) analysis of the achievement gap and Torrf's (2014) study of the rigor gap, in which he observed that underachieving students were often given less rigorous tasks.

However, research shows that low-advantage students greatly benefit from being exposed to critical thinking opportunities in class (Schraw and Robinson, 2011).

Summary

Education is ever-evolving and, for the most part, current instructional best practices and structures differ from those of the past (Donaldson, 2014). Proficiency-based education has been presented as a shift away from the factory model of education that school structures have been built around in the past and towards far greater individualization (Schwahn & McGarvey, 2011). However, PBE combines both behaviorist and constructivist philosophical approaches into a single functional structure in an attempt to balance accountability and personalization. When viewed through this lens, it is evident why inconsistencies in PBE implementation exist between districts (Sturgis & Silvernail, 2014). However, research has shown that higher-order thinking tasks should be an essential component of modern educational practices (Paige et al., 2013; Matsumura et al., 2008; Cain, 2002; Hess & Gong, 2014; Larson et al., 1998). Thus, it is crucial to better understand how and to what extent such tasks are being integrated into PBE classrooms.

Pavlov and Watson's discoveries with classical conditioning in 1889 and 1913 informed how educators initially understood instruction and learning. Common instructional strategies for all students, direct instructional methods, a focus on essential knowledge, and a common curriculum were prevalent practices for much of the 20th century (Direct Instruction, 2013; Engelmann, 1999; Moore, 2011). In particular, an emphasis on common standards that all students must complete in a specific learning progression is part of the behaviorist legacy that has shaped a key component of PBE classrooms (MDOE, 2015; Sturgis, 2016; GSP, 2015).

The second facet of PBE is a clear focus on individualization and differentiation (Sturgis, 2016; GSP, 2015; MCCL, 2015). These aspects originate from the works of constructivist philosophers such as Dewey, Hebb, von Ehrenfels, and Wertheimer (Talebi, 2015; Brown & Milner, 2003; Phillips, 1976). Along with an emphasis on the individual rather than the group, there was also a greater focus on process over content (Phillips, 1976). Knowledge is constructed, and social interaction and educational experiences are highly valued (Talebi, 2015; Vygotsky, 1978). In the modern PBE classroom, the concepts of learning pathways, student voice and choice, and differentiation and personalization (Sturgis, 2016; GSP, 2015; MCCL, 2015) align with this philosophy. The importance of integrating higher-order thinking tasks in the classroom is also a constructivist concept.

Higher-order thinking skill tasks are an essential component of middle and secondary school classrooms. Their inclusion has led to an increase in student participation, engagement, performance on assessments, and problem solving abilities (Matsumura et al., 2008; Paige et al., 2013; Miller, 2003; Cain, 2002). Additionally, employers have rated HOTS among their most desired qualities for future employees (Hess & Gong, 2014; Larson et al., 1998). Today, educators use multiple taxonomies to assess whether tasks call for higher-order thinking, many of which have been designed to apply to all content areas (Bloom, 1956; Krathwohl, 2002; Webb, 2007; Hess, 2009; Daggett, 2017) and illustrate a learning progression in which high-end tasks call for methods such as synthesis, evaluation, and creation. Mathematics-specific taxonomies have also been developed (Skemp, 1976; Porter, 2002; Doyle, 1982; Smith & Stein, 1998). A key component of these frameworks is the contrast between procedural and comprehension tasks.

In the proficiency-based classroom, the inclusion of higher-order thinking tasks has been referenced across multiple frameworks (GSP, 2015; MCCL, 2015; Sturgis, 2016). However, tasks that require the use of HOTS are often included after the completion of foundational tasks

(MCCL, 2015) or through additional projects or portfolios (GSP, 2015). Research has shown that there is a risk in excluding low-achieving students from HOTS tasks, as this can lead to a rigor gap (Torff, 2014). To prevent this from occurring and ensure greater engagement and success among students, teachers must identify ways to regularly integrate HOTS tasks in their PBE classrooms.

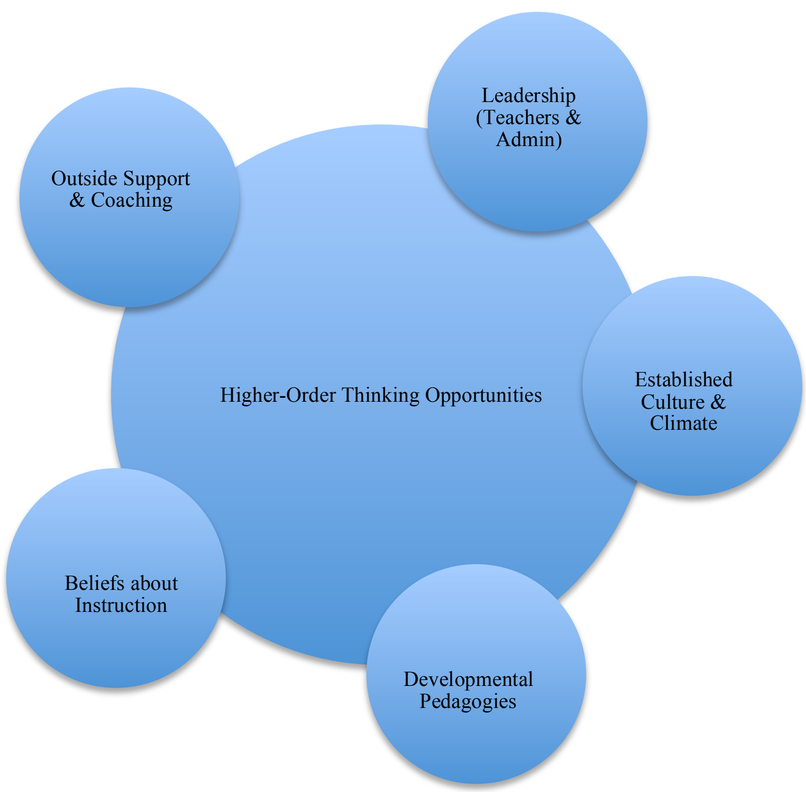
Conceptual Framework

The conceptual framework illustrates the multiple factors that may have an impact on the integration of higher-order thinking opportunities for students in a PBE system. Based on the literature review, each factor plays a role in whether HOTS opportunities are plentiful and meaningful for students. They are examined in greater detail throughout this research.

Each district that participated in this research has its own support structure for the implementation of PBE. One district is a member of a regional cohort, another works with an coaching organization, and the third uses its own district structure. The school administration and teachers had much to say about what elements would be emphasized when PBE is incorporated in their learning system. Will there be a focus on differentiation or individualization, or will there be more of a focus on grading and reporting? Some schools may find ways to balance all components of a proficiency-based system and incorporate opportunities for HOTS.

The culture and climate within a school play a key role in how change is perceived and acted upon. Are relationships and rigor key components of the current educational practices? Another question concerns how educational practices differ between middle and high school students. The IPI predicts that 18–25% of learning opportunities for students incorporate higher-order thinking tasks at the middle and secondary school levels. It is important to examine whether the interpretation of PBE in middle and high school has an impact on instructional opportunities. Finally, will the beliefs of educators themselves impact the level of learning opportunities that students can access? In mathematics, many believe that learning should be sequential. In other words, foundational knowledge must be covered prior to depth of knowledge tasks. This belief, coupled with a clear sequential order of standards and proficiency attainment, may impact the ability to incorporate HOTS opportunities for all students. The conceptual framework is represented in Figure 2.8.

Figure 2.8 Conceptual Framework Diagram



CHAPTER 3

RESEARCH AND DESIGN

The current research uses an explanatory sequential mixed-methods approach. Its goal was to better understand the frequency of opportunities for students to pursue authentic learning opportunities that require the use of HOTS in proficiency-based mathematics classrooms. In addition, it aimed to determine whether these opportunities differed according to professional support styles and between middle school and high school. The data is expected to provide educational leaders with a clearer understanding of how to assess and integrate higher-order thinking tasks in proficiency-based mathematics classrooms. In addition, it is expected to provide policymakers and educational leaders with a clearer picture of the impact that PBE efforts have on HOTS and its application in middle and secondary school classrooms.

Quantitative methods were used to better understand how frequently mathematics students in middle and secondary school are engaged in HOTS in proficiency-based classrooms. Classroom observations were conducted using the IPI (Valentine, 2007) to collect relevant data. The results of these observations were shared with educators in subsequent focus groups to solicit their help in interpreting the findings.

Qualitative methods were also used to better understand the perceptions of teachers who implemented higher-order thinking in their mathematics classrooms. Through focus groups, the data was analyzed to establish how mathematics teachers at the middle and high school levels used and implemented higher-order thinking in the classroom.

Research Questions Reviewed

The questions driving this research are:

1. How frequently are mathematics students engaged in high-order thinking in

proficiency-based classrooms?

a. Are there significant differences across professional support styles in the frequency of higher-order thinking opportunities?

b. Are there significant differences across grade spans (6-8, 9-12) in the frequency of higher-order thinking opportunities?

2. How and in what ways are mathematics students engaged in higher-order thinking when taught in a proficiency-based education system?

3. What factors enhance or impede educators' ability to implement higher-order thinking tasks in proficiency-based mathematics classrooms?

Definitions of Key Terms

Proficiency-based education: *An educational system in which time is flexible and students do not move on to the next skill until they complete the current skill (Maine DOE, 2015; Johnston, 2011).*

Higher-order thinking skills: *"Skills that enhance the construction of deeper, conceptually driven understanding" (Schraw & Robinson, 2011, p. 2).*

Research Design, Setting, and Participants

This research uses an explanatory sequential design. This method was chosen because the purpose of the study is to assess a trend in data and to explain the reasons behind it (Creswell & Plano Clark, 2011). In this case, the trend is students' rate of engagement in higher-order tasks in the classroom. Moreover, the use of focus groups enabled educators to explain the quantitative results. Three school districts in Maine participated in the research. For a district to be eligible to participate in the study, a middle school and the high school had to have implemented PBE. This means that each school had to have either been a member of a cohort or worked with a PBE coach for two or more years or used a 1 to 4 grading scale for two or more years.

The study took place at three middle schools of similar size (250–400 students) and three high schools of similar size (300–600 students) in Maine. Schools in Maine were selected due to the state's proficiency-based graduation requirements. Conducting the study in Maine was also of professional relevance to the researcher, as all participants provided information on PBE implementation in the state. These districts were selected based on their use of different PBE frameworks; the first district worked with the GSP, the second district was part of the MCCL, and the third district worked independently with different approaches for the middle school and high school. This enabled the researcher to examine several variations for the analysis of PBE approaches in Maine. Once the districts were selected, middle school and high school principals were contacted by phone to recruit their schools for the study.

A form of purposeful, criterion-based sampling was used to select participants for the study. For this research, it was important for participants to be from schools that had used PBE methods for at least two years. The appropriateness of school and participant selection was determined through the initial conversations with school principals.

Quantitative Research

Quantitative classroom data was collected for this study. Using the IPI, data was gathered and analyzed to answer the first research question. The data collection process consisted of classroom observations of 16 middle school mathematics teachers and 21 high school mathematics teachers. These observations were brief (1 to 3 minutes), as per IPI protocol. All school observations took place on separate days, with a total of 50 observations per school building per day. Each school was observed twice, which resulted in 550 total observations.

District 3 Middle School dropped out after the first round of observations. Therefore, only 50 observations in total were conducted at the school. The IPI was chosen as the data collection tool because it can help construct a school-wide picture of student learning, with an emphasis on engagement in higher-order thinking (Valentine, 2007).

The IPI is a classroom observation tool that was designed with the dual purpose of assessing engaged student learning and providing school faculties with data that can potentially improve instruction (Hunzicker & Lukowiak, 2012). Observers who use the IPI must rate student learning opportunities on a six-point scale in three broad categories (see Appendix C): student engaged instruction (5 and 6), teacher-directed instruction (3 and 4), and disengagement (1 and 2). The subcategories (or coding categories) are as follows: student active engaged learning (6), student learning conversations (5), teacher-led instruction (4), student work with teacher engaged (3), student work with teacher not engaged (2), and complete disengagement (1; Valentine, 2007). These are summarized in Table 3.1.

Table 3.1 Instructional Practices Inventory Categories

Broad categories	**Coding categories**	
Student engaged instruction	5. Student learning conversations	6. Student active engaged learning
Teacher-directed instruction	3. Student work with teacher engaged	4. Teacher-led instruction
Disengagement	1. Complete disengagement	2. Student work with teacher not engaged

The IPI observation protocol has very specific guidelines. The teachers being observed were informed at least a few days prior to the observation, and observations took place on a typical school day (i.e., excluding Fridays). The observations themselves lasted 1–3 minutes, and the observer recorded the learning experience seen upon entering the classroom. If students were having different learning experiences, the observer recorded the predominant pattern of learning. Additionally, scoring took place outside of the classroom to avoid any educational disruption. All observed teachers remained anonymous. Finally, all observers must be trained on the IPI process and receive a reliability rating of .90 or higher before using the tool for research purposes (Valentine, 2007). The researcher completed this training and received a 1.0 reliability rating.

Qualitative Research

Qualitative data was also collected for the study through two rounds of six focus groups consisting of five to seven mathematics teachers. The data was used to answer the second research question. Research recommends a group size of six to eight people to yield the best results (Morgan, 2009); the researcher aimed to stay as close to this recommendation as possible. The participants were all mathematics teachers from the same school. In total, there were five middle school focus groups and six high school focus groups. There were fewer focus groups in the former group due to District 3 Middle School dropping out of the research after the first round. The focus group format provided the advantage of a greater amount of contact with educators in a shorter amount of time (Eaton, 2017), with up to 90 minutes allocated for each focus group.

The focus groups took place at a comfortable school site and were audio recorded. All participants were given relevant information about the research and were informed that they could withdraw at any time prior to participating in the focus group. Assurances were made regarding their individual and school confidentiality. Each focus group began with a review of the IPI observation data gathered at the educators' school. Participants were then asked to help explain the data. To support the discussion, semi-structured focus group questions were used.

The guiding questions were the same for each group and focused on how teachers integrated higher-order thinking tasks into their mathematics classrooms (see Appendix A). Participants were asked questions about critical thinking routines, the explicit teaching of higher order-thinking skills, and student investigation and discussion. Emphasis on placed on whether the proficiency-based structure impacted opportunities for students to engage in these tasks.

During the analysis of data collected from the focus groups, conversations were examined at the group level rather than the individual level, as it is recommended for the group be the fundamental unit of analysis (Morgan, 2009). Additionally, efforts were made to distinguish between what participants found interesting and important. Open, axial, and selective coding were used to qualitatively analyze the focus group conversations. Coding was performed in NVivo, a qualitative data analysis program. Audio recordings of the focus groups were uploaded and transcribed prior to coding.

The coding process began with an initial round of open coding for all focus groups. The initial codes were grouped by like topics, and all fractured or split codes were given a new axial code. Once the axial codes were developed, they were organized into themes through the selective coding process. These themes are connected to the second and third research questions.

For example, the axial code *Clarity and Communication* was created with initial codes such as *Clear Expectations*, *Follows Directions*, *Clear Communication*, *Streamlined*. It represented a positive aspect of a school's move to a proficiency-based learning system. Thus, it was grouped within the theme of *Havens and Hazards of Proficiency-Based Education*. This theme then informed the third research question as a factor that enhances or impedes HOTS in the classroom.

Two additional themes emerged from the coding: *Attempting to Migrate Away from Standardized Learning and Instruction* and *The Continuum of Pedagogy and Impacts on Opportunity*. The latter informs the third research question, and the former informs the second research question, which aims to explain how HOTS is integrated in the classroom. The theme of *Attempting to Migrate Away from Standardized Learning and Instruction* consists of axial codes that represent instructional practices or structures in a PBE classroom, such as *Grading Practices*. This code comprises several open codes, including *Dual System, One Hundred Point Scale*, and *Four Point Scale, Rubric*.

The theme *The Continuum of Pedagogy* examines a diverse mix of learning opportunities within the classroom. It is organized with axial codes that represent IPI components. These made conceptual sense within the research and aligned well with the open coded data. An example of an axial code is *Teacher-Led Instruction*, which consists of open codes such as *Whole Class, Teacher Pace,* and *Direct Instruction*.

After the first round of focus groups, an additional round of 250 classroom observations was conducted. Again, the data was organized by school, then shared with teachers in the same focus groups to solicit their help in explaining the results. Particular attention was paid to any changes that took place since the previous round of data collection. Teachers were asked to explain what could have accounted for the changes to the best of their ability.

As previously mentioned, District 3 Middle School withdrew from the research before the second round of observations due to transitions taking place at the school. Thus, there were 550 observation points (50 observations per visit, for a total of 11 visits). In addition, 11 focus groups were held: one round after the initial observations and another round after the second observations.

Background of Participating Schools

The six schools that participated in the research all approached PBE implementation differently. Schools in District 1 relied on in-house professional development and their own understanding of what constitutes a PBE system. Schools in District 2 were members of the regional cohort (i.e., the MCCL), which provided an instructional model, suggested resources, and supported professional development for teachers. Finally, District 3 received coaching from the GSP and used its Framework for Proficiency-Based Learning.

District 1 Middle School. The middle school in District 1 is located in a former mill town with a population of around 5,000. The school serves Grades 5–8, but this research focuses on Grades 6–8. The school population is around 340 students. The student-teacher ratio is 11:1, and there is an 8% minority enrollment rate. As of the 2017–2018 statewide standardized testing, students in District 1 Middle School were 28% proficient in mathematics.

District 1 High School. The high school in District 1 is located in a former mill town that has a population of around 5000 residents. The school serves Grades 9–12. The school population is around 350 students. The student-teacher ratio is 13:1, and there is a 6% minority enrollment rate. As of the 2017–2018 statewide standardized testing, students in District 1 High School were 40% proficient in mathematics.

District 2 Middle School. The middle school in District 2 is also located in a former mill town, which has a population of around 2,900. In addition, neighboring towns with populations of 546 and 687 sent their children to the school. The school serves Grades 4–8 and has approximately 340 students. The student-teacher ratio is 12:1, and there is a 4% minority enrollment rate. As of the 2017–2018 statewide standardized testing, students in District 2 Middle School were 35% proficient in mathematics.

District 2 High School. The high school in District 2 is also located in a former mill town that has a population of around 2,900 residents. In addition, neighboring towns with populations of 546 and 687 send students to the school. The school serves Grades 9–12 and has approximately 305 students. The student-teacher ratio is 13:1, and there is a 5% minority enrollment rate. As of the 2017–2018 statewide standardized testing, students in District 2 High School were 18% proficient in mathematics.

District 3 Middle School. The middle school in District 3 is located in a central town in Maine with a population of around 3,300. At the time of research, this was one of two middle schools in the school district. The school serves Grades 5–8, and the school population is around 330 students. The student-teacher ratio is 15:1, and there is an 8% minority enrollment rate. As of the 2017–2018 statewide standardized testing, students in District 3 Middle School were 23% proficient in mathematics.

District 3 High School. The high school in District 3 is located in a central Maine town with a population of around 3,300 residents. Seven neighboring towns with an average population of 1,000 residents also sent their children to this high school. The school serves Grades 9–12, and the school population is around 570 students. The student-teacher ratio is 12:1, and there is a 7% minority enrollment rate. As of the 2017–2018 statewide standardized testing, students in District 3 High School were 40% proficient in mathematics.

Data Collection and Analysis

Data analysis for this research was both qualitative and quantitative in nature, taking place through the qualitative process of coding and generating relevant themes and the quantitative use of descriptive and inferential statistics. These processes were supported with appropriate software. Once the data was collected and analyzed, the results were merged and further examined.

Quantitative Data Collection and Analysis

The IPI tool was used for quantitative data collection. Exactly 550 total data points were gathered using the IPI; there were 100 mathematics classroom observations for all schools except District 3 Middle School, which had 50 observations. The data was examined both descriptively and inferentially.

Generating descriptive data from the IPI involved calculating the frequency, mean, mode, variance, and standard deviations for the rating score of each school, middle school and high school, and district. Each calculation was presented graphically and in table form to visualize the observation results. SPSS was used to analyze the quantitative data.

In addition, inferential statistical analysis was used in two ways. The first analysis was between the IPI scores generated between the middle schools and high schools as two groups. A two-sample t-test was used for this purpose. The null hypothesis was that the mean observation scores would be the same between the middle schools and high schools ($H_0 = MS = HS$). The alternative hypothesis was that there would be a difference between the means. A significance level of 0.05 was used. This analysis helped determine whether differing levels of higher-order thinking were observed between middle school and secondary school proficiency-based mathematics classes.

Secondly, an ANOVA test was run between school districts. In this case, the null hypothesis was that mean observation scores would be the same among all three districts ($H_0 = D_1 = D_2 = D_3$). The alternative hypothesis was that there would be a difference in the means. The significance level was also 0.05. This test helped determine whether districts implemented higher-order thinking differently in their classes according to the type of support structure (i.e., MCCL, GSP, or own system).

Qualitative Data Collection and Analysis

All focus groups conversations were audio recorded and transcribed with the transcription website Rev.com. The focus groups lasted about an hour each. Then, NVivo was used to code the transcripts. The coding followed a three-step process: open, axial, and selective coding. The open coding process focused on building categories of information, the axial coding process consisted of connecting the categories, and the selective coding process aimed to identify patterns in the codes (Saldana, 2009).

Explanatory Sequential Analysis

The study's mixed-methods design was chosen to allow the participants to interpret and explain the results of classroom observations made with the IPI. This design utilized prototypical follow-up explanations variant of an explanatory sequential design. In this method, the initial quantitative phase of research was prioritized, and the qualitative phase helped to explain the initial results (Creswell & Plano Clark, 2011). Qualitative data was used to answer the second research question (How and in what ways are students engaged in higher-order thinking when taught in a proficiency-based education system?), while the focus group discussions provided information on the types of structures that are currently in place at schools to integrate higher-order thinking in proficiency-based mathematics classrooms.

Regarding the first research question (How frequently are mathematics students engaged in high-order thinking in proficiency-based classrooms), IPI data on the frequency of higher-order thinking tasks in schools was compared to coded data from the focus groups from each school. The observation results were correlated and prepared for sharing with focus group participants (Creswell & Plano Clark, 2011). The third research question (What structures do educators feel enhance or impede their ability to implement higher-order thinking tasks in proficiency-based mathematics classrooms?) was addressed through the results obtained from the explanatory sequential portion of the focus groups (Creswell & Plano Clark, 2011). Ultimately, answering the three research questions provided a more complete understanding of how higher-order thinking can be best integrated in proficiency-based mathematics classrooms. Table 3.2 summarizes the research tools used.

Table 3.2 Summary of Research Tools

Research tool	**Type of data**	**Participants**	**Research question addressed**
IPI observation tool	Quantitative	37 total middle school and high school mathematics teachers Two rounds of 550 observations in total	Research Question 1 (a, (b)
11 focus groups (six focus groups in the first round and five in the second round; three middle school and three high school focus groups)	Qualitative	Groups of five to seven mathematics teachers	Research Question 2 Research Question 3

Trustworthiness

By design, explanatory sequential research enlists participants in reviewing the quantitative results and providing credibility to the study. The examination of three school districts with different approaches to PBE and a focus on middle and secondary schools increased the transferability of the results. Common themes or findings that were true across all three districts were more meaningful than those derived from a single school with a specific approach to PBE implementation.

The researcher completed training with the IPI observation protocol and obtained a 1.0 reliability rating; a rating of .90 is required prior to conducting research (Valentine, 2007). While there is potential for bias, a long-standing process was in place to address and reduce bias through training. The observational tool was used in the same manner from school to school.

Internal and External Validity

When examining educators' perceptions of student engagement and achievement, it is important to account for confounding variables that may impact the internal validity of the research, including quality instruction, student groupings, and school climate. Another factor is the degree of fidelity with which the chosen classroom is implementing PBE structures, which makes the screening process particularly important.

Regarding external validity, it is important to avoid making generalizations from one school to the next. Schools in this study use different variations of PBE, and the research focuses on how educators integrate HOTS opportunities. Thus, the findings may not be fully generalizable due to preexisting structures in other schools.

Ethical Issues

The most important ethical issue is confidentiality. Teachers in the focus groups were assured that their responses would not be shared with their colleagues or other participants. If they felt that their principal could gain access to their responses, they may not have given open and honest answers. A participant could choose not to answer a question at any time if they did not want to. This was made clear to them in the letter of consent prior to participation.

In addition, the study adhered to Institutional Review Board guidelines on participant selection and permissions received. Participants could opt out of the research at any time. All focus group data was stored on a password-protected device until it was no longer needed for research purposes. All audio recordings were destroyed once transcription was completed.

Possible Limitations

The first limitation of this research is that there is no universally accepted definition of PBE. Thus, one challenge could be generalizing results from one school or classroom to the next. However, the research aims to provide examples of PBE structures that are proven to engage students in higher-order thinking across districts and grade levels.

Another limitation is the lack of quantitative data connected to PBE structures. Due to the varied nature of PBE, this information was collected through the focus groups, which enabled the researcher to probe deeper into areas that teachers emphasize in relation to PBE. The administration of a survey would not have allowed as much depth in this regard.

Finally, the structure of the focus groups is a possible limitation because they consisted of faculty members from the same school. Thus, participants could have felt pressure to provide answers that aligned with the views of their peers in the group. It was the researcher's responsibility to guide the conversations around the focus group questions and ask appropriate follow-up questions to determine participants' true opinions.

CHAPTER 4

FINDINGS

This chapter presents the results of the explanatory sequential mixed-methods study, which were designed to answer the following research questions:

1. How frequently are mathematics students engaged in high-order thinking in proficiency-based classrooms?
    a. Are there significant differences across professional support styles in the frequency of higher-order thinking opportunities?
    b. Are there significant differences across grade spans (6-8, 9-12) in the frequency of higher-order thinking opportunities?
2. How and in what ways are mathematics students engaged in higher-order thinking when taught in a proficiency-based education system?
3. What structures enhance or impede educators' ability to implement higher-order thinking tasks in proficiency-based mathematics classrooms?

In keeping with the explanatory sequential approach, the quantitative data was collected and the qualitative results were subsequently used to identify the story behind the data. The quantitative data for all six participating schools was analyzed with descriptive statistical methods. Then, an independent samples *t*-test was used to compare the means of middle schools and high schools. Finally, an analysis of variance was conducted on data from the three participating school districts.

The qualitative data collected from the focus groups was transcribed and coded at three levels: open, axial, and selective. The open coding process helped to determine the initial codes, axial coding was used to combine split codes, and selective coding helped to identify emergent themes in the codes (Saldana, 2009). Coding at each of these levels generated clear themes from the data. This chapter includes several graphical representations of the coding results, along with quotations from the research participants to support the emergent themes.

Sample

In total, 37 mathematics teachers from three school districts (three middle schools and three high schools) participated in this research. Classroom observations with the IPI tool were conducted in a manner that allowed for as equal a distribution of data as possible. School principals provided a class schedule to the researcher, and the researcher mapped out a plan to gather 50 observation points over the course of a regular school day. Three minutes were spent in each class, then the researcher moved on to the next classroom. Each classroom visit resulted in the logging of an IPI score. In many cases, the researcher was able to return to a classroom and log an additional IPI score during the same class period. Mathematics classes observed at the high school level ranged from prealgebra to AP Calculus. At the middle school level, each school had a specific heterogeneous grade level mathematics class.

The focus groups included all teachers that participated in the classroom visit portion of the research. However, staffing changes occurred at District 2 High School and District 3 High School between the two rounds of research. This resulted in the participation of one different teacher from each school. The number of focus group participants remained the same between the two rounds. The breakdown of participants is presented in Table 4.1. Each focus group member was given an opportunity to speak. Equal weight was given to all respondents' answers when coding the focus group transcripts.

Table 4.1 Summary of Participants by School

School name	District 1 Middle School	District 1 High School	District 2 Middle School	District 2 High School	District 3 Middle School	District 3 High School	Total
Total number of participants	5	6	6	7	5	8	37
Focus group participants	5	6	6	6	5	7	35

Data Collection

The classroom observations and focus groups at participating schools occurred over a two-year period. The first round occurred in the winter and spring of the 2017–2018 school year. The second round took place in the winter of the 2018–2019 school year. When possible, the focus groups took place on the afternoon of an observation date. When this was not convenient for participants, an alternate date was selected.

The IPI tool was used for data collection during mathematics classroom observations. The researcher observed each class for three minutes, then scored the instructional practices on the IPI scale. The IPI categories are student engaged instruction (5 and 6), teacher-directed instruction (3 and 4), and disengagement (1 and 2). The subcategories, or coding categories, are student active engaged learning (6), student learning conversations (5), teacher-led instruction (4), student work with teacher engaged (3), student work with teacher not engaged (2), and complete disengagement (1). A score of 5 or 6 indicates that higher-order thinking is taking place (Valentine, 2007).

All data was collected on individual IPI data recording forms (see Appendix). A form was created for each period of the day, and a score was entered on each line. Anecdotal notes were avoided unless there was a justifiable reason to remove a score (i.e., a fire drill took place). Once the total of 50 observations for the day were completed, data collection ended. The observations typically took a full school day to complete due to the logistics of moving from class to class and the restrictions placed on observation times by class schedules. The data was logged into a secure spreadsheet and reviewed multiple times for accuracy.

Focus groups ranged in size from five to seven participants and took place in a comfortable setting at their own school. Once participants entered the room, the researcher reminded them that the data would be de-identified and that they could opt out at any time. Then, he asked everyone for permission to record the session on the Voice Recorder iPhone app. The sessions were later transcribed by hand and with the Rev.com transcription service.

Each focus group began with a review of the de-identified observation data that was collected in the participants' classrooms. They were given an opportunity to ask questions and discuss whether they agreed the results. Guiding questions were used to help direct the ensuing discussion (see Appendix A). Each focus group session lasted for around one hour.

Quantitative Data and Analysis

All IPI data was uploaded into spreadsheets and organized by school and observation date. For each school, the data was analyzed using descriptive statistics for frequency, mean, median, mode, variance, and standard deviation. The percentage of higher-order thinking instructional practices was also calculated; these consisted of the frequency of 5 and 6 scores in the total number of data points. These values were calculated for each school visit and for both visits combined. The summary data for each school is shown in Tables 4.2 to 4.7.

Table 4.2 Summary Data for District 1 Middle School

	Frequency	**Mean**	**Median**	**Mode**	**Variance**	**Standard deviation**	**% HOTS**
Visit #1	50	2.92	3	3	0.851	0.922	2.00%
Visit #2	50	3.26	3	3	0.727	0.853	6.00%
Total	100	3.09	3	3	0.810	0.900	4.00%

Table 4.3 Summary Data for District 1 High School

	Frequency	**Mean**	**Median**	**Mode**	**Variance**	**Standard deviation**	**% HOTS**
Visit #1	50	3.62	3	3	1.261	1.123	18.00%
Visit #2	50	3.88	4	4	0.802	0.895	18.00%
Total	100	3.75	4	3	1.038	1.019	18.00%

Table 4.4 Summary Data for District 2 Middle School

	Frequency	**Mean**	**Median**	**Mode**	**Variance**	**Standard deviation**	**% HOTS**
Visit #1	50	3.36	3	3	0.970	0.985	10.00%
Visit #2	50	3.78	3.5	3	1.073	1.036	26.00%
Total	100	3.57	3	3	1.056	1.027	18.00%

Table 4.5 Summary Data for District 2 High School

	Frequency	Mean	Median	Mode	Variance	Standard deviation	% HOTS
Visit #1	50	3.52	4	4	0.949	0.974	6.00%
Visit #2	50	3.94	4	4	0.792	0.890	18.00%
Total	100	3.73	4	4	0.906	0.952	12.00%

Table 4.6 Summary Data for District 3 Middle School

	Frequency	Mean	Median	Mode	Variance	Standard deviation	% HOTS
Visit #1	50	3.88	4	3	0.924	0.961	24.00%
Visit #2	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Total	50	3.88	4	3	0.924	0.961	24.00%

Table 4.7 Summary Data for District 3 High School

	Frequency	Mean	Median	Mode	Variance	Standard deviation	% HOTS
Visit #1	50	3.70	4	4	0.827	0.909	16.00%
Visit #2	50	4.02	4	4	1.163	1.078	32.00%
Total	100	3.86	4	4	1.011	1.005	24.00%

Notably, an IPI score of 3 (student work with teacher engaged) was the most common score logged during all middle school visits. All five of the middle school visits had a mode of 3. However, a score of 4 (teacher-led instruction) was most common among the high schools. Five out of six high school visits had a mode of 4, which had an impact on mean IPI scores. The following section summarizes total mean IPI scores by school, ranked from highest to lowest. This score represents the average of each visit.

Mean IPI Score by School

When using the IPI tool for middle school or high school classroom observations, school fall into the 18–25% range on average for higher-order thinking instruction (Valentine, 2007). A score of a 5 (student learning conversations) or 6 (student active engaged learning) is indicative of higher-order thinking practices. Table 4.8 summarizes the total percentage of HOTS observed at each school. They are organized from highest to lowest percentage observed per school.

Table 4.8 Mean IPI Scores and Percentage of HOTS Observed by School

	Mean IPI score	**% HOTS on the IPI (score of 5 or 6)**	**Frequency**	**School type**	**District**
District 3 Middle School	3.88	24.00%	50	Middle school	3
District 3 High School	3.86	24.00%	100	High school	3
District 1 High School	3.75	18.00%	100	High school	1
District 2 High School	3.73	12.00%	100	High school	2
District 2 Middle School	3.57	18.00%	100	Middle school	2
District 1 Middle School	3.09	4.00%	100	Middle school	1
Total	*3.63*	*16.00%*	*550*	*MS/HS*	*N/A*

The quantitative data was entered into the statistical analysis program SPSS. An ordinal variable was created for the IPI scores, and nominal variables were created for school, school type, district, and visit.

To determine whether there was a significant difference in IPI scores recorded at the middle schools and high schools, a two-sample *t*-test was used. The null hypothesis stated that mean observation scores between middle schools and high schools would be the same ($H_0 = MS = HS$). The alternative hypothesis stated that there would be a difference in the means.

For this analysis, the variable "SchlType" referred to middle school or high school. Middle school was designated as the first school type, and high school was designated as the second school type. To conduct the *t*-test, the means were first calculated for each school type, then compared (see Table 4.9).

Table 4.9 Comparison of Mean IPI Scores for Middle Schools and High Schools

School type	**Data points**	**Mean IPI score**	**Std. deviation**	**Std. error mean**
Middle schools	250	3.44	1.009	0.064
High schools	300	3.78	0.991	0.057

Next, the *t*-test was conducted, and the results showed a statistically significant difference between the mean IPI scores of middle schools and high schools ($p < .001$). Namely, high schools are statistically significantly more likely to have a higher mean score on the IPI. Thus, among the participating schools, high school students had more opportunities to engage in HOTS learning tasks in proficiency-based mathematics classrooms than middle school students. The results are shown in Table 4.10.

Table 4.10 Independent Samples *T*-Test

F	**Sig.**	***t***	***df***	**Sig. (two-tailed)**	**Mean difference**	**Std. error difference**	**Lower**	**Upper**
0.299	0.585	-3.974	548	< 0.001	-0.340	0.086	-0.508	-0.172

The next analysis was between school districts. An analysis of variance was conducted on the data from all three districts. The null hypothesis stated that the mean observation scores would be the same across all three districts ($H_0 = D_1 = D_2 = D_3$). The alternative hypothesis stated that there would be a difference between the means. The results demonstrated that there was a significant difference between districts. The significance level also fell below .001. The results are shown in Table 4.11.

Table 4.11 Analysis of Variance

	Sum of squares	***df***	**Mean square**	***F***	**Sig.**
Between groups	17.290	2	8.645	8.668	< 0.001
Within groups	545.553	547	0.997		
Total	562.844	549			

In addition, Tukey's honest significant difference test was run to determine which districts there was a significant difference for. A significant difference was found between District 1 and District 3, F(*df* between groups, *df* within groups) = F score, $p < 0.001$. In other words, schools in District 3 had a mean IPI score that was significantly different from schools in District 1 ($p < 0.001$, 95% C.I. = [-0.70, -0.19]). However, no statistically differences in mean IPI scores were found between District 1 and District 2 ($p = 0.056$) or District 2 and District 3 ($p = 0.111$). District 3 relied on external coaching to support its implementation of PBE in schools. District 1, on the other hand, relied on their own methods and did not seek external support. Moreover, this district had a distinctly different approach to PBE between middle schools and high schools. The results of Tukey's honest significant difference test are shown in Table 4.12.

Table 4.12 Tukey Multiple Comparisons

(I) District	**(J) District**	**Mean Difference (I - J)**	**Std. error**	**Sig.**	**Lower bound**	**Upper bound**
District 1	**District 2**	-0.230	0.100	0.056	-0.46	0.00
	District 3	-0.447*	0.108	< 0.001	-0.70	-0.19
District 2	**District 1**	0.230	0.100	0.056	0.00	0.46
	District 3	-0.217	0.108	0.111	-0.47	0.04
District 3	**District 1**	0.447*	0.108	< 0.001	0.19	0.70
	District 2	0.217	0.108	0.111	-0.04	0.47

*The mean difference is significant at the 0.001 level.

Qualitative Data and Analysis

All focus group sessions were audio recorded. Once created, the recordings were uploaded to a secure computer and de-identified. They were then transcribed and uploaded to NVivo 12 to initiate the three-round coding process. The latter consisted of open coding, axial coding, and selective coding. Open coding was used to determine the initial codes, axial helped to combine the split codes from the first round of coding. The selective coding helped to determine the emergent themes from the coding (Saldana, 2009).

First, the open coding round involved a line-by-line manual analysis of each transcript in NVivo. Each line was coded with a unique open code, or a new open code was developed for it. By the end of the open coding stage, there were 189 open codes. These codes were organized by topic, and fractured or split codes were combined during the axial coding process. Once axial codes were developed, a selective coding process was used to identify overall themes.

Each set of codes was also examined at the school level to construct narratives that connected the major ideas reported by participants in the focus groups and identified from the classroom observations. These individual school narratives provided a better understanding of the extent of HOTS integration into PBE classrooms, along with factors that aided or impeded their implementation. Moreover, the school-level coding and narratives informed the answers to the second and third research questions.

When all the schools were examined together, three major themes emerged. Two themes informed the third research question (What factors enhance or impede educators' ability to implement higher-order thinking tasks in proficiency-based mathematics classrooms?), specifically "Attempting to Migrate Away from Standardized Learning and Instruction," and "The Havens and Hazards of Proficiency Based Education." The third theme informed the second research question (How and in what ways are mathematics students engaged in higher-order thinking when taught in a proficiency-based education system?), specifically "The Continuum of Pedagogy" and "Impacts on Opportunity."

The first theme relates to educators' attempts to move away from traditional teaching and learning practices and towards PBE. Within this theme, teachers spoke about changes in grading practices, assessments, common standards, student support structures, and individual practices at their schools.

The second theme relates to educators views' about the havens and hazards of the PBE system. These findings are based on their experiences with the PBE implementation process. Participants viewed PBE as having a positive impact on clarity, communication, and individual relationship building. They also emphasized the importance of including direct instruction in their teaching practices. Direct instruction has been found to have a 0.6 effect size in a meta-analysis of teaching strategies (Hattie, 2012). However, challenges related to negative student outcomes, failure to build support for a move to PBE, and the necessary paradigm shift emerged.

The final theme is the continuum of pedagogical practices taking place within the classroom. This continuum was categorized using the IPI framework. Practices ranged from complete disengagement to students' full engagement in higher-order thinking tasks. By explaining their own instructional practices, teachers provided a descriptive framework through which to view the IPI observation data collected during the quantitative portion of the research.

Themes by School

After the qualitative and quantitative data analyses were completed, themes emerged from each participating school. Classroom observation data was used to triangulate the focus group themes and create a narrative for each school. These findings helped pinpoint the support structures at each school that foster more meaningful higher-order thinking opportunities in mathematics classrooms, thus informing the second and third research questions.

District 1 Middle School Findings. Teachers in District 1 Middle School used internal resources to build their own model for teaching mathematics in a proficiency-based system. The resulting structure heavily relies on students working at their own pace under the philosophy that mastery drives movement. Focus groups were held with mathematics teachers at this school over a period of two years. Despite having greater ownership over how instruction was implemented in the classroom, the participants' remarks revealed a lack of buy-in and understanding of PBE. In the first focus group, a second-year teacher shared her experience teaching in a new system:

> *My first year, I think it was just the first year was zero to 100. I think last year is the first year we used 1 through 4. I don't really care for it. I understand the idea behind it, but I don't know, I'm not a fan of it. I wish there was a way to do some hybrid model.*

In the same focus group, another teacher explained that he also did not enjoy the school's new grading structure:

> *Honestly, I don't know. I find the grading boring. One, 2, 3, or 4 over and over again. I don't know. I like having the standards. I think we should have standards and maybe if they're on pace, behind pace, or 1, 2, 3, 4.*

These teachers also experienced challenges when the school administration established a mathematics-specific professional development session for them. The trainer's aim was to help them integrate higher-order thinking tasks into their teaching practices. However, due to a lack of buy-in, the participants did not appear to have integrated any of these strategies. One teacher said the following about the professional development session:

> *It's interesting. We have someone from Math in Focus come to our school and do demonstrations. The math teachers don't necessarily agree with how she does it. It just seems kind of time wasteful. It takes a lot of time for …*

At this point, another teacher jumped in to complete the thought.

> *She wanted us to—sort of telling kids the formula for a square or something. She wanted us to draw pictures and try to get them to come up with a formula which okay, maybe for somebody who's more advanced, we could try that. But for an average kid or below-average kid, I just started giving the formula and work from there.*

When asked follow-up questions regarding the training and the program, one of the participants shared that she liked the program but wanted to "adapt it to how we think it needs to be used."

In mathematics classrooms at District 1 Middle School, students can be observed mostly working on their own at their own space, with a teacher situated at the front of the classroom or walking around the room checking in with individual students. On the IPI, this setup corresponded to a category of 2 (students working on lower-order tasks with the teacher not engaged) or 3 (students working on lower-order tasks with the teacher engaged), which was reflected in the school's mean IPI score of 3.09 based on 100 classroom observation data points (2.92 in the first observation year and 3.26 in the second observation year). The total percentage of higher-order thinking tasks observed was 4% (2% in the first observation year and 6% in the second observation year).

During the first visit, the participants said that they focused on using videos with students, which allowed them to learn on their own at their seat and move through the curriculum at a comfortable pace. The teachers support students who have questions.

Yeah. Six, seven, eight, we prefer to do the videos so kids can kind of learn at their own pace and just move on when they're ready and not feel rushed to move on.

In addition, each teacher played a role in the creation of the videos. Once recorded, the videos were shared with other mathematics teachers at the school for use with students.

I created the sixth grade videos and I share them with the other sixth grade teacher and then teacher made some seventh and eighth grade videos that him and the other seventh and eighth grade teachers use.

This approach provided some of the students with personalized content by their own teachers, while others learned from videos that were created by other teachers at the school. Once they watched a video, the students completed practice problems at their desk and submitted them to their teacher for correction. Some teachers used online software to correct the problems and provided the students with feedback. Participants were asked if they have noticed a change in engagement levels under this structure. One teacher explained believed that students seemed more engaged:

I feel like they are more engaged than previous years because where they're doing their own thing, it's more time for me to walk around to make sure they are at least pretending to be engaged. Whereas before, when I would teach up at the board, somebody could not be engaged and I might not really be able to tell.

Thus, the new structure provided time for individualized support and feedback, albeit perhaps at the expense of quality instruction from the teacher. Not all teachers spent the full class time moving from one student to the next. One participant explained that he used to have a method for students to sign up for help, but they now simply raise their hand. He said that they were independent and did not need to do this very often:

At the beginning of the year, I started with, I would write "help" on the board and they'd write their names and I'd just go in order of who needs help and then, as the classes have moved on, it's morphed. My homeroom, they can just raise their hand. They're pretty independent, so they don't need help too often, so they just raise their hand and bop around the room.

In this classroom instruction model, students work independently on lower-order problems most of the time while the teacher is disengaged at their desk. This setup would be scored as a 2 on the IPI. During the second round of observations, the researcher noticed that teachers began to spend more individual time with students. As a result, the mean IPI score increased from 2.92 to 3.26. Although this is not a significant leap, participants mentioned the need to do this was mentioned during the focus group discussion. During the second-year focus group, a teacher shared the following:

> *Now the time I see, a lot of time that students are working on problems with support in the school. And that typically, like the homework-type problems that they, you make sure they have time in school to work on those.*

Some teachers even began to push students more at a teacher pace to keep them up closer to the pace of the class. Participants explained that, in the past, some students did not make much progress; thus, they had to make changes to benefit lower-level students. One teacher said the following:

> *That was a big change I made this year, where it was all at your own pace last year, we found all those lower students that I could have a kid who's still in Chapter 2 right now.*

During the second year of the study, the teachers still asked students to work independently. They spoke about the resources that they used and found that an online tool called ASSISTments to be very helpful. It grades student assignments for the teachers and provides immediate feedback, which eliminates the need for assignments to be submitted to the teacher for grading.

> *I think the website that we use for our book, the ASSISTments website, I think that is huge for those kids who are going at their own pace instead of them doing, they go through the book, they have no idea whether they're getting it right or wrong and they could think they're doing great and have that all wrong. And that ASSISTments site, after they watched that lesson is giving them that instant feedback. Yeah, this is right. You're doing well. Or if they're getting them wrong,*

they're like, "Oh, I need to go back and check something and ask for help" versus them just thinking, "Oh yeah, this is great."

However, the teachers still spoke about confusion among students and parents during the second-year visit. Of all the participating schools, District 1 Middle School made perhaps the greatest structural change. This may still have been a need for clarifying purpose and building support and understanding of the change prior to implementation. One participant described how the students felt about the change:

It's tough for the kids. They're used to it by the time they get to eighth grade, but where I teach seventh and eighth grade, those seventh graders, they have no idea what passing is. And they still ask, "Is this a passing grade?" We're almost to February now, we're a semester through, and they're still unsure. They're just so used to from kindergarten through sixth grade, that 1 through 4, that would be a 65 or an 82, both, "Is this passing?"

Another teacher explained that "parents don't get it." It is easy to see why they may have trouble with this particular instructional approach. The teachers shared that students watched videos at their desk and worked on practice problems, moving at their own pace. Some teachers frequently moved from student to student, while others waited for a hand to be raised or a student to write their name under the "Help" sign on the board. This mode of mathematics instruction was found to be devoid of higher-order conversations or enthusiastic learning.

District 1 High School Findings. Teachers at District 1 High School also used internal resources in their implementation of a proficiency-based instructional system. At the high school level, they continued to use a direct instruction model while paying close attention to the needs of individual students. Their focus was on grading and assessments in the proficiency-based model. Focus groups took place over two academic years. During the first year, participants shared that the proficiency-based system for grading, which allows retakes, was detrimental to mathematics students.

> *They don't do well, they can retake it. So, especially the smart lazy kid knows that I'll just take a look at the test, and if I don't like it, I'll just redo it anyway. They don't have the pressure on them to do well as they will when they do go to school. So, I think it's some ... They learned, but I think they've learned something negative.*

In the second focus group, participants said that the retake policy was recently changed, which made holding students accountable a bit easier.

> *It was hard for me at first, but I think the more we do it, I'm just in the groove. And then we switched this year. We have a new policy where if kids don't pass a standard, they can retake it. But they can only get a 76 on it. Whereas before, they could retake it 10 times if they wanted to get 100 on it. So, it's made them a little bit more aware of being right the first time kind of thing. So, that's helped, I think, a lot in my room.*

Participants believed that the common standards have been helpful for both teachers and students. A teacher in the second focus group explained that he was in favor of the move to grade requirements on the makeups and the standards:

I like the standards. I like still keeping that 76. I don't know about you guys, but I like that. I'm glad that they can't retake them as many times as they want to. But I still like using the standards because then I can say, "Okay, we've already learned this skill and you wouldn't have been able to do this had you not got a [inaudible], at least you're more proficient in that skill." And I'll build on it a lot. Especially with math, it's easy to build on those skills and to bring those back up again, you know.

However, the participants did not fully buy into the concept of standard mastery. They believed that repetition was important for students to truly master the content.

It may hold them accountable for that standard, but let's face it, I think we all can admit that I can give a test to kids today and next Friday the results would be very, very different. In other words, they got the skill for now. But do they really have it? You know what I mean?

To ensure that all students progress through the content, one of the participants introduced a teacher pace in their classes. In each class, a new lesson was taught, and the students followed along with it and completed related work. If they had not completed the previous work or assessment, they were still responsible for it.

They have to do everything basically for me. If there's a teacher pace, then let's keep up with, but there's no skipping parts. They have to work their way to the summit. So, they have to get it done, then they can move on. So, you'll have some students that may be behind teacher pace, then others that are ahead of teacher pace.

Another participant said that the students were most engaged when he worked through a direct instruction model :

I think for me, when it feels the best is like they've learned how to do something and then we're doing problems on the board and I'll kind of let them take over and it kind of becomes a competition, okay. Who's going to get this right. Who's going to get this right. And someone will shout out an answer, and someone else will look and say that can't be right because of whatever, because of this and this. And then they'll look at it and then they'll find their mistake. And there's a lot of back and forth and a lot of interaction that way.

In addition, participants shared that the instructional model worked well due to the relationships that they had built with students. One participant said that she understood the needs of students in her class and that the relatively small size of the school helped to make this possible:

Part of that, I think, comes from the kids that we teach, we know them. We know which one probably didn't have dinner last night and they were hungry this morning. So, if you want an apple, go in my back room. You know, those kinds of things. We know who did well in the basketball game last night so you can say, "Good job on game last night." And I don't know that you'd be able to do that in a bigger school. If you would have those connections. Those kinds of things. So, it's kind of easy to be connected because we know them so well.

For the most part, mathematics teachers at District 1 High School used a more traditional instructional model in their classes. They frequently taught lessons at the board and engaged their students in question-and-answer sessions about their current learning. From Year 1 to Year 2 of the study, classroom observation data showed a slight shift to teacher-led instruction. During the first year, the mode and median IPI scores were 3 (individual student work with teacher support). During the second year, the mode and median IPI scores were 4 (teacher-led instruction). The mean IPI score also increased from 3.62 to 3.88, which indicates that scores of 4 were assigned more frequently to classroom observations. In the second focus group, a teacher explained that he used to let students work on problems on their own, but he now integrated more opportunities for the entire class to discuss when he felt that they were struggling:

> *In addition, sometimes I bring them back to the overhead and start going over a problem to reengage them, too. That's what I try most times when they're trying to not do the work.*

During both rounds of data collection, the researcher observed that HOTS were implemented in the classroom 18% of the time. The participants provided examples of HOTS opportunities that they offered their students. In each case, the opportunities directly tied into daily lessons.

> *I did one of the other day, trigonometry, where something was wrong on two questions and they had to figure out what was wrong. And this was after they had already learned about the intro and all that stuff. And they practiced it a little bit so they can see, that's wrong because they didn't divide. They multiply instead of divided. So, they know the process. And this is wrong because that should have been a negative sign instead of a plus sign. So, we're trying to work on problems like that in our classes because we found that that's actually helping them to understand it better.*

Another participant said that she began by building a foundation through daily learning, then worked up to more challenging questions for her students:

> *And then it progressed to like a problem where they actually had to find x. But it was just a triangle and word problems, real-world problems. And then it escalated to the word*

problems and putting it all together. So, it definitely is like a clear process of they're starting at a lower level and then escalating to that. And same thing with that class, because we started yesterday with just getting to know the initial piece of that group.

District 2 Middle School Findings. District 2 Middle School is a member of a regional cohort (i.e., the MCCL) that provided guidance and curriculum resources for the school's initial integration of a PBE system. The resources included guidance on making standards clear and using rubrics to score students. In the first focus group, a teacher described the impact of the cohort:

I definitely communicate more with the kids about what we're doing and what the expectations are. What the standard is and why we are doing it... Why we need it now. I'm definitely having more conversations with kids just based on we had it all written out of what needed to do be able to succeed.

Participants shared that their instruction mainly focused on Level 3 work, in which students attempt to attain proficiency on a standard. Higher-order work is mainly connected to the attainment of 4s.

The target that we have in math are generally the higher of thinking would be the 4s, so that I think you could see a lot of the Level 3 type of things going on to reflect the targets we're required to meet with all the kids.

While opportunities for Level 4 work exist, they seem to be more the exception than the rule. Teachers attempt to find ways to provide Level 4 learning opportunities for students who need it at the time rather than the entire class.

I have a bit of a challenge, at one point our GT kids, have been it's just a boost of an extra grade so we don't really have all the standards in a particular basis so *it's challenging when you have that student and to be able to keep everybody learning.*

A level of inconsistency was found among participants, as some teachers would add Level 4 problems into their assessments and ask students to attempt them there. In other words, students problems did not focus on these problems during regular class instruction time but only on assessments. A teacher described his expectations for students on these questions:

> *It's an expectation in my class that the if you hand me a test that you have to at least attempted a 4.*

During the first school visit, the class observations did not indicate that students received many higher-order thinking opportunities during their regular class time. The mean IPI score was 3.36, and HOTS opportunities occurred 10% of the time. There was a marked increase in the IPI data recorded during the second-year observations; the mean IPI score rose to 3.78, and the percentage of HOTS opportunities grew to 26%. When this data was shared with participants in the second focus group, they said that the mathematics teachers had introduced a new curriculum resource called enVision 2.0 prior to the second year. It offered more opportunities for students to engage in HOTS in class and required them to share and discuss their thinking. In addition to this change, the teachers shifted away from the minimum proficiency requirements recommended by the MCCL. When teachers were asked what accounted for the change, one participant responded as follows:

> *enVision 2.0, and that's why there is a high order thinking questions in every lesson. So, we really didn't know and realize what kind of level we were supposed to teach. Because before, we were, we thought we were teaching higher-order thinking. We're obviously not. And that's why we see the type of questions, we see, [teacher's*

name] made for us, for every lesson there are questions that students have to decide as a group. That's my answer, enVision. That's so key.

While participants said that their lessons and the level of HOTS opportunities in the classroom differed from day to day, the new curriculum resource gave them the tools to incorporate HOTS at a greater rate. Although enVision was not a perfect fit for students' current understanding of the material, they adapted it to make it fit.

I think for me, it was the day. I mean, it depends on what the topic is. For the eighth grade enVision, I have to do a lot of backtracking. For seventh grade skills, I do a lot of frontloading to be able to get to the enVision topic. So, it depends on what the topic is. Today, when we were doing the decimals activity, I mean that just lent itself really nicely to just go back and revisit unit rating. And that lent itself to some higher-order thinking. There was another one that I wish I had done. I let them do it on Friday. And it was one that was a little more basic, but it was equally as good an activity for them to do. I try not to do, like today it was a lot more teacher-based instruction. It just depends on the day, what the topic is.

A clear difference that emerged during the classroom observations was that students explained their thinking or problem-solving process much more frequently. Teachers seemed to be very proud of this change in their practices, and it was clear they saw a difference in the classroom.

Another thing, in enVision, in many problems they ask students to explain, which is higher-level thinking. So, in the beginning they say, I know the answer, but I can't explain. So, that's what the enVision was teaching them how to communicate with each other, how to ask questions, how to analyze problems. That's valuable as well.

The change in teaching philosophy was aptly summarized by one teacher. She explained the impact of setting higher targets and giving all students the opportunity to use HOTS in her class:

Even if they're doing it together. There are times when we have a pretty high-level target and we're working and I'm like, "Let's just do a few of these problems together." They've at least experienced them. They've seen that level of problem. And more often than not, they'll say, "Try this one." And they can do it. And they never would've had that opportunity last year.

District 2 High School Findings. Mathematics teachers at District 2 High School received guidance and resources from the same regional cohort as the middle school. However, there was not much evidence to show that this resulted in a major change to teaching practices with regard to content delivery. Grading was the main focus among teachers during the transition to PBE. They expressed considerable frustration about the new grading practices and believed that the changes led to a lack of accountability among students and negatively impacted student learning. One participant's experience with returning students summarized a common philosophical belief among teachers:

Kids are either dropping out of college or coming back from college and telling us point blank, "I was not ready for college. I was not ready for one attempt to the test, and I was not ready for one test to count 30% of my grade, and I hadn't studied that material. I don't know how to study. I don't know how to cram. I don't know how to really prepare overall. Homework means something in college," one kid said that to me yesterday. One of the alumni from last year came back yesterday, and she said, "I wasn't ready for homework to count." Because in college, homework can be 10, 20, whatever, *percent of your grade, whatever the professor decides. And she's like, "Here, homework didn't count, so no one did it if we could get away without doing it."*

The teachers were frustrated that summative scores comprised a student's entire grade and that homework was not graded. They were also opposed to the idea of unlimited retakes for assessments, which were elements that they believed were implemented by the administration and supported by the regional cohort. In addition, they do not feel that they had a voice in these classroom practices, which fueled a culture of apathy within the school.

Apathy is like the word of the school. It's just complete apathy. They don't give a crap about anything until that final grade. The week before, "How can I pass?" Or the good students, "How can I get a 4?" But it's not until the end. They just want to see that final grade. They don't care about the learning aspect of it at all.

According to participants, there was a level of inconsistency among mathematics teachers with regard to retake practices and the grades that students received after completing a retake. There was also a clear disconnect from students. All these elements are evident from one participant's explanation of his retake process:

Back to your question about if it's the same throughout the math department—no, it's not. I mean I can't even stay consistent, as much as I try, in my own class, let alone be on the same page as everyone else. I mean, if a kid retakes a test three or four times and they're finally at an 80% or an 85%, I'm like, "Yes, you got your 3. Just get out." I know that other teachers are different. We have one math teacher who pretty much requires 100% with the exception of maybe a very tiny mistake, to get your 3. I usually let them have one or two wrong, as long as those one or two aren't the same exact skill over and *over. But there's no set rule for it. Whereas the set rule before was, if you got 8 out of 10 right, you got an 80. If you got 9 out of 10, you got a 90.*

It was clear that the building administration worked with teachers to try to shift their instructional practices from a lecture style to one that involved the students in learning to a greater extent. However, this was not fully accepted by teachers at the school. During both observation visits, the mode and median IPI scores were 4, which indicated that a direct instruction model was in place. The following quotation demonstrates the participant's belief that the only alternative to teaching the entire class was for students to work independently. As a result, she was opposed to her administrator's advice about her teaching style.

Like you said, sort of a lecture at the board, but more just showing them examples, because, trust me, I've been watched, observed by our administrators, and they always say, "Let the kids do more of the work. You're working too hard. Let them learn it." And I will try that, just to appease them. Kids don't learn math from a book. Even my top students do not learn math from a book.

Despite the teachers' hesitancy to change their practices, there was a clear shift in recorded IPI scores between the first and second round of observations. During the first-year visit, the mean IPI score was 3.52, and the percentage of HOTS observed in the classroom was 6%, which are both low. During the second-year visit , the mean IPI score jumped to 3.92 and the percentage of HOTS observed in the classroom increased to 18% (low average). The second focus group conversation shed light on what had changed from the first to the second visit. The school moved away from heterogeneous groupings in their mathematics classes and towards homogenous groupings. It also adopted ability-based classes for each grade level. A greater level of HOTS was observed in the classes that contained advanced students. When asked about the

HOTS opportunities in the classroom, one participant explained that only the advanced classes had access to such opportunities:

Exactly. And so that advanced group, that stuff that I did with them today, I will probably not be doing that with my regular group.

The practice of dividing students into advanced groups was stopped during the initial implementation of PBE, but it was brought back. A teacher described the effect of this change on her precalculus class:

So, I have an advanced precalc versus a regular precalc group. Last year, they got rid of the honor sections. This yea,r they brought them back, but they're not allowing us to call them "honor sections." We're calling them an "advanced section."

The teachers also explained that advanced groups were added for all mathematics subjects at the high school: "this year have advanced groups for every subject." Overall, the teachers were happy with the addition of the advanced groups. They believed that the heterogeneous groupings during the previous year partly explained why they could not provide more advanced opportunities for their students.

Yes. Because last year without them, it was just really, really hard to have that super advanced kid in a group that was mixed and just feeling like I couldn't push forward with the curriculum.

District 3 Middle School Findings. Teachers at District 3 Middle School only participated in the first round of research. Thus, data was collected from one focus group and one round of observations, which yielded 50 data points. The teachers at the school were more positive and seemed to have ownership over the decisions that they made in class. They did not express many major concerns about their own buy-in to a proficiency-based structure. However, they explained that there was not full buy-in from families and that many parents still expressed confusion about grading and curriculum content. Nevertheless, the participants shared an eagerness to collaborate with each other and a willingness to introduce projects in their classes.

Most instructional practices that were in place prior to PBE remained in place in the classroom. When asked about the components that made up her class, one teacher responded as follows:

I'm still using, you know, summative assessments, formative assessments, some homework, participation in class.

Participants felt that the PBE structure allowed them to give students a clearer picture of what they had learned and used the standards to explain why learning was important. The mathematics staff appeared to have bought into the new approach for the most part.

So, I think it has actually shifted for the better. I'm going to repeat myself. It just makes them more aware, and it gives us more purpose for why they have to learn it. Not just because you get an A, B, C, or D. You need this because this is why you need to meet standard on all of these before you send it up.

Despite their positivity about the move to PBE, the participants all said that families did not fully understand it. There was clear confusion about what they were doing and what the Common Core Standards are. In addition, although Common Core served as a foundation for the teachers, they created their own district standards.

It's hard for the parents to understand. One thing, one thing that I have seen with parents especially is that they blame, in math anyway, they blame everything on Common Core. Everything is Common Core. If they would stop teaching Common Core ... well, to have the discussion with them that we're not teaching Common Core. We're *given a set of ... we have written a set of standards for our district, first initiated perhaps by Common Core, but our district has their own set of standards and our own assessments and things like that.*

The mathematics teachers had regular time built into their professional development schedule that allowed them to collaborate. They expressed a desire for more time to gather and discuss teaching and learning.

Now that we ... we do have a math curriculum person, but she also is a teacher and so, we're not able to spend as much time together because we're all just very busy and having the time to get together. I used to enjoy that very much, and we've talked about that, that we need more of that. Not just half workshop days.

The mean IPI score for District 3 Middle School was 3.88, and 24% of observations indicated the use of HOTS in the classrooms. This higher percentage of HOTS appeared to be linked to teachers' willingness to include projects and other hands-on learning opportunities in their lesson plans. In fact, students seemed to be the most engaged when such activities were offered.

For me, mine are more engaged in when we're doing a project, because I tried to give them guidelines but not too many. So, that way they still have some self-choice on either ... like we bought a car. That's one of our projects is they buy a car. I let them choose the car, but I only gave them a certain amount of money. They had to still all have the same graphs and things like that. But you know, they were able to do what they want as far as purchase. So, that was fun for them.

Another participant described how they used architecture to introduce the concept of volume to students. For the assignment, the students were given certain guidelines and had to complete a three-dimensional drawing that fit the requirements:

They had to, for the constraints, they had to have five buildings. The volume had to be over 10 for at least two of them. They had a bunch of criteria, and then they use isometric dot paper, and they had to draw it 3D.

The participants believed that there had to be a balance between the use of such projects and learning the foundational skills necessary to be successful. One teacher explained this thinking and how it related to assignments that allow student choice:

The ratios and proportions one, they had choice assignments. So, they could pick from a bingo board of what they wanted to do. I thought that they would love that, and that was in this big of a hit. Maybe it was because it was ratios and they weren't swayed. But they like the projects typically. I also explained to them, too, that we can't do all projects all the time. If you're going to be ... you got to do the hard work and understand what we're doing first so you can understand how to do the project correctly. So, they do have that understanding that math can't be games and fun all the time.

District 3 High School Findings. The IPI observations revealed that students at District 3 High School received the most HOTS opportunities of any participating school. This was true across all ability levels. Participants went into great detail about their belief in student exploration and shared problem solving strategies. In addition, they explained that the move to PBE impacted how they approached assessment and reassessment, in addition to grading practices. As they became more comfortable with the system, they found ways to more frequently incorporate HOTS in their instruction.

In the past, the teachers at this school had a much more traditional approach to assessment; students had one assessment and were not able to retake it. This changed, and the teachers believed that the change had mixed results. On one hand, it holds students more accountable to learning; on the other hand, there is less accountability in doing the work well the first time.

> *Like we used to be very traditional, where if you took a test and you failed it, you better do more the next time. So, we've made, I think, a good and a bad push towards reassessment. Like today, I had somebody in here really working and working and working, and he learned to factor, where he had a zero before. And that was a great success. I feel like it used to be, "No reassessment no matter what." And now it's, "Reassessment no matter what you've done." So, I think we still have some work to do there. In theory, we talked about, "Oh, they can't take a reassessment unless they've earned it and they've been keeping up," and all those kind of good things, but our grading system fights with that.*

The high school emphasized exceeding the standards. Students who exceeded multiple standards in a course received honors status on their transcript. This is a strategy that the school used to help students see greater value in going above and beyond and that appears to have been made after the initial PBE implementation based on feedback.

> *For their transcript to get, like, with honors distinction for math, they have to exceed the majority of those indicators. They have to have met all of them, and then they have to exceed the majority of them. So, some of my kids who just work medium and medium and happy-go-lucky, when I remind them that at the end of the year and they look, and they've had a nice grade, B, maybe an A-, and they look and then they're like,*

"Oh, I only exceeded two." I'm like, "Well, that's lovely. You have my credit, but you don't get to have honors." So, some of them are like very excited to take a final because … And that never happened before because they have another chance to exceed.

The move to PBE appeared to help teachers be more purposeful in their assessment building. They gave considerable thought to the questions added to student assessments, and the results provided them with specific feedback on students' understanding of each standard being assessed.

I think we have now … We used to teach the unit and then walk over to them, "Do you have a test for systems?" Or you'd just use somebody's test. And now our tests are like, "Oh, I want to see if they can." Like, each test question is just a little bit different. Like, we will make sure that we intentionally put in some for systems or some for substitution in like each different form. Like, every question is on there for an intentional purpose. It's, like, misconceptions that we're looking for, we write a test question for that. It's not just, "How many do I need? Oh, here's five."

During the first round of observations, the school had a mean IPI score of 3.70, and 16% of classroom observations involved the use of HOTS. These increased to a mean IPI score of 4.02 and a HOTS rate of 32% during the second round of observations. The students were observed talking about their thinking and strategies for solving problems. There were also multiple opportunities for students to try solving a problem without an obvious solution or a scripted algorithm. The students would then discuss the methods used in small groups or as a class. When asked about the observed increase in HOTS utilization, participants freely shared their teaching strategies.

Talking it out and having those conversations. That's where the rigor comes in. Just doing stuff on your own. If a teacher is going to run from person to person, that's just, this is how you do this and now you copy me. That's not as rigorous as trying to justify your answer. Figuring out where those connections are or why is this working? Or would this be the same if I got rid of these pieces? They're like no, yes, I don't know. They have to stop and think about that, and that's a level of frustration. The math that thinking about things instead of just always knowing and repeating back.

One of the main observations that participants made was that conversations and thinking were the most important elements for higher-order tasks. Often, teachers moved directly from typical practice problems and traditional learning experiences to a project. They believed that projects often did not increase the level of HOTS at all. Instead, they can be repetitive and do not exceed a standard.

So, I mean, we've heard about project-based learning and project-based assessment and alternative assessments, but the problem is, you either lose some rigor or it feels fluffy. I mean, like, you made a bridge out of Popsicle sticks, you get an A, and you meet an indicator. That's fabulous.

This teacher explained that she used a question or series of questions to elicit estimation, thinking, and analysis from students. These types of questions are accessible to students of all ability levels. In the following quotation, she explains a problem in which her students were asked to determine the length of a car.

And that was the third class we had looked at that question. The first time, they were presented the question and they had to work in their groups and they wanted... The question they had is, okay, well, how long is a car? And I was like, well what do you think

> *is reasonable? Knowing that's what I wanted them to write about, is how reasonable was their solution, which was based on how... For a car length. I have said the average car length. So, I was like, okay, but.... I'm like, yeah, I about 10. And I would stand there, too and their group agreed that 10 feet was reasonable. So, they continue to work and do their task and I'd have to push them later on. Okay. So, you guys thought... I see you follow 10 feet. And I was like, "So, you guys were really thinking bumper to bumper? Like these people are touching? What about spaces between cars?" And then have go and okay... Consider a space. Nope, they're bumper to bumper. In New York, they're literally bumper to bumper. We know. Okay.*

In summary, each of the six participating schools had differing levels of success with regard to the implementation of HOTS in PBE classrooms. If the schools are examined as a group, it can be seen that the teachers made multiple references to factors that were a major focus for them in the move to a proficiency-based instruction model. In turn, these factors drove the inclusion of HOTS opportunities for students.

Attempting to Migrate Away from Standardized Learning and Instruction

Participants in each focus group discussed their attempts to incorporate a PBE system in their classrooms. Although each school and district's focus differed, common themes emerged across the focus groups. This section describes the elements that educators focused on when implementing PBE at their school; these can either enhance or impede HOTS opportunities in the classroom. Thus, the results can be used to inform the third research question.

Grading Practices. Regarding grading practices, the focus group participants shared whether the move to a proficiency-based system impacted the way that they graded students. Many teachers spoke about the specific impact of grading practices on learning outcomes.

However, there was considerable inconsistency in approaches from district to district regarding the impact of a proficiency-based system on the grading of students. The most significant change for teachers was a move from a grading scale of 0 to 100 points to a rubric of 1 to 4. A younger teacher from District 1 Middle School described her experience adapting to the new system in her second year of teaching:

> *My first year, I think it was just the first year was zero to 100. I think last year is the first year we used 1 through 4. I don't really care for it. I understand the idea behind it, but I don't know I'm not a fan of it. I wish there was a way to do some hybrid model.*

A lack of buy-in among teachers for the switch to the 1 to 4 grading rubric was a common theme heard in the focus groups. Teachers were in favor of some components of proficiency, but grading was an obstacle for many. A more experienced teacher from District 1 Middle School shared her negative experience with the new grading rubric:

> *I find the grading boring. One, 2, 3, or 4 over and over again. I don't know. I like having the standards. I think we should have standards and maybe if they're on pace, behind pace, or 1, 2, 3, 4.*

Regarding the grading system, there also seemed to be mixed feelings among teachers and the leadership at schools and within districts. Teachers reported feeling at odds with administrators or others who had advocated for the grading component of the proficiency-based system. This led to some divide in the ranks and mixed opinions with school committee members. One teacher from District 2 High School described what he had heard about the feelings of school board members:

There are school board members, I don't know if it's a majority, but I know a couple school board members personally who want to go back to the old system and who want to have that be part of the superintendent search, finding a superintendent that's going to support going back. Now I don't know if that's a majority, and I don't know if that's the guy or girl we're going to end up hiring, but I've heard that straight from their mouth from certain school board members that they do want to go back to the old system. It's just a matter of getting someone in charge that would be in favor of it.

Many teachers expressed an interest in reverting to the traditional grading scale of 0 to 100 points. They believed that this was still possible and that they could retain common standards. Moreover, many participants said that the 1–4 rubric did not allow them to reward students who truly excelled in class, whereas the 0–100 grading scale enabled them to differentiate students from each other. When participants were asked how they would change the grading system if they could, one teacher from District 2 High School advocated for returning to a traditional grading scale:

Well, step one, As, Bs, Cs, and Ds. Step two, I would talk to the staff. It doesn't have to be a dictatorship, but if we want to continue to give retakes, how will we average scores together? You got a 50 the first time, study hard, get a 90 the second time, you get a 70. You know what I mean? Not like, "Oh, you got a 90 so you get the same grade as the guy that got the 90 the first time." But you don't have that freedom in the 1 to 4 system.

Inconsistency was also a theme in each of the focus groups. Inconsistencies were seen from school to school and teacher to teacher. Many teachers still used some form of grade conversion, which means that they scored students on a 0–100 grading scale and then converted their grades using the 1–4 rubric. However, this was not an exact science and could have led to an implementation of proficiency-based grading that lacked fidelity. A teacher from District 2 High School spoke about the inconsistencies that she observed in her classroom:

> *Back to your question about if it's the same throughout the math department, no, it's not. I mean I can't even stay consistent, as much as I try, in my own class, let alone be on the same page as everyone else. I mean if a kid retakes a test three or four times and they're finally at an 80% or an 85%, I'm like, "Yes, you got your 3. Just get out."*

Assessment. Participants frequently mentioned their assessment procedures and whether they were impacted by the move to a proficiency-based system. When asked about their assessment structure, one teacher from District 1 High School succinctly explained the purpose of formative assessments in the classroom and added that the change in structure led to fewer homework assignments and a looser teacher unit structure:

> *So, it's almost more of the standard sort of thing. This formative assessment, the homework you're talking about, is informing them and me if it's working. So, I'm way less structured than I used to be. I may put in a half a dozen homeworks now a quarter.*

At the high school level, there was an impact on the traditional structure that once included midterms and final exams. Many teachers no longer give these because students would have already demonstrated proficiency on an earlier assessment. As a result, summative assessments focused more on a specific standard than the culmination of multiple units. A high school teacher from District 2 High School described this situation as follows:

> *So, it's like when the learning target's all wrapped up, you get a 1 through 4 score on that particular learning target, but that takes away our ability to test,*

and retest, and mini-test, and all that type of stuff, so finals are out the window because they've already tested on something, and I can't retest them on something like that in a final setting.

Retake tests were a point of frustration for teachers. Many felt that students learned just enough material to pass the retake tests to allow them to move on to the next standard. Participants also believed that the retake structure impacted the need for students to prepare for assessments themselves. Many teachers did not have banks of different questions for retake assessments; thus, students could focus on gaining a partial understanding of the concept and still attain proficiency on the make-up test. A teacher from District 2 High School described her impression of some students' attitude towards the make-up assessments:

I can ask the teacher a bunch of pointed questions before I take a retake, and then I can learn just the bare minimum, and then I'm not going to retain any of that for the next lesson.

However, make-up assessments were not viewed as a challenge at all schools. At District 3 High School, teachers developed an accountability system that they believed helped students study rigorously before the first time that they took a test to avoid a retake. According to this accountability system, students would receive a lower grade if they needed to retake a test. Although this practice may not fully align with the proficiency-based grading protocol, the teachers believed that it had a positive impact on student success. A teacher from District 3 High School explained how students perceive the accountability system:

The other thing too for assessments. They know if they do not meet the assessment the first time, they have to have all of their homework done in order to reassess. Because *how do I know what you practiced was good practice? So, there's that chance. What do you want to gamble? Not do it, and oh, you just made it? Or gamble and crap.*

Teachers at District 1 High School took a similar approach. Their retake policy only allowed a certain grade attainment on the retake. This emphasized the importance of the first assessment for students. The teachers believed that this represented a favorable change for the latter.

We have a new policy where if kids don't pass a standard, they can retake it. But they can only get a 76 on it. Whereas before, they could retake it 10 times if they wanted to get 100 on it. So, it's made them a little bit more aware of being serious the first time kind of thing. So, that's helped, I think, a lot in my room.

At District 1 Middle School, participants shared that they had developed a common assessment system and common classwork that were used by all teachers and designed to create a uniform experience for students, regardless of their class. One participant explained the expectation that common assessments would be used throughout the mathematics classes:

We use the exact same things, and the other grades are supposed to be doing that as well, having common assessments. We have a document that has all of our tests in it.

Upon further probing, it seemed as though the consistencies that existed were with teachers who felt comfortable collaborating with each other. Some collaborative pairs shared materials; such collaborations were more present within grade levels. For example, a sixth grade teacher at District 1 Middle School explained that he and his grade-level colleague used the same assessments:

Yeah, sixth grade, myself and [a fellow teacher], we use the exact same homework, the exact same tests, we just think that keeps it kind of fair.

When participants were asked about the number of assessments given per unit, the responses were mixed. The main inconsistency seemed to concern the number of formative assessments that were used. In addition, there did not seem to be a common definition of a formative assessment in the system. The following quotations represent responses from three teachers at the same school (i.e., District 1 Middle School) to a question about the number of formative assessments that they typically gave within a unit.

> *I like to have at least four formative and at least one summative, minimum if not more. It depends on the standard, because some standards have a lot more chapters in [them] than others. Sometimes, they don't get their formatives in before they take the summative, so sometimes I don't know.*

This was the inconsistency observed at the school that had fully adopted a common assessment system. The other schools had not yet developed common assessments.

Standards. Each focus group discussed how standards drove instruction. In fact, having clear standards that students must achieve was the most consistent component mentioned by participants at all six schools. Participants described the origins of their own standards and were in many cases part of the development process in some way. The focus group from District 2 High School discussed the standard-writing process and how the standards evolved.

> *We were part of the writing team during that cohort process. And so, I had all those files from what we worked on in the past, and I took those, and we kind of said, "This hasn't worked. This would probably work better after experiencing a year or two with that," and we've been redefining them ourselves for next year.*

Participants from the above school were in the district that worked the MCCL; thus, their school initially had specific standards, then adapted them for both the middle school and high school. This was the only focus group that did not mention using the Common Core Mathematics Standards as a starting point, although this does not necessarily mean that they were not referenced at some point. Teachers at the middle school and high school in District 2 appear to have implemented cohort-designed standards, which one high school teacher explained as follows:

> *So, the cohort of schools that developed those learning targets initially, we took that and said, "Okay, now make it work for us." And so we took that as a starting point.*

Although District 2 Middle School initially implemented standards based on the cohort's recommendations, the teachers and administrators did not see the achievements that they hoped for. As a result, they shifted to a mathematics curriculum aligned with Common Core. During the second round of focus groups, a teacher from District 2 Middle School explained this curriculum choice, which had been implemented before the second round of observations and focus groups:

> *enVision is aligned to Common Core and our targets in Empower, and this way we start to see that we really need to do some shifting. And we did it last spring in the vertical. So, right now I'm teaching what I used to teach in fifth grade years ago and what used to be sixth grade. So, we're really shifting.*

Thus, teachers in District 2 have begun to change their way of thinking and are working on raising the bar for mathematics students. Instead of making standards easier when students cannot complete the work at the current grade level, they use the standards to shift their curriculum and practices.

> *Well, at one point we had control over our curriculum, and we would say, "Okay, well, here is the learning target. Oh no, our sixth graders can't do that." So, we put it in seventh grade. Whereas we should have been saying, "Our sixth graders will do*

math." So, yeah, it will be interesting to see a couple of years down the road. (District 2 Middle School)

In District 1, teachers were also part of the development process for the mathematics common standards. Participants explained that they began with the Common Core Mathematics Standards and developed their own power standards from there. This ensured appropriate grade level content alignment and a focus on content that would prepare students for the next level. A teacher from District 1 High School shared their experience with the process:

> *We took the performance indicators straight from Common Core and then developed power standards, or we call them "graduation standards."*

Middle school teachers in the same district purchased materials that aligned with their standards and the Common Core Standards. These are the same teachers who spoke about their common assessment system. At the administrative level, there has been a push to ensure that standards are aligned and that teachers are assessing students on these standards. When asked about the Common Core, a teacher from District 1 Middle School explained how their resources are aligned:

> *We have Common Core. Yeah, and the textbook, it aligns the sections with the Common Core.*

District 3 also aligned its standards with the Common Core, and teachers in this district were part of the standard-creating process. However, participants explained that, although they used the Common Core as a reference, their district created its own set of unique mathematics standards.

We're given a set of ... we have written a set of standards for our district, first initiated perhaps by Common Core, but our district has their own set of standards and our own assessments and things like that. (District 3 Middle School)

The level of flexibility at the school level seems to be connected more at the indicator level, where high schools have been able to customize to ensure its appropriateness for graduation requirements. This seems to have had an impact on the student experience.

We have a certain number of indicators under every standard. To graduate, they only need a majority of each of those indicators to just graduate. So, like you teach geometry, which is done all in that one year. Once they get their majority, they technically have enough geometry. So, there's this, like, "What do you do?" (District 3 High School)

Student Supports. Regarding the successful implementation of proficiency-based structures, many participants referenced student supports that have been established. These supports include ways to provide attention to students who are working individually or help them meet a specific standard. One example is time made available to students outside of regular school hours. Teachers at District 1 Middle School said that they provided opportunities for students to stay after school for additional support.

I would say per grade level, there's probably at least one grade-level teacher available for kids to stay afterwards every day.

Many schools introduced web-based resources to help students focus on their individual skill deficits. For example, District 1 High School instituted a personalized, web-based learning system called IXL to facilitate the assessment of students and determine their current skill level. The program then provides lessons and practice exercises at the student's own level. Teachers have found ways to integrate such programs into regular class time. One participant explained how IXL was used in her class:

> *So, everyone's doing something different on IXL. It's not related to the specific class. They had their own tracking sheets based on diagnostics.*

The program is also used in other districts. Participants greatly valued IXL as a supplement to their own instruction. Some teachers wanted to devote extra time to practice skills, and textbooks did not always provide enough practice problems to give students this opportunity. A teacher from District 2 Middle School explained that this was one reason why she valued IXL:

> *We're fortunate enough to still have IXL as well. In my book, there's not a lot of rote practice. There's a few problems and then they go into applications of that. And they need to have that rote practice as well, and that's instant feedback, and we can hang onto IXL as well. That's good.*

Due to the nature of the proficiency-based system, students often demonstrate proficiency on a standard ahead of their peers. Thus, teachers must find additional work for them. If students are not allowed to move ahead to the next standard or work on an assignment that requires greater depth of knowledge, they are often recruited to support their peers, which adds an additional layer of support for lower-achieving students in the class. A participant from District 3 Middle School described how this process worked in his class:

> *If they don't test out, they wouldn't have time. If they do, they might finish a lesson early, but they're either a peer tutor at that point or where it's me, them trying it, and then with a peer, and then them doing on their own.*

Individual Pace Learning. Individual pace learning and moving at one's own student pace are frequently mentioned components of a proficiency-based system. In this research, individual pace learning was found to be more prevalent at middle schools than at high schools. Teachers who adopted all aspects of a PBE system asked students to work individually in a flipped classroom model while they played the role of a tutor and made their way around the class. In such classrooms, students moved at their own pace rather than teacher pace, and the teacher focused more on individual or small group instruction rather than whole-group instruction. One participant from District 1 Middle School explained how the independent structure began to take shape at her school:

> *What I see is they transition from the primary school, with the fifth grade, which is pretty structured, to more independent responsibility with the sixth grade, with the flipped classroom. And then as they go into seventh and eighth, they're starting to build even more independence and responsibility for their own work.*

By contrast, high school teachers did not fully buy into the idea that students should move at their own pace. Many believed that students were not ready to do so or would not take enough accountability for their own learning. When asked why they had not fully adopted an individualized structure, one teacher from District 1 High School responded as follows:

> *You're putting a lot on them, learning on their own, and some kids aren't made for that or disciplined enough for that. That's tough to do. It's too easy for some of them to sit in a corner and just be quiet.*

This was a relatively common remark among the high school teachers who participated in the study. They seemed to feel that students needed extrinsic motivators to complete the assigned work. A teacher from District 3 High School explained their rationale for keeping the students together in the following way:

> *We're pretty well all staying in the same, but I think the knowledge that they can reassess later makes that a little different. It seems okay. We have tried to stay away from the "everybody doing something different." That still gives me a lot of cause for worry. When you don't have a lot of self-motivated students, that can be disastrous.*

However, some high school teachers implemented selected tenets of an individual pace structure. One participant from District 2 High School noticed that students did not complete their classwork in a timely manner and believed that they procrastinated until the last possible minute, which had a negative impact on their learning:

> *That's the way they interpret it, at least, here in this culture, they say, "Oh, since I can work at my own pace, that means I can put this off for three weeks," or whatever the case is. So, instead of actually working at a pace that's reasonable for them, they're not doing it at all until they absolutely have to.*

Some teachers kept the students together as a group, then integrated individual pace time at the end of class once the group work had been completed. This enabled students to focus on individual skill building while keeping the class together to complete the assigned unit work. A teacher from District 3 Middle School said that she used a specific program that students can access to work at their own pace at the end of class:

> *There's not a whole lot of time for them to finish early. But if they do, then I put them on Freckle. And that's at your pace.*

Some teachers used a flipped classroom model, but it seemed to be individualized. At District 1 Middle School, teachers used videos created by a high school teacher in the district to instruct students in upper-level classes. This teacher was the only teacher in the high school with the flipped model, but this work seemed to align well with the structure that the middle school was using.

> *We have really good support from the high school with that, the high school teacher who teaches algebra, [teacher's name], he has sort of flipped classroom. And he has a great website with all of his videos on it. And has the assignments listed, has the answers, which is a good thing. So, as they work through their homework, they can check their progress and they can watch the videos.* (District 1 Middle School)

The Havens and Hazards of Proficiency-Based Education

Throughout the focus groups, participants were able to articulate the positive outcomes and challenges that they associated with the implementation of a PBE system. Within these havens and hazards, educators shared their concerns regarding student learning outcomes and their feelings on how the change had benefited them. There were many differences in opinion throughout the focus group conversations, but common themes emerged. The havens and hazards inform the third research question, as they are factors that enhance or impede HOTS in the PBE classroom.

The Haven of Increased Clarity/Communication Impacting Learning. A clear advantage associated with a shift to a proficiency-based structure is increased clarity regarding standards and learning expectations for students. Participants explained that they often played a role in the creation of standards, which clearly communicated to students what they needed to achieve to attain proficiency. Although clear learning expectations already existed in many classrooms, this has helped teachers develop a common framework for communication. A teacher from District 2 Middle School explained how the move to PBE prompted her to communicate more with her students about the standards that they were working on:

> *I definitely communicate more with the kids about what we're doing and what the expectations are. What the standard is and why we are doing it. Why we need it now. I'm definitely having more conversations with kids just based on we had it all written out of what needed to do to be able to succeed.*

The PBE system prompted teachers to explain to students exactly what they needed to do to attain proficiency on a standard, which contrasted with students missing questions on an exam and moving on to the next unit. Students now had an opportunity to slow down and truly understand the missed content. A participant from District 3 High School explained how this has led to greater clarity for her students:

> *Honestly, I feel like that's been a positive, too, is that the kids who are struggling where they know that they can focus on Level 3, so as opposed to in a traditional-style class where there are just questions that they don't get and so they wind up with 80% or whatnot. If they can show that they meet ... Like separating that this is what you need to know, this is an extension would-be-nice-to-know stuff. And I think they feel less bad about, like, missing questions because it's labeled out for them. These are hard questions that not everyone's supposed to get.*

The impact of switching to a proficiency-based system on students is difficult to quantify. No common responses emerged from participants in the focus groups. However, some teachers believed that the PBE system was beneficial for students. Although there may not have been enough data to determine a clear correlation, one educator from District 1 High School felt that the move to PBE may have had an impact on students' skill base. Every year, students take an adaptive assessment in the fall and the spring. According to the participant, students' performance on this assessment had improved since the implementation of PBE:

> *In the last five years in the NWEA testing in the fall and the spring, I've seen huge increases. Individually because it's based on their growth from fall to spring. I've seen much higher increases in my kids with standards than before. So, I really think that that shows something.*

Participants explained that the PBE system tended to work well for some of their students, but it had been a struggle for others. Notably, teachers felt that some of their classes worked well overall within the system, and others encountered challenges as a group. A participant from District 1 Middle School remarked that the students who have excelled the most in the PBE system were the ones who were passionate about mathematics:

> *It depends on the student. It depends on the class. I have some classes that are really excited to be there, like math is their favorite. The others are much more ELA-focused, and they'd rather be reading than solving algebra problems.*

The Haven of Relationship Building and Student Engagement. The move to a proficiency-based model underscored the need for individualized student support. Participants frequently discussed how they implemented structures in their classroom to support students and ensure that they all attain proficiency. Individualization was connected with an increase in opportunities for teachers to build relationships with students. One teacher from District 1 High School experienced this first-hand in his classroom:

> *Yeah. Well, I like to be able to walk around too, because it gives me a chance to say stuff like, "Good job in the basketball game." But then I will notice problems that they are not doing well on. So, you kind of know them and all that stuff. And In addition, when*

I go over the homework it gives me a chance to ask them questions about it too, and say, "How'd you guys do this?" And so, I can get some feel-good with that.

Teachers in many of the focus groups mentioned the value of relationships. They believed that relationships were a factor in whether students succeeded or struggled in their classrooms. A participant from District 3 High School said that relationships had an impact on what a student will attempt for problems and how deep or how far they can be pushed in their learning:

So, I don't know. I feel like some kids, you have a strong enough relationship that you can push them, and push them, and push them to go a little deeper. And other kids while you're still working this relationship, can't quite get their state because it starts to crumble a little.

Participants were specifically asked about student engagement and when they believed that students were the most engaged. There were multiple responses, including comments about the time of day or when students worked on specific assignments. However, there was no consensus about engagement and the implementation of a PBE system. However, some respondents expressed that students were engaged when they were quiet and had mathematics problems in front of them. Many participants seemed to believe that students doing what they were asked to do constituted engagement. Thus, the definition of engagement was open for interpretation. One teacher from District 1 Middle School believed that the PBE system had helped with engagement in her classroom:

I feel like they are more engaged than previous years because where they're doing their own thing, it's more time for me to walk around to make sure they are at least pretending *to be engaged. Whereas before, when I would teach up at the board, somebody could not be engaged and I might not really be able to tell.*

At District 3 Middle School, a participant explained that her students were the most engaged when they were working on a project. Although projects were not frequent, this teacher still found opportunities for students to pursue them in the proficiency-based structure. She believed that offering individual choice and broad guidelines rather than a much more prescriptive assignment increased engagement levels among students:

For me, mine are more engaged when we're doing a project, because I tried to give them guidelines but not too many. So, that way they still have some self-choice on either ... like we bought a car. That's one of our projects is they buy a car. I let them choose the car, but I only gave them a certain amount of money. They had to still all have the same graphs and things like that. But you know, they were able to do what they want as far as purchase. So, that was fun for them.

The Haven of Teacher-Directed Instruction. Despite the push for a more individualized classroom instruction structure in a PBE system, many teachers chose to continue with whole-class, teacher-directed instruction. The latter was prevalent in five out of six participating schools. Teachers believed that whole-group instruction helped to ensure that students received appropriate instruction to attain proficiency. A participant from District 1 High School explained that keeping students together created a bit of extra time at the end of the unit, but he still believed it to be the best method:

I keep my kids kind of on the same pace and there's always some lag a little bit, but before we take the next summative, everybody has become proficient on the last one.

In addition, the school administration requested more of an individualized approach in the classroom, which meant that students would move at a student pace. However, teachers have not found a way to effectively teach mathematics content under this structure. They said that the use of examples was key to their instruction and that a whole-class approach was more conducive to this method. When a teacher from District 2 High School was asked about his class structure, he responded as follows:

> *Like you said, sort of a lecture at the board, but more just showing them examples, because, trust me, I've been watched, observed by our administrators, and they always say, "Let the kids do more of the work. You're working too hard. Let them learn it," and I will try that, just to appease them. Kids don't learn math from a book. Even my top students do not learn math from a book.*

At the high school level, participants mentioned the importance of whole-group conversations, which help students better understand the content. A whole-group structure is important to allow teachers to create opportunities for these conversations. A participant from District 3 High School shared her views on this issue:

> *As much as it's not in vogue right now, most kids that I have in, like, primarily direct-instruction classes, I think want the instruction time. And if everyone's in a different place, then you couldn't really, like, what are you going to talk about?*

When asked about levels of engagement in the classroom, participants said that students were often focused when the instructor presented to them. Teachers also appreciated knowing that students were concentrated on the content at hand. When they worked individually, it was sometimes difficult for teachers to know what was on their computer screens. According to a participant from District 2 High School, his students were most engaged when he instructed the entire group:

> *Honestly, I think they're more engaged when I'm doing problems up on the board. As much as I hate to say that, and as much as all this new research says kids want to work in groups, and they don't want the teacher teaching them, I don't see that at all.*

The Hazard of Negative Student Outcomes. The move to a proficiency-based system is rooted in the belief that the change will have a positive impact on student learning. However, many participants said that challenges associated with this shift impacted students' ability to learn optimally. The most common concern was student preparation and performance on assessments, as students prioritized passing rather than excelling on the latter. As a result, participants believed that students took shortcuts that they would not have taken in the past. A teacher from District 1 High School cited the example of memorizing information to pass the test:

> *It may hold them accountable for that standard, but let's face it, I think we all can admit that I can give a test to kids today and next Friday the results would be very, very different. In other words, they got the skill for now. But do they really have it? You know what I mean?*

As previously mentioned, the assessment retakes were another point of contention. Retakes are intended to allow students to learn content at a comfortable pace. Additionally, they are based on the idea that students should not be penalized if they take longer to understand a concept than their peers; the point is that they understand the concept. However, many participants believed that a PBE system led to students not working as hard to prepare for the assessment the first time around. This in turn impacted their long-term progress, as one teacher from District 1 High School explained:

> *They don't do well, they can retake. So, especially the smart lazy kid knows that I'll just take a look at the test, and if I don't like it, I'll just redo it anyway. They don't have the pressure on them to do well as they will when they do go to school. So, I think it's some ... They learned, but I think they've learned something negative.*

Participants at District 2 High School described this phenomenon as a level of apathy from students. Theyd believe that students did not care as much when they knew that they could retake assessments. According to one educator, many students were unprepared for assessments as a result and began to care more when it was time to assign final grades:

> *Apathy is like the word of the school. It's just complete apathy. They don't give a crap about anything until that final grade. The week before, "How can I pass?" Or the good students, "How can I get a 4?" But it's not until the end. They just want to see that final grade. They don't care about the learning aspect of it at all.*

In a more traditional system, teachers give multiple quizzes to students before the final unit test. The quizzes allow students to study a smaller amount of content at a time and are scored as part of their final quarter grade. In a PBE system, quizzes are replaced with formative assessments that do not count or count minimally towards a summative score. As a result, participants believed that students did not take the formative assessments as seriously as they did in the past. This in turn led to greater challenges on the summative assessments, as described by a teacher from District 2 High School:

They're not stupid. Now, the kids that don't remember anything for that final test, they're screwed because they don't have a couple of those good quiz grades along the way. It's just not a good system.

Participants also felt that students were not as impacted by missing or low grades as they were in the past. Teachers frequently assigned a zero for a missing assignment, which would dramatically lower a student's grade. In response, students checked in with the them to try to increase their grade or discuss the assignment. The same was true when students obtained a low quiz score; they worked harder on the next assessment to increase their overall grade. Since students now understand that summative assessments make all the difference and can be retaken, these strategies no longer have the same impact. A teacher from District 2 High School summarized the situation as follows:

And you can even tell them, "Oh, I'm going to put this quiz as, like, a holding grade until the test," "But in the end, it's going to get replaced?" "Yes." Then they're like, "Oh, okay." It would help them learn the math and retain the math, but they don't care about that. They care about that grade, and they know it's not going to be part of that final grade.

Many teachers struggled with the independent component of the proficiency-based structure, as they felt that students learned best when they instructed them. When asked about opportunities for students to work independently at their own pace, many participants said that this approach did not tend to work for them. At District 1 Middle School, one educator believed that her students needed less independence and learned best when she introduced the material to them. She referenced that the system may have hindered her being there for the students as much as she would have liked.

Well, I think they need less independence. I think they need me way more than what I'm there for.

The concept of individual student pace was not fully supported by participants. As mentioned earlier, teachers tend to believe that students frequently do not work optimally when left to their own devices, which has an impact on the amount of content that they can master over a school year. A teacher from District 2 High School expressed the following:

That's the way they interpret it, at least, here in this culture, they say, "Oh, since I can work at my own pace, that means I can put this off for three weeks," or whatever the case is. So, instead of actually working at a pace that's reasonable for them, they're not doing it at all until they absolutely have to.

Many teachers share the belief that a misuse of the system's intentions has had a negative impact on student learning. They felt that formative assessments, practice problems, summative assessments, and retakes were not being used as they should be, which has led to a lack of effort from students. A participant from District 2 High School felt that only a small percentage of students appropriately used the system:

We have two or three kids in the whole school that I can think of that actually use the system as intended and are decent students, and try their best the first time. As it is written in paper, I can think of one or two that really appreciate the system for what it is, but the rest of them just abuse it.

Other educators shared a concern that the PBE system was grounded in the belief that most students are intrinsically motivated. An intrinsically motivated student would carefully prepare for every assessment and see the value of learning for the sake of improving at mathematics. One teacher from District 3 High School felt that intrinsically motivated students were now in the minority. In other words, students who wanted to do their best felt a bit out of place in front of their peers because a critical mass of students did not work optimally in the PBE system:

> *We went through a period of time here where we had a core group of kids I think that were very, very intrinsically motivated. And they kind of pulled the school along and I think we've lost some of that. The ones who are really motivated now, they're not the leaders.*

In addition, high school teachers were concerned that the proficiency-based model set too low of a bar for students and would clearly contradict the expectations that students would encounter upon entering college. The concepts of exam retakes and moving at an individual pace are still not widely accepted in the postsecondary world. If students expect to receive second chances on all exams, this may impact their adjustment to a postsecondary setting. A teacher from District 2 High School shared that students had already told him that they were not prepared for college:

> *Kids are either dropping out of college or coming back from college and telling us point blank, "I was not ready for college. I was not ready for one attempt to the test, and I was not ready for one test to count 30% of my grade, and I hadn't studied that material."*

Moving to a PBE system represents a significant change, which raises the question of how students feel about it. This would undoubtedly differ from school to school, as there has been considerable inconsistency in approaches to PBE implementation. At District 2 High School, it was clear that mathematics teachers did not support the change. Focus group participants shared the results of a student survey at their, which showed that students did not see PBE as beneficial.

They did a survey about PBL and stuff, and I can't remember every stat, but I remember a couple. Ninety-five percent [of students] didn't think they were getting prepared for college, and like 92% hated the grading system. So, I mean it was overwhelming among the student body. Those two things, the grading system sucks and it's not preparing us for college because in college we're not going to get retakes and professors are going to give us deadlines. It was resounding, yet no one listens.

The Hazard of Failing to Build Support for PBE. A challenge that was frequently mentioned by participants was the difficulty of building support for PBE. Parents did not fully understand the system; the same was also true for many teachers. Instead of fully shifting to a proficiency-based grading system, many teachers chose to adapt their current system in a way that still allowed them to maintain many past practices. These implementation methods have led to confusion for families, as what they see in one class may be very different in other classes. To address this confusion, some teachers intentionally changed their grading approach to one that does not implement PBE with complete fidelity. For example, a participant from District 1 Middle School explained how she changed report cards for students who were still trying to meet a standard:

Yep, 'cause I was talking to Miss [teacher's name] about that and she was saying the same thing. She wanted to give all these kids 2s, trimester two because they're still working on that standard, but she thought parents and maybe even administration might be like, "Why is everybody getting a 2?" and not understand it.

Many teachers mentioned that parents simply did not understand the concept of assigning 1 or 2 as a grade if students had yet to attain proficiency. They wanted their children to do their best and felt that they should be working harder if they received a 2. The difficult part for teachers is that they may have only given students the ability to attain a 2 based on the work that has been presented thus far. Thus, communication was much more important for teachers in a PBE system. A participant from District 1 Middle School said that parents had difficulty understanding the proficiency-based system and why the school has invested significant effort into making this change:

> *Because the parents don't get it. We sometimes, we're having parents who don't understand the work.*

This was a common sentiment across the participating schools. Another concern was that there was a general misunderstanding of the Common Core Standards for Mathematics among parents. Many parents believed that "Common Core" was another term for mathematics being taught differently than they learned it growing up. This leads to lack of buy-in from parents when teachers explained that their standards were rooted in the Common Core. A teacher at District 3 Middle School said that many parents linked PBE with the Common Core, which made it difficult for them to support the system:

> *It's hard for the parents to understand. One thing that I have seen with parents especially is that they blame, in math anyway, they blame everything on Common Core.*

In addition, there was widespread confusion among students. Students in both middle school and secondary school experienced the adoption of the PBE model and the transformation of the grading system. The goal of a PBE model is for students to attain proficiency on content area standards. A teacher from District 1 Middle School explained that students struggled to understand what was considered "passing" in the proficiency-based system:

> *It's tough for the kids. They're used to it by the time they get to eighth grade, but where I teach seventh and eighth grade, those seventh graders, they have no idea*

what passing is. And they still ask, "Is this a passing grade?" We're almost to February now, we're a semester through, and they're still unsure. They're just so used to from kindergarten through sixth grade, that 1 through 4, that would be a 65 or an 82, both, "Is this passing?"

Over time, teachers have adjusted their own practices within the PBE model to foster greater accountability. Although they understood that this was not a faithful implementation of the model, they felt that it was best for their students. For example, District 3 High School imposed restrictions on when a summative reassessment could happen. A teacher from the school described the changes that had been made and the fact that they were at odds with the grading system:

So, I think we still have some work to do there. In theory, we talked about, "Oh, they can't take a reassessment unless they've earned it and they've been keeping up," and all those kind of good things, but our grading system fights with that.

In other cases, teachers did not attempt to switch to a PBE structure first. Instead, they continued to use methods that had been used in the past, then converted results to align with the system that they were required to report in. For these teachers, little changed in terms of the design of their assessments. For example, a participant from District 2 High School explained his continued use of the grading structure that he had used in the past:

I don't even quiz them on Level 2 basics, so to speak. So, when I give them a test, and it's all the advanced material to begin with, I just in my own way give them a percentage grade the traditional way I would, and say, oh, you got 90, that's 3.5, or you got a 95, that's a 4. You got an 80, that's a 3.

In addition, participants noted that teachers within the same school often took a very different approach to grading. In one class, students may have been required to obtain 100% to attain proficiency. However, teachers in other classes may have required a lower percentage of correct answers for students to attain proficiency. This discrepancy indicates a lack of understanding about how to implement the PBE model for mathematics assessments. A participant at District 2 High School said the following:

> *I know that other teachers are different. We have one math teacher who pretty much requires 100% with the exception of maybe a very tiny mistake, to get your 3. I usually let them have one or two wrong, as long as those one or two aren't the same exact skill over and over. But there's no set rule for it. Whereas the set rule before was if you got eight out of 10 right, you got an 80. If you got nine out of 10, you got a 90.*

Other teachers did not seem passionate about teaching in a PBE system. In a traditional system, they had much more leeway for devising their own grading structures. Certain questions counted for a specific percentage of a student's grade, and the course grade consisted of various components, such as tests, quizzes, homework, and projects. Under a PBE system, a student's was determined by their performance on summative assessments; often, the best that a student could achieve was a 3. One teacher from District 1 Middle School summarized her frustrations with the grading structure:

> *Honestly, I don't know. I find the grading boring. One, 2, 3, or 4 over and over again. I don't know. I like having the standards. I think we should have standards and maybe if they're on pace, behind pace, or 1, 2, 3, 4.*

In some districts, the move to a PBE model was accompanied by a new mathematics curriculum, which was met with support or hesitancy depending on the type of consensus-building activities that took place prior to its implementation. At District 1 Middle School, a curriculum coordinator chose a mathematics program and presented it to the teachers. The latter viewed it as the wrong fit and expressed a concern that there was a disconnect between the program and what they saw in the classroom. In the following quotation, an educator from the school described her experience with the presentation of the curriculum, which her students attended:

> *Not for the way that we felt she was describing it, how she wanted us to approach some things. I can't remember exactly what it was that fifth grade, when they did a topic, she had taught with the implied knowledge that they knew rounding decimals or something they hadn't even learned yet. The kids were kind of like what?*

Once the curriculum was adopted, another concern among teachers was a lack of support and appropriate resources to implement it. This is a real concern in a PBE model, as students' progression through the standards is very important. Participants said that the order was not correct in the classroom resources, which created additional work for them. One teacher at District 1 Middle School said the following:

> *Gosh, I don't know. It's just ... Our sixth grade book, it's like I told you, the sections line up with Common Core standards, but it's 1.1 and 9.3 go with a standard. One of them, they start in Trimester 1 and they don't finish that standard until Trimester 3.*

Some participants felt that the old system worked for them; thus, they had difficulty understanding why a change was needed. Such responses indicate whether there was staff buy-in at the school prior to PBE implementation. A teacher from District 2 High School explained why there should not have been a change:

You don't need to do research on something that's been done for 150, 200 years. It's already been done. You know what I mean? The research is the history. It worked for 200 years. We don't need to do more research on it. It worked.

That same teacher expressed that the proficiency-based model may work for other content areas, but it was not appropriate for math. His reasoning was based on the typical grading structure in a mathematics class; scores are based on percentages, and a 1 to 4 grading rubric did not accurately represent student knowledge. He said,

Give me a poster on World War II, and if the poster has this, this, and this, it's a 4, and if it's lacking something, it's 3.5. I could do that, but math is not like that. It's about giving the teacher the right answers, and what percentage do you need to give a teacher to get the 3. You know? It hasn't been defined, and this system doesn't make sense for math people.

The Hazard of the Paradigm Shift to PBE. During the shift to PBE, districts developed implementation plans that made the most sense for them. Implementation often began with lower grade levels, then progressed to the middle and secondary school levels. An important component of the implementation process was the provision of appropriate training and resources for teachers. A concern that emerged among teachers was the unequal allocation of resources and supports. A participant from District 1 Middle School described the frustration at their school:

Like [teacher's name] said, the administration has really put the focus on the fifth and sixth graders. Yeah, that's going to help them then. It's not giving them the independence that they need when they get to seventh grade. And then we've pulled all that support from them, and they're just, they're floundering, some of them.

The number of content standards assessed in each grade has an impact on the depth of understanding with which content can be covered. The proficiency-based diploma holds all to a new standard—namely, to ensure that all students demonstrate proficiency on the standards needed to graduate or move on. Therefore, if a teacher has slowed down to spend more time on a unit in the past, this may not be possible in the new model. An educator at District 2 Middle School explained his decision to sacrifice depth to ensure that specific learning targets were met:

> *My plan is to get through, because I think it's easier for us to, with the lower levels of the program than it is with [teacher name's] class. That's tough when you're starting out assuming they have this background stuff that they don't, as eighth graders. So, my plan is to get through all of our targets, not in maybe the depth that we've done fractions, decimals, percents, and that type of stuff.*

The philosophical change required by assessments was one of the more difficult shifts for high school teachers to adjust to. Most participants had not allowed reassessments in the past, let alone multiple reassessments. This led to challenges in logistics and teaching strategies, which were not small changes for high school mathematics teachers. A participant from District 3 High School described these changes as a completely different way of thinking:

> *I feel like it used to be, "No reassessment no matter what." And now it's, "Reassessment no matter what you've done."*

Another paradigm shift associated with assessments was teachers' ability to reassess at strategic points over the course of the school year. In the past, students may have been assessed on certain material on a quiz, a test, a midterm, and a final. Under the current system, the general interpretation was that students were not assessed again once they had demonstrated proficiency. Participants believed that this impacted a student's long-term understanding of the content, as they only memorized it for an upcoming assessment. This situation was summarized by a teacher from District 2 High School:

> *So, it's like when the learning target's all wrapped up, you get a 1 through 4 score on that particular learning target, but that takes away our ability to test, and retest, and mini-test, and all that type of stuff, so finals are out the window, because they've already tested on something, and I can't retest them on something like that in a final setting.*

In addition, participants reported that finding a good place for students to take make-up assessments was a challenge. In a traditional model, all students work quietly during a test and the instructor monitors them throughout the process. Now, if students need to retake a test several times, they may need to find a corner of the classroom or a neighboring classroom in which to work. This impacts classroom management, as the teacher must supervise many students who are working on different assignments at the same time. A participant from District 3 High School described how she attempted to send students to a neighboring teacher's classroom for a reassessment:

> *But the hardest part is, you say, "Well, Mrs. [teacher's name] got a free period, so I'm going to send you over there to take the test." And they're like, "No, I'm going to take … I don't want a new space. I want it with you." And, you know, like they're just … They don't like change either.*

At some schools, online grading resources and other tools were not ready for use during the initial PBE implementation. Multiple online gradebooks have been introduced to help teachers track proficiency instead of averaging grades on a 100-point scale. District 3 Middle School struggled with the use of an online gradebook that was not the right fit for the school's new grading style. Despite making the needed changes in her class, one teacher described her difficulty recording grades in the current online system:

> *And I expose the kids to ... like, we're working on this standard right now, but our grading system is still based on regular assignments.*

As students reached high school, there was increasing pressure to ensure that proficiency on standards was attained. In middle school, students still had considerable time to go before obtaining their diploma. In general, students were not held back in middle school if they did not achieve proficiency. Therefore, teachers and students alike felt greater pressure in high school than in middle school. For example, a participant from District 2 High School shared the following:

> *Well, and another thing we had to do for next year because of this system, like at the middle school, they can teach something to a Level 2 in seventh grade, give them a 2. It doesn't hurt because there's no GPA, and then teach the Level 3 and 4 in eighth grade. That 2 being on their grade in the seventh grade doesn't hurt them. It doesn't make them failing. It doesn't hurt their GPA.*

Participants expressed a hesitancy to shift to a more student-centric teaching and learning model in which students moved at their own pace. They even resisted the idea of a workshop model in which students learn in smaller groups based on factors such as abilities or learning topics. As mentioned earlier, teachers tended to use whole-class instruction methods. A participant from District 3 High School explained why she had concerns about a workshop instructional model:

> *You know, we talk about workshop model a lot, but workshop model is beautiful when to get everyone kind of to the same place by the end of a class period, but in order to do*

it for a whole unit, I just, I see chaos. And then what do you do with that group if they've already taken that assessment? What do I do with them in the interim? Because I can't teach them the next unit and try to get them through this unit, too.

At some of the schools that were most negatively impacted by PBE, teachers said that the shift to PBE was detrimental to their belief in students. Students worked towards achieving scores of 3 or proficiency instead of trying to excel in their understanding of the content. As a result, teachers reduced their expectations to ensure that all students could reach proficiency on the standards needed to graduate. A participant from District 2 High School explained her views on this issue:

I feel like the system has made us dial back our expectations, because it's now at the point where it's like, "I've done this, I've done this, I've done this, can I have my 3 now?" It's like, "Well, we'll give you 2.5." I feel like we're giving away 2.5s when people don't really earn them, and I think that's system-wide, not just the math department. That's a system-wide thing happening.

The Continuum of Pedagogy and Impacts on Opportunity

The continuum of pedagogy describes whether HOTS opportunities are happening in the PBE classroom and thus informs the second research question. The IPI tool was used to gain an understanding of the level of higher-order thinking that students engaged in. Classroom observations were scored on a scale of 1 to 6. The scores were represented in the following way: 1 for complete disengagement, 2 for student work with teacher not engaged, 3 for student work with teacher engaged, 4 for teacher-led instruction, 5 for student learning conversations, and 6 for student active engaged learning. A score of 5 indicated that conversations included more than one student. A score of 6 indicated that students worked individually. Thus, scores of 5 and 6 showed that higher-order learning opportunities were present in the classroom.

In the following subsections, participants' responses about the types of learning opportunities that are available in the classroom are organized according to the IPI continuum. This structure serves as both an organizational tool and as a means of unpacking the qualitative and quantitative results, which are further examined in Chapter 5.

Complete Student Disengagement (IPI Score of 1). Complete disengagement was infrequently observed with the IPI tool, but it is very clear when it is seen. Participants shared their experiences with disengagement in the classroom. They frequently mentioned the time devoted to student homework in class, which was typically the last 30 minutes or so of a class. Students often did not use this time to work on their homework; instead, they were completely disengaged from learning. Participants did mention efforts made to reengage students or change their class structure. A teacher from District 2 High School shared the following experience, which seemed to be typical among participants:

> *Probably the last 30 minutes are for them to get their work done with me there. You know? Which should be a good thing. Work on your homework while the teacher's still around rather than take it home. Now I can answer questions, solve any confusion, but with the exception of a few, they don't. They see it as free time.*

At District 2 Middle School, another teacher similarly reserved the last 30 minutes of the class for students to work on their homework. Again, while a few students opted to complete their work, others were disengaged:

If I set them free to work, like I say, I give them probably a half hour to work on the homework, I'm not saying there's no kids, a few kids will start doing a few problems, but no kid works for that full half hour, and a lot of them don't work at all.

It remains unclear why this practice continues if it is ineffective. An educator at District 1 Middle School was also concerned about students' lack of engagement at the end of a period or around transition times. When one student began to disengage early, others soon followed:

Maybe towards the end because if they see it's a minute or two minutes before, some of them will start picking up and it's like a domino effect. When it's one or two ... But then also with our double blocks, our back-to-back double blocks, sometimes at that bell between the two blocks, they start getting restless 'cause they heard the bell.

Student Work with Teacher Not Engaged (IPI Score of 2). Most student learning opportunities that were categorized as a 2 consisted of students working on practice problems or homework while the teacher sat at their desk or engaged in another activity. In such cases, students often did not collaborate with their peers because they focused on their own work. This would not always be the case though, as they may be working together on lower-order type problems. A teacher at District 1 Middle School implemented a fully individualized student pace model, which explains the lack of collaboration:

Everybody's working on their own little thing. We don't ever come together taking one of those things and say, "Okay, this was wrong. What made this wrong?" So, they don't get that. I don't know what the word is, but they don't get that conversation, I guess.

Typically, teachers intervened or checked in when they felt that a student was stuck on a problem or asked for support. At District 1 Middle School, this individualized structure drove all classes. The teachers believed that challenges would be less likely to emerge outside of school if students completed their work in class. One participant believed that this approach reduced the number of difficult interactions with parents:

> *So, they can't help the kids. So, we do them, we give them the opportunity to do a lot of practice at school to reduce the amount of phone calls, problems that could occur at home.*

High school teachers still devoted a portion of the class to instruction. However, they ensured that considerable time was left at the end of class for homework. As stated earlier, this time was sometimes used effectively and other times not. The problems that students typically worked on were similar to the examples presented in class, which did not represent higher-order thinking questions. A teacher from District 2 High School explained how class time was used for homework:

> *Well, that's just it, I use the term "homework." With block scheduling, you've got plenty of time to get it done in class. It's in-class work, you shouldn't have to ever take anything home unless it's maybe the day before a test and you need to study a little bit. The work itself, I give plenty of time to do.*

The same practice was also reported at middle schools that did not fully follow a student pace. A teacher typically planned a lesson for the first part of class, but time was always set aside at the end of class for students to work independently on their homework, which was due the next day. A teacher from District 3 Middle School described how she implemented this practice in her classroom:

> *We do practice problems, and then I'll give them the homework. They always have time at the end of class to work on homework. So, that's generally what it is.*

The previous examples show that many teachers limited the amount of work that students are asked to do outside of the classroom. This means that students did not spend time conducting investigations or deepening their understanding of a topic but rather practicing the material that was due at the beginning of next class. Many participants reported seeing this change occur over time. A teacher from District 2 High School said that he now had lower expectations for students:

I remember my math teacher in high school would give 60, 70 problems a night. I give about 10, and I give them a half an hour to do it. There's no reason for homework.

A similar sentiment was reported at multiple schools. Participants felt that homework completion levels have decreased. As a result, students needed more time to do their homework in class; otherwise, it would not be completed. This challenge was compounded when students already had learning deficits. A teacher at District 1 High School who also works in the school's Response to Intervention program shared the following concern:

As an interventionist, I see that very few of the students do the homework at home.

Moreover, class time was used for the rote practice of skills. In many cases, the teacher's presence was not required, as students used a computer program (e.g., IXL) that gave them instant feedback about whether their answers were correct. With IXL, students were initially assessed and given practice problems at their current skill level. Many teachers allocated time during class for students to use the program and develop their skills. A participant from District 2 Middle School described how she used IXL in her classes:

We're fortunate enough to still have IXL as well. In my book, there's not a lot of rote practice. There's a few problems and then they go into applications of that. And they *need to have that rote practice as well, and that's instant feedback, and we can hang onto IXL as well.*

Many high school teachers felt that it made sense to provide students with answers to homework problems to allow them to correct problems themselves and avoid the teacher's corrections. To a certain extent, it enabled students to pinpoint the questions that were challenging for them, which they could later ask the teacher for support on. This shifted responsibility to the student rather than the teacher. A participant from District 3 High School shared her rationale for giving all the homework answers to students:

That shows a respect for their time is equally as important as my time. And kids, it's great, kids are just using that app and taking a picture and copying it down anyway. They can get answers to everything. You're just fooling yourself if you think they're just doing all that work by hand. They take a picture and they copy it down and they're done. So, that's a waste of my time to look at that.

The challenges associated with students working at an individual pace have become evident to teachers who have experimented with this practice. During the first year of this study, all students at District 1 Middle School worked completely at their own pace. However, teachers began to notice that some students did not make the same progress as their peers and fell increasingly behind. In response, a few teachers tried to keep students together as a class when possible. One participant explained the changes that he made over the course of a year:

That was a big change I made this year. Where it was all at your own pace last year, we found all those lower students that I could have a kid who's still in Chapter 2 right now.

Student Work with Teacher Engaged (IPI Score of 3). When a classroom observation is assigned an IPI score of 3 instead of 2, this means that the teacher is actively engaged in the learning of their students. Often, they make their way from student to student or work closely with a small group. Students still work on lower-order tasks during this time. Many participants felt that the period when students worked on practice problems on their own represented an opportunity for them to receive support from their teacher. Among the teachers who left students to their own devices during homework time, there appeared to be as many—if not more—who advocated for full support during this time. Although students worked individually during homework time, a participant from District 1 Middle School explained that the purpose of this setup was to give them the opportunity to receive support on areas that they did not know while at school:

> *Now the time I see, a lot of time that students are working on problems with support in the school. And that typically, like the homework-type problems that they, you make sure they have time in school to work on those.*

In all the focus groups, participants said that students needed support with foundational skills. Thus, time was devoted to skill building in class. These were specially designed or handpicked items that teachers chose based on what they viewed as student needs. Teachers used strategies such as preteaching or teaching mini-lessons with students or groups of students to support their foundational understanding. A participant from District 2 Middle School described this process:

In sixth grade, for instance, the first lesson they had to multiply, divide, decimals, and mixed numbers. And they didn't know this. We have to preteach before you even get to the first lesson. So, that's the problem. That's the problem this year. And maybe we'll be here next year, but with less.

Another strategy for addressing the gap in foundational skills was to begin the class with one or several warmup exercises, which were typically skills-based and unrelated to the lesson topic for that day. Teachers worked with students to correct their answers to the problems and understand the foundational skill in question. An educator from District 2 Middle School described how she structured this time in her class:

I still have a group that I don't think it matters what I do. And they just don't know their multiplication facts, but I've just tried to instill in them how much easier everything else is. I mean, if you don't know your multiplication facts, it's pretty hard to work with fractions. That's just a huge setback. And I use that probably two or three times a week just as a warmup. And then when I don't, I do a problem of the day.

Participants expressed that the lack of foundational skills was a major challenge, as it impeded opportunities for students to work on problems that required greater depth of understanding. Furthermore, this appeared to be a vicious circle. Teachers did not devote time to higher-order thinking tasks in class because they did not feel that students were ready to attempt them, but students will never be ready if they do not attempt them. According to a participant from District 3 High School, it was difficult to engage in interesting tasks when few students had the necessary foundational knowledge:

At this point, fewer have enough of a base down that I can try to do more interesting things.

In addition, many participants felt that the proficiency-based grading structure gave them permission to allow lower-achieving students to only focus on attaining proficiency and not worry about anything beyond this scope. As a result, much of the work assigned to lower-achieving students was lower on Bloom's taxonomy or other measures of HOTS. However, participants believed that this was the best choice for lower-achieving students, as they did not have the skills needed to develop a beyond-proficiency understanding. A teacher from District 3 High School explained that lower-level students focused on Level 3 problems:

> *Honestly, I feel like that's been a positive, too, is that the kids who are struggling where they know that they can focus on Level 3, so as opposed to in a traditional-style class, where there are just questions that they don't get and so they wind up with 80% or whatnot. If they can show that they meet … Like separating that this is what you need to know, this is an extension would-be-nice-to-know stuff. And I think they feel less bad about like missing questions because it's labeled out for them. These are hard questions that not everyone's supposed to get.*

However, participants reported that students greatly appreciated seeing many practice problems, then having the opportunity to attempt them and receive feedback from their teacher. This structure provided students with an algorithm to follow and apply to similar questions, which they continued to use on the assessment. According to a teacher from District 3 High School, all students loved this method, and class time passed rapidly for them:

> *They're just waiting for it. And the class goes by so quickly for that class if I'm doing 75 minutes of just, "This is the problem. This is what you do. Here's a different problem. Here's…" The information, like helping them pull up that information and then applying it to where they need to go.*

When students can move at their own pace, higher-level students typically have two options, which are dictated by the teacher. The first option is to attain proficiency, then move on to the next standard. The second option is to attain a level beyond proficiency on the current standard. However, if a student chooses the first option, they may never have the opportunity to understand the content with a greater level of HOTS. If clear opportunities to achieve a level beyond proficiency are not provided, students will simply move on to the next standard. Thus, one teacher at District 1 Middle School felt that the PBE system benefited higher-level students because they could continue to move ahead:

I think it's really important for those kids that have the desire to move on that they're not held back. And I think the way that our system is set up right now, that really suits that, that population of kids. So, that's really super important in all of our grades.

The concept of moving ahead after attaining proficiency can be true for all students in the class. Participants said that, once the average student attained proficiency, they could choose to skip more advanced problems that would require the use of HOTS. In some cases, teachers forced advanced students to try the higher-order problems; in other cases, they allowed them to move on like the rest of the class. This seemed to be dependent on the teacher. One participant from District 1 Middle School allowed some students to skip the higher-order problems:

Typically, I'll let the regular kids skip it if they don't think they can do it, but I don't like high fliers to skip it. I say they must attempt it.

Teacher-Led Instruction (IPI Score of 4). Despite a call for a student pace and individualized instruction, teacher-led instruction remains a mainstay at five out of six schools that participated in this research. This approach may consist of the teacher using questions and answers or talking students through examples. During this time, students are attentive and follow along. Many participants seemed most comfortable with this model and believed that it had a true impact on students' learning. A teacher from District 1 High School described learning under this model:

> *I think for me, when it feels the best is like, they've learned how to do something and then we're doing problems on the board, and I'll kind of let them take over and it kind of becomes a competition, okay. Who's going to get this right. Who's going to get this right. And someone will shout out an answer, and someone else will look and say that can't be right because of whatever, because of this and this. And then they'll look at it and then they'll find their mistake. And there's a lot of back and forth and a lot of interaction that way.*

The question-and-answer component of teacher-directed instruction provided immediate feedback about students' understanding of the topic. However, it sometimes only provided information about how the most vocal students were doing. A participant from District 1 High School used questions and answers to help him better understand his students' progress:

> *I like to go around and look because I'll ask them individually, it only takes a second, "Were you okay with this thing?" "Yeah, I got this." And then the next kid will say, "Well, I did these, but I couldn't do these" or something. And then that gives you a pretty good idea of kind of start that day. You can say, "Okay, a lot of people couldn't do this problem. So, let's do this one together." So, it's a quick way for me to assess and see what I need to.*

Participants explained that direct instruction mostly consisted of a step-by step process that students could use to solve problems. For example, teachers often shared an algorithm that could be implemented by students. However, less time was spent explaining the "why" behind the learning; rather, most instruction concerned how to perform a task. A teacher from District 2 High School explained the instructional method that she used in her classes:

> *Show some examples on the board, list out the steps, so if they do take notes, if they're a person that needs that visual step by step. I'd say it's even a little longer than hers. I usually give them probably a good half ... I mean we've got a block schedule, so we have hour and 10-minute periods.*

Whole-class instruction was valued by many participants, who believed that this was a time when students were more engaged. Moreover, they felt that students wanted this time to better understand how to complete their homework and attain proficiency on the current standard. A teacher from District 3 High School felt that whole-class instruction was beneficial and that students desired it:

> *As much as it's not in vogue right now, most kids that I have in, like, primarily direct-instruction classes, I think want the instruction time. And if everyone's in a different place, then you couldn't really, like, what are you going to talk about?*

Student Learning Conversations (IPI Score of 5). An IPI score of 5 indicates that higher-order thinking opportunities are integrated into classroom instructional practices. In other words, higher-order conversations are taking place between students or student groups. Often, these opportunities will replace an open-ended task that students must complete together. Two participants from District 3 High School joined a professional development network for mathematics teachers, which helped them identify ways to incorporate more open-ended tasks in the classroom. One participant said the following:

> *For my algebra class, I do more task-oriented things. [Teacher's name] and I are part of a network where they've been training us to do that for Algebra 1 for Better Math*

Teaching Network. So, I have more, like, group project task-oriented things that they do, and they have, like, a protocol that the kids run through, and they've been kind of trained to do this.

Open-ended tasks evolved into open middle tasks, which help make problems more accessible to all students. There are clear guidelines about what is asked, and students are given the opportunity to plug in numbers and make calculations to attempt to to find a solution. This approach to problems requires students to synthesize information that they have been given, make predictions, and develop a hypothesis. A teacher from District 3 High School explained why this approach may be more accessible to all students:

Like, everybody's always talking about an open-ended task, and they get a little lost in those. When they can go anywhere, and then I'm trying to close it up, trying to … I'm not as good at that. I like open middle because they can all do something in a different way, but I still have this thing at the end. It feels more … I'm a control freak, and it's easier for me, and it's almost more satisfying to them when they know if they got a right answer or not.

Participants believed that open middle tasks engaged students because they found them to be fun. Students worked together in small groups to tackle challenges, which resulted in many mathematical conversations. The teacher presented problems to the class, then monitored students' progress as they worked together to solve them. The following quotation shows how a participant from District 3 High School set up an open middle problem:

So, we have, like, if you were trying to solve equations, we'd do that forever. Sometimes it's just blank boxes in the middle of them. And you might say, "Use the numbers 1 through 4 and make an equation that has the highest answer possible or the *lowest answer possible." So, they have all these … There's actually [a website called] Openmiddle.com. There's tons of stuff out there. But those kind of things. Something where you take away some of the information and then you're often trying to do something extreme, not just solve, but find the biggest one possible, which can be a little more fun or … I did that one for you guys, where it was just, it was an accident.*

In addition, projects were a traditional method used by teachers to incorporate higher-order collaboration and conversation between students. Under a proficiency-based system, the number of projects did not seem to significantly change. However, some participants felt that the number of standards that had to be covered limited their ability to offer projects that required greater depth of understanding or were directly connected to the learning standards. a teacher from District 3 High School assigned a project to his class this year after not assigning one for the past couple of years. However, students were mainly concerned about how their grade would factor into the attainment of proficiency. He said,

> *But the last time I did it was two years ago, or three years ago. And I brought it back this year because I was a little ahead of the game, and I was like, "Maybe they will find this fun." And the only question I got repeatedly was, "So, is this a summative? Is this informative?" And I'm like, "Well, your reflection and your notes and your math are informative." And they're like, "It's only 15%." And then they just didn't do it. Well, they did a bridge, but it had nothing to do with the scale factor and the design that they drew.*

Many teachers viewed projects as "fluff," not true learning experiences. They did not consider them as an opportunity for students to work collaboratively on an assignment that required greater depth of understanding. A teacher from District 3 High School said that some projects that she had seen in the past were a waste of time for students and represented a low level of rigor:

> *So, I mean, we've heard about project-based learning and project-based assessment and alternative assessments, but the problem is, is you either lose some rigor or it feels fluffy. I mean, like, you made a bridge out of Popsicle sticks, you get an A, and you meet an indicator. That's fabulous.*

In addition, projects did not necessarily allow students to think critically. Teachers at District 1 Middle School assigned a mathematics project every March to coincide with a parent evening. This was notable because the school used a student-centered model throughout the rest of the year. In other words, project-based learning was only deployed once per year when the school planned to showcase the students' work to parents. One teacher said,

> *We have our celebration of learning every March, so I always do projects for that celebration.*

Many participants believed that students were more engaged when they worked on a project. Often, students worked together in a group to complete a task. The guidelines tended to more open, or there was opportunity for student choice. A teacher from District 3 Middle School described a group project in which students had to decide how to purchase a car:

> *For me, mine are more engaged when we're doing a project, because I tried to give them guidelines but not too many. So, that way they still have some self-choice on either ... like we bought a car. That's one of our projects is they buy a car. I let them choose the car, but I only gave them a certain amount of money. They had to still all have the same graphs and things like that. But you know, they were able to do what they wanted as far as purchase. So, that was fun for them.*

High school teachers felt that group projects were difficult to frequently assign. They tended to do so when the project had a connection to the real world and the content that they were currently teaching in class. For example, a teacher from District 2 High School explained that, as a unit, probability lent itself well to a project:

I couldn't make my class all project-based. I mean, I do a couple projects around probability and things that are more real-world. I have a project that I like to do with a murder mystery thing based on Newton's law of cooling for the logarithm unit.

To better teach students at their current ability level, District 2 High School implemented ability grouping. The teachers believed that this allowed advanced students to move more rapidly through the content and not have to wait for their peers to achieve proficiency. As a result, they had an opportunity deepen their understanding of a topic. A teacher from District 2 High School explained how this change impacted her teaching:

So, what I do with my group is they have exactly the same learning targets. It's just that the advanced group moves faster through them, so I don't have to spend 15 minutes on something that should only take five minutes or things like that.

Another approach that created opportunities for deeper learning was peer tutoring. Once a more advanced student attained proficiency, teachers often paired them with peers who were still trying to understand to understand the material. Tutoring may have helped advanced students understand the content at an even deeper level. However, there was also a risk that students did not have the opportunity to move beyond the proficiency target. A teacher from District 2 Middle School paired gifted students with lower-level students in a mixed-ability classroom:

As I'm working with another student, another one has a question and sometimes, I have in one class, I have four gifted and talented students and then I have some very low-

level. So, some days I'll just have them pair up and work with and it's very cool to see them kind of doing that.

Some teachers spent some time at the beginning of class to "hook" students. A hook activity is designed to engage students in thinking about the topic of the day and create opportunities for higher-order conversations to take place. Teachers use hooks to activate students' thinking and prepare them to participate in learning during the class period. A participant District 3 High School described how she used a hook activity with her class:

Once they have that hook, then they'll go into it deeper on their own. Well, and you don't have to do hook, after hook, after hook, right? If you get something interesting going, their brains are all going... Then they'll do some of the boring stuff, too, and then it just has a different feel.

At District 2 Middle School, a change in mathematics curriculum resources prompted students to engage in more higher-order conversations. The enVision program challenged students to solve more difficult problems and impacted how they communicated with each other. According to one participant, the use of enVision encouraged the students to value group conversations:

Another thing, in enVision, in many problems they ask students to explain, which is higher-level thinking. So, in the beginning they say, I know the answer, but I can't explain. So, that's what the enVision was teaching them how to communicate with each other, how to ask questions, how to analyze problems. That's valuable as well.

Many higher-order conversations that occurred in the participants' classes were connected to real-world problems. Students were encouraged to guess or estimate the correct solutions to the problems. They discussed reasonable answers with each other and refined their thoughts. This approach appeared to be accessible to students of all levels. A teacher from District 3 High School gave an example of a problem that helped all students use HOTS:

> *The first time they were presented the question and they had to work in their groups and they wanted... The question they had is okay, well, how long is a car? And I was like, well what do you think is reasonable? Knowing that's what I wanted them to write about, is how reasonable was their solution.*

Some teachers found ways to integrate real-world problem solving for students by examining what took place around them at their own school. For example, District 3 Middle School was in the process of constructing a new school building. The students received a tour of the latter and had an opportunity to incorporate their observations in a small-group assignment that required them to examine the building’s blueprints. According to their teacher, they used their own school as a basis to engage in real-world problems:

> *We were able to take a tour, and I know all the grades didn't. We took a tour of the new school as it was being built. Then I got a copy of the blueprints. So, we did surface area and volume based on the blueprints of our new school.*

Student Active Engaged Learning (IPI Score of 6). An IPI score of 6 indicates that students are fully engaged in higher-order learning. This means that hands-on learning and problem-based learning are taking place. In addition, students are fully engaged in thinking and learning in authentic projects. However, a challenge that high school teachers sometimes encountered when trying to incorporate problem-based learning in their classes was a dependency on available resources. When they wanted to challenge their students, participants often had to comb through the textbook to find appropriate questions, as described by a teacher from District 1 High School:

I try to self-select higher-order thinking questions in the book. I do a lot of labs, too, in that statistical class, and they have a lot of analytical questions and thinking, like, a whole project they work on data collection.

Some participants developed their own banks of books or resources to pull class projects from. A common theme was that teachers sought advanced material for students who had completed regular class assignments and demonstrated proficiency before their peers. For example, an educator from District 1 Middle School used a supplementary book to source projects that advanced students could work on:

Yeah, then also, seventh grade doesn't have them, but I have a ton of books. They give you assessments which we don't use 'cause they're zero through 100 and then there's extra practice and reteach and enrichment. There are sources to use, and sometimes I'll pull projects out for some of the advanced kids.

Many of the additional assignments that teachers gave to students who had attained proficiency early appeared to be more engaging and incorporate HOTS opportunities. However, such opportunities were usually only accessed by a small number of students. A teacher from District 1 Middle School provided an example of a project that advanced students could complete:

One that one of them did was, it's called a pixel person. The student had to make a person, or some people did an animal, but out of two-inch by two-inch squares and they got to make their own person, have fun with it, make a face on it and then they had to calculate area and perimeter and they had to find the fractional pieces of every color that they used in their pixel person.

However, it was an ongoing challenge for some participants to find work that was at an appropriate level for advanced students. Some advanced students sought more and more problems to work on once they had completed their assignments, as one teacher from District 1 High School explained:

> *Yeah. They always joke, "When I finish, do I really have to keep going?" And then none of the kids that ever finish ahead of time want to stop, so they want me to find more and harder things for them to attempt. I'll do like the Algebra II section of rational equations, radical equations, then I'll do some trig with them as well.*

Some teachers tried to meet these students' needs by preplanning a different assignment for them. While most students were assigned work at the basic level of proficiency, more advanced students received more challenging word problems. A participant from District 3 High School explained this process:

> *With my mixed group, which is every one of my groups except for one, with my mixed group, though, I can take my high fliers and I can give them a set of word problems or something that's like, "Work backwards. I gave you the area, so what would half of the radius be?" You know, and then, I give them a series of all of those. Or shaded region problems in geometry. I can take my high fliers and I can say, "Okay, if you clearly do not need help finding the area anymore, you go do this."*

Some teachers provided several ways for students to attain proficiency on a certain unit. For instance, students worked on typical practice problems, while others chose a project or presentation. When these options were given, advanced students typically chose the project. A statistics teacher from District 2 High School said that only a couple of his students opted for the latter:

In my stats group, a couple of kids did a project. This was last year's stats group, but so it's a year old now, but same general kids, you know.

When students were grouped into an advanced class, teachers designed more challenging opportunities for them. At District 2 High School, which recently divided students by ability level, participants reported that advanced students typically worked on word problems, which offer more opportunities to engage in higher-order thinking. A teacher from this school explained what differentiated students in her advanced group from those in the regular group:

And also within a chapter of a book, sometimes there's that last section in the chapter that just is a more challenging word problems, whatever. And I'll do that section with the advanced group.

Advanced students often had leeway to move through the curriculum more rapidly than their peers. If they worked alone, they were sometimes allowed to continue progressing through the material and demonstrate that they understood the material through summative assessments. According to a teacher from District 1 High School, he often looked beyond the current textbook to find material for more advanced students to work on once they had completed the regular material:

The other thing I do, I have extra units that I have for the advanced kids as they finish my normal curriculum. I have a couple units from Algebra II and Geometry that those kids will do, push more of that.

According to the grading rubric for a PBE classroom, students were typically awarded a 3 when they attained proficiency. They had to complete tasks at a level beyond proficiency to be awarded a 4. Opportunities to achieve a 4 differed from classroom to classroom and teacher to teacher. In some cases, they were well-established and clearly delineated to students from the beginning of the unit; in other cases, they were developed on the fly by teachers. In many cases, they were not available at all, as students did not access them. For the most part, participants said that higher-order thinking opportunities were most frequently integrated into the mathematics classroom when students worked to attain a 4. A teacher from District 2 High School provided an example from his own classroom:

> *Like I said, I do a Level 4 project with logarithms, but it's just Level 4 work, it's not the whole unit. You can't teach the whole unit. There's so many things that go, step by step, they go into it that, like she said, they don't recall anything. If you tell them to do a project on this, they're just going to do the project but not actually learn the stuff along the way.*

Moreover, he explained that if a Level 4 project was offered to all students at the beginning of the unit, they would not develop foundational knowledge on logarithms. This is because the project only required students to access a certain portion of basic knowledge on logarithms. The teacher said,

> *It's a Level 4 thing. It's a cool Level 4 project, but I can't teach my whole unit that way, because then they would only learn how to solve that one equation, and not anything else about logarithms. You know what I mean?*

In the current grading structure, students are allowed to decide whether they want to pursue a 4. According to participants, only a small group of students attempted to obtain a 4. Average students wanted 4s to increase their grade, not to develop a deeper understanding of the content. A teacher from District 2 High School shared her thoughts on the pursuit of 4s in her classroom:

Yeah, only the top students pursue it, or there's the mediocre students, so to speak, that will do 3s all along, and then all of a sudden in the last week of school ... I had a kid today say, "Oh, I need to work on some 4s" who hasn't done a 4 all semester. They want to get on honor roll at the very last week of school by getting everything up to a higher average.

Many teachers created opportunities to obtain a 4 in the form of extra problems on homework assignments and assessments. In theory, these problems should require students to use HOTS. However, they were not a focus of instruction but more of an add-on to assessments. It remains unclear how or when students are provided with tools or strategies to solve these problems. A teacher from District 1 Middle School discussed the practice of adding Level 4 problems to student assignments:

Every homework, every assessment has a chance to get a 4 on it.

In addition, some teachers asked students to at least attempt Level 4 problems on their own. For example, a Level 4 question might be included at the end of an assessment and all students had to try solving it. Then, teachers typically reviewed the answer when the entire class reviewed the assessment together. A teacher from District 2 Middle School said that he expected all students to try solving the Level 4 problem on the test:

It's an expectation in my class that if you hand me a test, you have to at least attempt a 4.

Participants were asked to explain opportunities for higher-order thinking in their classrooms. One of the most common responses was word problems. Many textbooks structure chapters in a way that moves from basic practice problems to more challenging word problems. The latter typically ask students to apply what they learned earlier in the chapter to solve a problem. In addition, they must interpret what to do after reading a paragraph that describes the problem. A teacher from District 2 High School gave a straightforward response about higher-order thinking opportunities exist:

> *Yeah, so traditionally speaking it's the word problem stuff. If you want to put it in the most basic of math terms.*

Participants visibly brightened when asked to provide examples of higher-order thinking opportunities in the classroom. They described problems that were clearly fell under this category. Again, students would only access questions that required the use of HOTS if they aimed to obtain a 4. A teacher from District 2 High School gave an example of a higher-order word problem:

> *And then there was another question on the Level 4, about like theater seating where the rows get wider as you go up, so how many people are you going to be able fit, based on that, after 10 rows, it goes up by another seat on each end. Then also, I think there's another question on it about how much money could you get if you sold out, based on this many rows being this price. So, you're adding a couple of different steps together. So, it's a word problem with multiple steps.*

Most participants believed that students needed to develop their understanding before they were ready to attempt more challenging word problems. More specifically, this meant practicing the basics, demonstrating understanding through formative assessments, and demonstrate proficiency through a summative assessment. If they completed these steps and were ahead of their peers, they were given the opportunity to tackle more challenging problems. A teacher from District 1 Middle School explained her understanding of this progression:

> *And then it progressed to like a problem where they actually had to find X. But it was just a triangle and word problems, real-world problems. And then it escalated to the*

word problems and putting it all together. So, it definitely is like a clear process of they're starting at a lower level and then escalating to that.

At other schools, some teachers added Level 4 problems to their assessments and let students decide whether to attempt them. When students were not asked to attempt them, only a small group of students opted to do so. A teacher from District 1 High School explained that, although higher-order thinking questions were added to exams as a bonus, most students did not try to solve them:

Yeah, I think our goal was to kind of just try to put some problems on an exam, where it really was, did they even make an attempt at it. We talked about making it a little bit easier then, instead of making it a bonus question. Most of the time on the SATs, we see just a ton of them just stop it, they'll skip it. We want to get them more used to at least trying it.

One strategy that emerged among teachers to persuade students to participate in HOTS opportunities was the power of justification. If students could explain their thinking and why they took certain steps to solve a problem, they began to evaluate and not simply recall an algorithm. A teacher from District 3 High School believed that justification led to rigor:

That's where the rigor comes in. Just doing stuff on your own. If a teacher is going to run from person to person, that's just, this is how you do this and now you copy me. That's not as rigorous as trying to justify your answer.

The same teacher explained that she worked with all students to help them talk through their thinking and justify their work. She was a proponent of group conversations and open middle tasks. She also advocated for all her students to be able to talk about their process when she checked in on them. This teacher further explained her push for justification as follows:

And talking about what you were thinking. Even just having them share out their observations on whatever we might be starting or doing. Or what did you notice that kid... I'm trying to push, right now, my students in their justification.

Although some teachers supported the idea of using a problem-solving approach to teach mathematics to all students, others did not fully see the value in this method. Many participants felt that the initial focus should be on building foundational understanding; otherwise, students would never be ready to solve problems. Educators at District 1 Middle School had access to a professional development session on problem solving for teachers, but it was not well-received by all. Many believed that the strategies shared at the meeting should only be used with advanced students. One participant explained her reaction to the session:

She wanted us to—sort of telling kids the formula for a square or something. She wanted us to draw pictures and try to get them to come up with a formula which okay, maybe for somebody who's more advanced, we could try that. But for an average kid or below-average kid, I just started giving the formula and work from there.

Ultimately, teachers agreed that it was important to teach problem-solving strategies to students. However, they had to balance this endeavor with the need for all students to demonstrate proficiency on the standards. This recalled the old adage of "breadth versus depth." This challenge was aptly summarized by a teacher at District 2 Middle School:

It's the quality, not the quantity...We'll really have to slow down and teach kids to think in deep and analyze. And there's some things, actually, they have to know.

CHAPTER 5

DISCUSSION

The purpose of the explanatory sequential mixed-methods study presented in this thesis was to identify the frequency of higher-order thinking opportunities for middle school and high school students following the adoption of proficiency-based diploma requirements in Maine. In addition, it examined how HOTS opportunities are integrated into the classroom and the factors that impede or enhance this process. This chapter discusses major findings in these areas. Furthermore, it presents implications for research and practice, the study's limitations, and recommendations for future research. The chapter concludes with a brief summary of the research.

The three questions driving this research are:

1. How frequently are mathematics students engaged in high-order thinking in proficiency-based classrooms?

   a. Are there significant differences across professional support styles in the frequency of higher-order thinking opportunities?

   b. Are there significant differences across grade spans (6-8, 9-12) in the frequency of higher-order thinking opportunities?

2. How and in what ways are mathematics students engaged in higher-order thinking when taught in a proficiency-based education system?

3. What factors enhance or impede educators' ability to implement higher-order thinking tasks in proficiency-based mathematics classrooms?

The frequency of HOTS opportunities was determined through a series of classroom observations at six different schools (three middle schools and three high schools). The observation tool used was the IPI. The results of a two-sample *t*-test showed that instructional practices at the high schools were scored higher than those at the middle schools at a statistically significant level. At the district level, schools in the district that received coaching from the GSP scored higher than schools in other districts, which had developed their own models. This difference was statically significant.

The most common components of PBE integration that participants discussed in the focus groups were (a) grading practices, (b) assessments, (c) individual pace learning, (d) standards, and (e) student supports. Each of these areas presented educators with implementation challenges. However, there has been some success in the areas of (a) positive learning impact, (b) relations, (c) engagement, and (d) clarity and communication. However, factors such as (a) negative student outcomes, (b) failing to build support for PBE, and (c) the paradigm shift to PBE created multiple challenges for teachers. However, educators still found ways to integrate HOTS opportunities within the PBE classroom.

Interpretation of Findings

This section reviews each research question in light of the quantitative and qualitative findings from the study. Moreover, it discusses factors that educators focused on during PBE implementation in the classroom and how that aligned with the levels of higher-order thinking. Two major impacts were revealed in the findings: independent learning structures and the disparity of higher-order thinking opportunities.

The Frequency of Higher-Order Thinking in Proficiency-Based Classrooms

Regarding the first research question ("How frequently are mathematics students engaged in higher-order thinking in proficiency-based classrooms?"), the findings from the quantitative analysis suggest that there is a difference between middle schools and high schools as well as

District 1 and District 3. The explanatory sequential nature of the study allowed the qualitative findings to be reviewed and cross-referenced with the results of the quantitative classroom observations. Participants in each focus group discussed the components of a PBE system and how they chose to integrate them at their own schools. Varying levels of attention were given to certain practices. For instance, in District 1, there was a greater focus on individual pace learning and student support. In District 2, grading practices and their impact on students was a major topic of conversation, along with ways to support students in the PBE system. District 3 tended to view PBE through the lens of formative and summative assessments, along with individual pace learning practices.

There was also a difference in PBE focus between middle schools and high schools. The middle schools tended to focus on individual pace learning and student support, while the high schools tended to focus their implementation efforts on grading and assessment. The findings revealed that higher-order thinking practices were more present in the school district that sought outside coaching. In addition, the study found that HOTS opportunities were more present at the secondary school level.

Differences Across Professional Support Styles. The three districts that took part in this research followed three distinct professional support styles in their implementation of PBE practices. In District 1, professional development was conducted internally with external support. District 2 was a member of a regional cohort, the MCCL. The latter provided training and resources to staff during the implementation of PBE. Lastly, District 3 teachers received coaching from the GSP and had access to frameworks and individualized support.

There was a clear difference in mean IPI scores and the total percentage of higher-order thinking opportunities observed in the classroom between the districts. In fact, the ANOVA and

Tukey's honest significant difference test found a statistically significant difference in mean IPI scores between District 1 (self-developed professional development) and District 3 (external coaching). Moreover, there seemed to be a clear reason for the difference in mean IPI scores and the percentage of HOTS observed in the classroom between districts. Specifically, the focus on individualized instruction in District 1, particularly at the middle school level, led to many observation points being scored as a 3 or a 2. In middle-school classrooms in District 1, students were seen working on their own or with individual teacher support on practice problems that were at a lower level on Bloom's taxonomy. This could have accounted for the district's much lower scores on the IPI.

The practice of allowing students to move at their own pace appeared to put them at risk of not receiving as many higher-order thinking opportunities in the mathematics classroom as their counterparts who remained in a whole-class structure. This is because they only accessed more challenging problems by choice in many cases, and there were not as many opportunities for mathematical discussions.

District 2 began its PBE journey as part of a regional cohort that drove the content area standards used in its territory. The level of HOTS opportunities in the classroom in District 2 increased from the first visit to the second visit two for two main reasons. First, the middle school decided to move away from the lower-level standards that were used during the first year of the study and adopted a new mathematics curriculum that integrated more higher-order thinking opportunities in the classroom. Moreover, all students had access to these opportunities since the entire group stayed together for whole-class instruction.

However, the high school in District 2 was resistant to change and largely viewed PBE through the lens of grading. Teachers were troubled by the accountability connected to the proficiency-based diploma and a system that they saw as harming students' aspirations. To incorporate more HOTS opportunities in the classroom, the teachers shifted to ability-based grouping. As a result, advanced students were observed participating in HOTS opportunities during the second visit. Although this increased the school's total mean IPI score, this practice may have limited opportunities for average and lower-level mathematics students.

The higher IPI average scores in District 3 can be linked to the teachers' search for professional development opportunities and efforts to incorporate mathematical discussions and open middle questions in their regular whole-class instructional plan. Thus, they focused on assessment structures within the PBE system, allowing retake assessments to ensure that all students can attain proficiency. However, they also instituted accountability measures such as minimum goals that needed to be met prior to reassessment and score caps on makeup assessments. This approach enabled a much more fruitful experience for students. The differences in mean IPI scores and the percentage of HOTS utilization in the classroom between districts are shown in Table 5.1.

Table 5.1 District Data

	Mean IPI Score	**% HOTS**	**Professional support style**	**Main PBE Themes**
District 1	3.42	11.0%	Self-created	1. Individual pace learning 2. Student supports
District 2	3.65	15.0%	Regional cohort	1. Grading practices 2. Student supports
District 3	3.87	24.0%	Coaching	1. Assessment 2. Individual pace learning

Differences Between Grade Spans. When middle schools and high schools were examined separately, the high schools clearly had a higher mean IPI score and percentage of HOTS opportunities in the classroom. In addition, this difference was statistically significant on a two-variable *t*-test. A contributing factor was the greater frequency of teacher-led instruction and whole-group activities at the high school level. Overall, the high schools scored in the average range on percentage of HOTS opportunities in the classroom, but their approaches considerably differed from school to school. By contrast, the individualized nature of instructional models at the middle schools meant that students worked on their own and engaged in fewer group conversations than high school students.

The factor that had the greatest impact on the lower percentage of HOTS opportunities at middle schools was the incorporation of individual pace learning. This was the most frequently recorded PBE theme among participants who taught at middle schools. When students were observed individually working on mathematics problems in the classroom, they typically followed an algorithm. To implement individual pace learning, teachers must provide more tailored support for students. Thus, student supports was the second most common PBE theme in the middle school focus groups. Individual student pace structures clearly impacted HOTS opportunities. Differences in PBE implementation between middle schools and high schools can are shown in Table 5.2.

Table 5.2 Grade Span Data

	Mean IPI score	**% HOTS**	**Top PBE themes**
Middle schools	3.44	13.6%	1. Individual pace learning 2. Student supports
High schools	3.78	18.0%	1. Grading practices 2. Assessment

Student Engagement in HOTS Opportunities in the PBE Classroom

The second research question examined the extent to which students engage in HOTS opportunities in the PBE classroom. A common remark among participants was that students had to master foundational skills before they could attempt more complex work. On the surface, this sounds logical and, in many cases, entirely appropriate. However, a flaw in this logic arises when certain groups of students are not provided with frequent opportunities to engage in higher-order thinking because they have not mastered the basics; this appears to be a fundamental misunderstanding about who can access higher-order thinking tasks. However, this belief was not shared by all educators. Teachers at District 3 High School shared examples of how students of all ability levels can engage in HOTS opportunities in the classroom.

It was clear that the students at District 3 High School classrooms felt comfortable contributing to the learning community. They did not fear ridicule from their peers, nor take offense to the teacher or classmates questioning their assumptions. The students all seemed to understand the assigned task and were encouraged to explain their thinking. These are all factors that have been shown in the QUASAR research to foster student engagement in HOTS (Silver & Smith, 1996).

Regarding the typical PBE grading structure, minimum proficiency targets in mathematics were not always set at a rigor level that required students to think beyond the application level. This means that, to surpass this level, students had to complete Level 4 learning opportunities. By design, students who reached this point had already completed lower-level work. In many cases, teachers allowed students to skip Level 4 problems and move on; in other cases, advanced students were asked to work on them. According to participants, Level 4 tasks tended to have real-world connections and be more engaging for students. Sometimes, they were comparable to a project. However, this learning structure meant that most students worked on mastering the basics while advanced students had the opportunity to learn more about the content through projects.

Moreover, teachers sometimes included problems that were designated as Level 4 questions at the end of assessments. Several participants did not require students to complete these problems and reported that the only students who attempted them were advanced students who were aiming for a grade of 4. Other teachers asked all students to at least attempt the problems. However, this may not be a true indication that all students are engaging in a higher-order thinking opportunities, as the complexity of the problems themselves may prevent some students from accessing such opportunities without more formal guidance from the teacher. Thus, students who examined a Level 4 question and were unsure about how to proceed may not have persevered in attempts to solve it.

Over the course of this study, a notable phenomenon was observed at one high school. During the first round of research, all students were heterogeneously grouped in mathematics classes. The classroom observations showed that students mainly listened to their teachers deliver whole-group instruction or worked independently on homework. The percentage of higher-order thinking opportunities recorded in the classroom was 6%, which was well below the range of 18–25% recommended in the IPI. When observations were conducted again during the following school year, advanced students had been grouped in their own classes. As a result, the level of HOTS remained the same in the regular classes but dramatically increased in the advanced classes. This increased the overall percentage of HOTS opportunities to 18%, which represented the low end of the average range. Thus, the school found a way to incorporate more HOTS opportunities in mathematics classes, albeit by restructuring and providing these opportunities to advanced students. Although this represented a gain, it raises the question of whether it is possible to find ways to provide these opportunities for all students. For example, research has shown that heterogenous class groupings may allow all students to succeed at a high level, with a nearly nonexistent achievement gap. Schools that have had success with heterogeneous instruction have implemented regular workshops and professional development time for teachers. Perhaps these teachers would have been more successful and supported if such a structure had been implemented at their school (Linchevski & Kutscher, 1998).

Clear inequities exist in a system in which advanced students are the only students who consistently engage in higher-order thinking. In general, students are much more engaged in learning when they can analyze, create, and synthesize information and see the relevance of what they learn. Therefore, lower-level and average students who do not have opportunities to engage in HOTS are at risk of losing interest in mathematics and falling into a cycle in which their lack of interest is reinforced by more worksheet-style problems. Students who are not engaged in mathematics class may be less interested in STEM-related fields upon graduation. If teachers believe that students must master all foundational skills before engaging in real-world mathematical applications, they may never feel that it is important to master the basics.

PBE Factors that Enhance or Impede HOTS

The third research question concerns factors that may enhance or impede HOTS in the PBE classroom. When students are truly engaged in higher-order thinking, they create, evaluate, synthesize, and analyze information. On the IPI rating scale, these components correspond to student learning conversations or student active engaged learning. In a classroom in which these activities are taking place, it is clear that in-depth conversations are occurring and that students are engaging in authentic and meaningful work or trying to move beyond the application of a memorized algorithm.

Instructional Beliefs. The study results show that educators' instructional beliefs impact opportunities for higher-order learning in the classroom. A positive outcome of the introduction of PBE systems is an emphasis on increased clarity. According to focus group participants, the system required them to be clear about learning targets and goals, how assignments would be graded, and what rubrics would be used. As a result, students understood what was required of them to meet or exceed proficiency. In most cases, exceeding proficiency was tied to the use of HOTS; therefore, it is much clearer how to engage in these opportunities.

However, higher-order thinking opportunities for all students were not frequently observed when students had to work through the curriculum at their own pace. Instead, they were seen when teachers engaged in whole-class instruction that required all students to engage in a higher-order task. The tasks could be as simple as several students being asked to explain their thinking with different mathematical operations that would lead to the next term in a sequence of numbers. They could also consist of students making predictions about a real-world situation. The key was that all students had an opportunity to participate and think more deeply about a topic.

Certain factors appeared to impede students' access to higher-order opportunities in classrooms designed to allow them to move at their own pace. One teacher prepared a bank of practice problems and assessments linked to the mathematics standards. Prior to attempting the practice problems, students received some form of instruction, whether through prerecorded videos or an online tutorial. Sometimes, the teacher provided support to individual students or small groups of students. At any given time, they might be supporting students who were working towards achieving proficiency on five different standards. With only one teacher in the classroom, it was difficult for educators to determine how to examine a given topic in greater depth.

Throughout the focus group conversations, participants said that the students did not prioritize depth of understanding. Instead, they focused on completing work for the current standard and moving on to the next. If teachers did not require students to at least attempt Level 4 problems for the standard, many students considered their work to be complete when they attained minimum proficiency. In some cases, this resulted in top students racing to complete the standards to stay ahead of their peers rather than trying to understand the content on a deeper level.

Another challenge was that higher-order opportunities linked to 4s were often not fully developed by classroom teachers. If students worked on their own and completed the assigned content, a teacher might then ask them to help their peers rather than think about their learning in a different way. Tutoring can be considered a higher-order task in some cases, as reflecting on how to teach a topic to another person may require a greater level of synthesis of thought. However, this is different than engaging in a higher-order thinking learning opportunity designed by the teacher.

Individualized pace learning may also have an impact on how students learn content. If they are left to follow an online video or a tutorial, they may forego opportunities to ask questions about why a particular mathematical operation or algorithm works. Successful teachers understand that there are potential pitfalls and aspects of a lesson that they should emphasize with students to support a clear understanding of the material. Otherwise, students may try to memorize an algorithm or steps to solve a problem without truly understanding why these approaches work. Students would then replicate these steps on an assessment to attain proficiency and move on. However, this type of memorization is positioned at the lower end of Bloom's taxonomy. In such cases, students miss out on higher-order thinking opportunities and may not retain what they have learned.

Finally, the individualized nature of this type of learning means that students may not have opportunities to engage in learning conversations with their classmates. When students discuss why an approach worked or did not work through the mathematical process, they learn. Often, a teacher must be an expert in the content area to identify the right questions to ask and ways to engage students in these types of conversations with their peers. If these opportunities are absent, students must raise their hand and ask the teacher for help when they are stuck on a problem to process.

School Culture and Administrative Support. The focus group conversations showed that some districts were better placed than others in terms of school culture and administrative support. For example, District 3 made a clear commitment to professional development and establishing the purpose of PBE among teachers. The latter had a voice in the process and believed in their instructional methods. By contrast, many teachers at District 2 High School and District 1 Middle School had to implement a system that they did not believe in. Participants made multiple references to not having a voice in the process or the dysfunctionality of the PBE system. In addition, these two schools had the lowest percentage of HOTS opportunities observed in the classroom.

Implications for Theory and Research

In Chapter 2, multiple PBE models and higher-order thinking frameworks were reviewed. In this section, these are examined in light of the research findings.

Proficiency-Based Learning Model

The concept of attaining proficiency in one standard before moving on to the next underscores a clear system of accountability for students and teachers. There are common expectations for all, which are grounded in local and national standards such as the Common Core. Often, in education, one topic builds on the previous one; thus, students must demonstrate their understanding on one topic before moving on to the next. Additionally, PBE requires teachers to develop a clear picture of where every student stands in their attainment of proficiency.

In shifting to a proficiency-based model, the true challenge for educators is to contend with systems that are based on the longstanding factory model. At both the middle school and secondary school levels, classes meet on set bell schedules, and students move from one grade to the next at the end of each academic year. Teachers have certain curricula that must be covered, and students are not free to pass from one grade to another when proficiency is attained. Each class has one teacher, who cannot effectively teach multiple lessons or units to students who are at different points in their learning journey. Although differentiation is a foundational component of PBE systems, teachers cannot differentiate with fidelity with class sizes of 20–25 students. Smaller class sizes would facilitate the success of instructional models such as PBE. Hattie (2012) demonstrated that class size reduction had an effect size of 0.10 to 0.20. Without proper funding, however, major initiatives such as this are destined to fail. As a result, educators fail to buy in to PBE because they believe it will not last, as was the case with past initiatives.

Therefore, educators seem to have developed their interpretations of PBE. Some participants focused on grading practices, while others introduced a system in which students move at their own pace through the curriculum. The study clearly showed that no educators were thoroughly focused on all aspects of a PBE model simply due to their limited capacity and inability to make structural changes to the school's educational model at the classroom level. As a result, the participants reported a lack of buy-in or logistical challenges.

Nevertheless, there was consensus on certain elements that should exist in a PBE system. The five most common elements were clear learning expectations, mastery drives movement, formative assessments, student voice, and differentiation of instruction. In the following subsections, these are individually reviewed in relation to the findings from this research.

Clear Learning Expectations. Perhaps the most essential element in the implementation of a PBE system is the creation of clear expectations for all (GSP, 2015; MCCL, 2015; Sturgis, 2016). This point was frequently mentioned by participants, specifically common standards and being clear with students about what constitutes proficiency. In addition, this clarity and the use of common language to refer to learning outcomes among staff emerged as havens in the research findings.

While there was clarity in the communication of proficiency expectations and common standards, there was not necessarily the case when it came to determining a score that exceeds proficiency. In addition, there was some inconsistency regarding teachers' use of common materials, such as assessments. Nevertheless, it could be said that participants viewed clear expectations as a positive outcome of PBE.

Mastery Drives Movement. The belief that students should only move on to the next standard once they have demonstrated proficiency on the current one forms the basis of proficiency-based graduation requirements (MDOE, 2015; GSP, 2015; MCCL, 2015; Sturgis, 2016). From the classroom observations and focus groups with teachers, it was clear that only one of the participating schools had fully embraced this concept. As mentioned in the findings on individual pace learning, this structure appeared to be an impediment to the incorporation of higher-order thinking opportunities in the classroom.

Furthermore, logistical challenges emerged for teachers who implemented this model when it came to retakes of summative assessments. Participants shared that students were frequently unprepared for the first assessment and given multiple opportunities to retake it. They believed that this structure lowered accountability and expectations.

Formative Assessments. Within a PBE system, there is a strong need for frequent formative assessments, as teachers must know how students are progressing towards the standard that they are attempting to master (GSP, 2015; MCCL, 2015; Sturgis, 2016). Participants shared multiple ways in which they used formative assessments, which were confirmed through the classroom observations. Homework remained a consistent element, although many teachers allowed time at the end of the class for students to complete their homework.

When working on practice problems in class, students would either receive instant feedback from an online system, be able to refer to answers in front of them, or check in with their teacher. On many occasions, educators were observed moving from student to student to provide support. This style of support was more frequently seen in middle schools than in high schools. Student work with the teacher engaged occurred 56.4% of the time in middle school observations and 32.3% of the time in high school observations.

Student Voice. Student voice refers to the belief that students should have some input in their learning, which may include contributions to learning design or learning pathways (GSP, 2015; MCCL, 2015). However, this element was not as present at the participating schools. When asked whether they had attempted to incorporate student voice, teachers did not necessarily believe that this would be helpful to students' learning.

Perhaps the most common example of student voice consists of students being allowed to choose how to exceed a standard. According to the participants, students could choose to support their peers or work on a specific project or set of problems. Other students were given the choice to move on to the next standard.

Differentiation of Instruction. When teachers differentiate instruction, they are ensuring that each student receives the support that they need to engage in learning (GSP, 2015; MCCL, 2015; Sturgis, 2016). Instruction may be differentiated on an individual basis or in small groups and based on content or learning style. This element was present at most participating schools, as students were given the opportunity to retake assessments and received individualized support from their teachers in preparation for the reassessments. At District 1 Middle School, students all worked at their own pace and the teacher provided individual support and answered questions based on their needs.

However, no clear differentiation based on learning style was found, although there were differences in the students who could access higher-order material between many schools. Once they attained proficiency, advanced students often had an opportunity to pursue additional learning avenues.

Higher-Order Thinking in the Classroom

In modern society, information is easily available. This is also true of mathematics; if an individual wants to know a basic mathematics fact, Google or Alexa can retrieve the answer in seconds. Wolfram Alpha and other web-based resources can even solve complex mathematical equations. For many students, memorizing how to implement an algorithm in mathematics class is more of an exercise in recall and executive functioning.

Career Focus. The current educational system has not incorporated developments from the past 10–15 years to update course requirements and offerings. Students who are interested in pursuing STEM fields still need to complete traditional mathematics courses and memorize differentiation rules and integration strategies in a senior calculus course. However, the skills needed to succeed in modern careers are ever-changing. Automation is present in most fields, and the basic skills once needed in many professions have been replaced by computer programs. These developments make the development of higher-order thinking in the classroom even more important than in the past.

Lack of Foundational Skills. Participants repeatedly said that they would love to challenge all students to work on complex problems in the classroom, but their foundational skills had to be developed first. If the foundational skills have not been developed yet, are there ways for students to use the tools that are available to the masses to help them approach complex thinking? This does not mean that the teaching of foundation skills is unimportant; rather, it should not be viewed as a roadblock to higher-order thinking opportunities in the classroom.

Opportunities to use HOTS are important for all students in modern middle school and secondary school mathematics classrooms. However, the current research suggests that the implementation of PBE models has not sufficiently fostered such opportunities. Models that only provide opportunities for the creative application of knowledge after students demonstrate proficiency with foundational skills creates that all but the highest achievers will be disengaged in the classroom. For youth to succeed in an ever-changing society, all students must have access to these important opportunities.

Mathematics-Based Observations. Most current classroom observation models require school administrators to look for levels of student engagement in the classroom. These models often connect engagement to the level of thinking taking place in the classroom. To outsiders, engagement in mathematics often looks like students working on problems. Administrators should not view mathematics any differently than they do any other content area. True engagement and higher-order learning happen when students share ideas, talk about why something works well or not, synthesize information to better understand a complex situation, or devise ways to visualize what they are learning. Thus, observations of mathematics classrooms should focus on the balance between the teaching of foundational knowledge and HOTS conversations and learning.

HOTS Frequency in PBE. This study did not find that the proficiency-based structure itself supported the incorporation of more frequent higher-order thinking opportunities for students. Among the participating schools, HOTS opportunities in the classroom were observed from 4% to 24% of the time. Students who worked at their own pace were less likely to engage in higher-order thinking opportunities. This may have been due to the logistics required for teachers to support and provide rigorous learning for all students when they needed it. Therefore, it would be beneficial in future research to focus on instructional models that require students to work individually and not receive direct instruction from a teacher, such as the flipped classroom model. Future research should examine the level of higher-order thinking that occurs in mathematics classrooms that use the latter.

Rigor Gap. Another important area for future research is the rigor gap observed between low-level and high-level mathematics students in middle school and high school. Lower-level students received considerable support in building their foundational skills in mathematics classes, while high-level students were encouraged to pursue more engaging learning opportunities that would challenge them to understand the content in greater depth. Thus, it would be valuable to examine this rigor gap in a longitudinal study to assess students' progress from elementary school to high school.

Implications for Practice

If mathematics teachers shifted their focus of practice to incorporate more HOTS opportunities in the classroom, this could impact the classroom routine for many students. Participants in this research either retained a traditional structure or adopted a learning structure based on the idea that mastery drives movement. In the former, the teacher reviews homework, teaches a new lesson, and give students practice time at the end of class. In the latter, students pick up where they left off at the beginning of class and move at their own pace.

Resource Development

Incorporating more HOTS opportunities in the classroom requires teachers to generate their own internal resources and highlight real-world connections in their unit planning. They also need to plan the facilitation of student conversations that require students to make predictions or explain their understanding of complex situations. This could impact the amount of time that teachers have to teach the content in their current mathematics curriculum. Such a change would push educators to focus only on the most essential material and collaborate on redefining standards before the beginning of the school year.

This shift would also make professional development and collaboration even more important. Developing true HOTS prompts for students is challenging and lies outside most educators' experience. In most cases, the textbook provides all worksheets and problems that teachers use in class.

Analytical Thinking and Engagement

Why is it important for teachers to focus on HOTS in the mathematics classroom prior to mastery of the basics? The answer is twofold. First, an increase in opportunities for higher-order thinking would increase students' capacity to think analytically and in turn their capacity to truly understand the concepts that they are learning. Secondly, with higher-order thinking comes the opportunity for greater relevance and with that engagement. Students who understand why a concept is important to learn are more likely to fully engage in learning.

It is not uncommon for mathematics teachers in middle schools and high schools to hear students express their frustrations and dislike of mathematics class. They must memorize rules that they do not fully understand, practice the application of these rules and algorithms on worksheets, and try to replicate these on an assessment. Teachers use different techniques to help students memorize rules that they must apply. For example, when introducing geometric concepts related to the coordinate plane, a teacher may require students to memorize the midpoint or distance formula to determine the length or center of line segments. Students may see these formulas and learn to plug in the numbers that correspond to the variables in the equations, but they do not always fully understand why they work. Thus, it is important for educators to not only help a student memorize and repeat but also truly understand it.

Another way that mathematics teachers might introduce the midpoint formula is to give students an open-ended question without a formula. The teacher may begin the class by giving students a coordinate plane with a line segment drawn on it. The ordered pair that represents the midpoint of the segment is given to students, who are asked to identify the ordered pair that represents the two end points. Without a formula, they must rely on their own thinking to determine the correct answer. Some students might count spaces horizontally and vertically;

others may try to measure the distance by hand and work backwards. Some students may even derive the midpoint formula on their own.

Thus, these students would have an opportunity to develop a clear understanding of the problem. Then, they could share their various approaches with the class and realize that there are multiple ways to obtain the same answer. Moreover, they could discuss why some answers are incorrect and challenges that emerged during the problem-solving process. After the learning activity is complete, the teacher could provide the formula to students to simplify the problem-solving process for future applications. Notice the difference between these two approaches. Ultimately, all students are given the formula. However, in the second approach, students develop a much more complete conceptual understanding of the coordinate plane and learn to appreciate the formula's usefulness.

Real-World Connections and Relevance

Ensuring that the learning material has real-world connections or relevance is another way to make HOTS accessible to students in the mathematics classroom. Real-world applications also make mathematics classes much more engaging. However, the greatest challenge for teachers is the extra work that designing such lessons entails.

That being said, it is unnecessary to design a hands-on, real-world lesson that incorporates rigor and relevance for each class. Practice remains very important, and not all learning must incorporate HOTS. According to IPI research, HOTS opportunities tend to occur 18–25% of the time in typical middle school and high school classrooms. Teachers should aim for the high end of this range and target higher-order tasks approximately one third of the time.

So, how is it possible to make mathematics class more relevant? As in the last example, students should simply be given a formula and asked to memorize and practice it. For instance, if high school students are learning applications of the Pythagorean theorem, the teacher could take them to the school's baseball field, which features many right triangles and lengths that could be measured. Students might be given a measuring tape and asked to gauge the distance between bases. They could then use their knowledge of the formula to make predictions and calculate various lengths. They could even be asked to recalculate if the distance between bases was something other than 90 feet. Once educators devise ways to make learning more relevant to students, the possibilities are endless. When students go home at the end of the day and their parents asked them what they learned, they can explain that they moved bases around on the baseball field using the Pythagorean theorem. Such experiences are memorable for students and have the potential to transform their perceptions of mathematics class. It would be difficult to argue that repetition and worksheets would yield better results than a more hands-on approach.

Whole-Group Learning

Although the proficiency-based diploma statute has been repealed in Maine, schools across the state retain remnants of its implementation. In many cases, the mastery-drives-movement component of a PBE system remained in a flipped classroom model. In this structure, material is reviewed prior to class and applied during class time. Learning is facilitated through videos or other forms of instruction outside of class and through teacher support in class. A meta-analysis of studies on the flipped classroom model showed that this structure had a small positive effect on student learning and that the level of teacher support for students did not decrease under it (van Alten et al., 2019).

If teachers decide to adopt the flipped classroom model for mathematics classes, research shows that it is important to incorporate ways for the entire class to regroup to engage in higher-order learning opportunities. According to the results from this study, individualized models in which students work at their own pace drastically impacted the number of higher-order learning opportunities.

Most focus group participants believed that students had to show competency with foundational skills before taking part in higher-order learning opportunities. However, if all classes were designed in this manner, many average and lower-achieving mathematics students would not have the opportunity to participate in engaging discussions and projects. Thus, teachers should try to ensure that all students can access higher-order thinking opportunities at their own level. For example, educators at District 3 High School used strategies such as open middle problem solving, in which students must describe their thinking and collaborate with peers, to design instructional practices. Such opportunities should not be limited to advanced students.

Policy Implications

Currently, the mathematics education guiding documents provided by the state of Maine make clear references to higher-order thinking in the classroom. Multiple standards in the Maine State Learning Results are designed to help students build towards using the learning content to solve real-world or practical problems. In addition, the state's guiding principles for education require students to become creative and practical problem solvers. The Department of Education has also unpacked what this may look like for beginners through emerging experts (MDOE, 2015).

HOTS at Earlier Ages

The current research demonstrates that these aspects of mathematics education considerably differed from school to school at the district level. Despite the resources, the provision of HOTS opportunities in the classroom is largely determined by the teacher's assessment of a student's readiness level and ability. In addition, the research found that, within PBE mathematics classrooms, high school students are exposed to more opportunities for higher-order thinking than middle school students. Again, this may be because there is more of a focus on foundational skills.

This study's results show that the importance of the guiding principles of Maine's learning system should be emphasized for earlier ages. Currently, high school students must demonstrate attainment of the guiding principles before graduation. Thus, one way to potentially incorporate more HOTS opportunities at the middle school level would be to require students to demonstrate some fulfillment of the principles before leaving eighth grade. This requirement would also establish a foundation for the work that they must take on in high school and provide students with similar opportunities across schools.

Decrease Class Sizes

To begin changing instructional practices at the middle school level, educators must have the flexibility to work with students more personally. Typical class sizes in middle school and high school currently approach 25 or more students. As a result, the average educator struggles to engage all students in higher-order thinking opportunities if there is little space for authentic learning conversations. When policymakers consider school funding, they must be mindful that more teachers are needed at schools if they want to make the greatest impact on students.

Class sizes are not the only factor that currently impedes HOTS opportunities in the classroom. The current factory model of education drives mathematics educators to continue teaching the way that they have taught in the past: they instruct students in the classroom, then give them a homework assignment. During PBE implementation, most teachers did not attempt to change this structure. Instead, they paid attention to assessments and grading. The only participating school that allowed students to fully move at their own pace was impacted by the logistics that governed student movement from grade to grade. If educators want students to learn differently, they must think of ways to change the current structure, which has been in place since the days of Horace Mann. They must think about what is needed for students to have authentic learning experiences rather than trying to fit such experiences into the current structure.

HOTS in Assessments

Another factor that impacts what is taught in the classroom is the state assessment test. Communities, administrators, teachers, and stakeholders judge a school's success by the success of students on statewide assessments. The latter drive the curriculum and promote breadth over depth, as teachers work hard to get in all of the information that will appear on the assessment. Assessments such as Smarter Balanced push for students to incorporate greater levels of HOTS to answer questions, and they even cross walked the assessment with Webb's depth of knowledge. However, students were not very successful on this assessment, and Maine moved away from it relatively quickly. If students do not have these experiences in class, they will struggle on the assessment.

HOTS Focus in Teacher Evaluations

At the administrative level, principals are tasked with supporting the teaching staff through professional growth and evaluations. When they visit a classroom, administrators examine factors such as student engagement and evidence of learning. Additionally, they ensure that learning targets are clear for all students and that there is appropriate communication. The use of formative assessment is also closely monitored. When administrators review engaged learning, it is essential that they provide teachers with feedback to help them integrate higher-order thinking opportunities in their lesson structure. Regarding supervision and evaluation tools, school district leaders should focus on tools that incorporate a measure of rigor, real-world connections, and higher-order thinking in the classroom.

Limitations

This study's mixed-methods design enabled data collection over a period of two years and conversations with mathematics teachers in a focus group setting. A potential limitation is that, over the course of this two-year period, schools may have begun to change their approach to instruction in a PBE model. For example, a district that originally received support from a regional cohort moved away from it, thus changing their approach to mathematics standards and the mathematics curriculum. At another school, students were grouped into ability-based classes between the first year and the second year. Such changes impacted the year-to-year consistency of the results. However, the use of focus groups allowed the researcher to better understand changes in the quantitative data.

Another impact of the two-year timeframe was that one middle school dropped out of the research after the first round. Although the data could still be used for the study, this led to an uneven number of observation points and focus groups between middle schools and high schools. Nevertheless, there was enough data to conduct a quantitative analysis.

Moreover, it would have been beneficial to gain a clearer understanding of whether challenges reported by participants were also encountered by other educators across Maine. The participants worked in educational systems that were developed in three different districts. In many cases, they expressed dissatisfaction with the PBE structure that had been recently implemented at their schools. A survey of other educators across the state could have provided a clearer picture of whether these challenges were widespread.

Conclusion

When Maine moved to proficiency-based diplomas for all students, school districts were forced to develop a system to ensure that all students attained proficiency in each designated content area standard before graduation (MDOE, 2015). The nature of this system suggested that there was a sequential progression to learning in which foundational skills had to be mastered prior to attempting more complex problems. Additionally, this structure required greater accountability for all but especially high schools, which were the gatekeepers of graduation.

Due to the sequential nature of PBE, a question emerged about whether all students would have opportunities to engage in higher-order thinking in the mathematics classroom or whether the focus would remain on foundational skills for some students. Thus, the following research questions were used to guide the study.

1. How frequently are mathematics students engaged in high-order thinking in proficiency-based classrooms?

   a. Are there significant differences across professional support styles in the frequency of higher-order thinking opportunities?

   b. Are there significant differences across grade spans (6-8, 9-12) in the frequency of higher-order thinking opportunities?

2. How and in what ways are mathematics students engaged in higher-order thinking when taught in a proficiency-based education system?

3. What factors enhance or impede educators' ability to implement higher-order thinking tasks in proficiency-based mathematics classrooms?

This research examined the nature of the learning that took place at participating middle schools and high schools after they had implemented their own version of a proficiency-based model. This explanatory sequential study included two rounds of research conducted over two academic years. In total, six schools (three middle schools and three high schools) were part of the research. In each round of research, 50 classroom observations were conducted at each school. Then, focus groups were held with mathematics educators from the participating schools. In total, there were 11 focus groups; this number is odd because one of the schools dropped out of the study after the first round of research.

Classroom observations were recorded using the IPI tool, which allowed the researcher to quantify the level of higher-order thinking taking place in each classroom on a scale of 1 to 6. The tool provided a continuum in which higher scores indicated a greater level of engaged learning. Thus, a score of 5 or 6 indicated that higher-order thinking was taking place.

On a quantitative level, there were two major findings. The first was that the high schools had a higher mean IPI score than the middle schools. High schools recorded a mean IPI score of 3.78 over 300 observation points, while middle schools recorded a mean IPI score of 3.44 over 250 observation points. This difference was determined to be statistically significant through a two-sample *t*-test. At the high school level, 18% of the observation points indicated higher-order thinking in the classroom. At the middle school level, 13.6% of observation points indicated higher-order thinking. On average, schools fall in the range of 18–25% on the IPI.

The second finding was that there was a statistically significant difference in mean IPI scores between two districts. District 3 had a mean IPI score of 3.87, while District 1 had a mean IPI score of 3.42. All three districts were examined using an analysis of variance. In District 3, HOTS opportunities occurred at a rate of 24%, while only 11% of classroom observations in District 1 indicated higher-order thinking. This data was later interpreted using insights from the focus group conversations. For example, schools in District 1 were impacted by the individualized nature of the PBE program, and HOTS opportunities in District 3 were observed when teachers specifically planned them for the entire class.

The qualitative research findings can be divided into three themes. The first theme focuses on the process of moving away from standardized learning and instruction. Instruction and learning practices were discussed in each focus group. Certain themes emerged from these conversations, namely the components of an instructional model. The most common themes were (1) grading practices, (2) assessment, (3) standards, (4) student supports, and (6) individualized pace learning. Participants discussed their schools' move to a 1–4 grading system and the impact that this had on their instructional practices. Teachers expressed that their feelings about the grading system were mixed from the school administration.

Moreover, participants discussed their use of summative and formative assessments. They mentioned the challenges associated with assessment retakes and shared their concerns that accountability measures were reduced during the transition. However, there seemed to be some consistency in approach in the area of standards; teachers said that standards were used to drive instruction and that there was clarity about standards in mathematics.

Participants understood the need to provide individualized support for students who were working to attain proficiency on standards. They shared systems that were in place for individual check-ins, afterschool support, and peer support. Individual pace learning differed from school to school. At District 1 Middle School, students moved entirely at their own pace through the grade level content. At other schools, the summative assessment process enabled individualization through reassessments and individual support.

The second theme in the qualitative findings were the havens and hazards of PBE as defined by mathematics educators. the havens can be described as (1) increased clarity and communication impacting learning, (2) relationship building and student engagement, and (3) teacher-directed instruction. Participants were very clear on a key point: the move to a PBE system helped them improve the clarity of learning standards and expectations for students. They now used a common language to communicate with students. In some cases, teachers believed that this change helped them better identify the needs of students and had a positive impact on student learning.

The individualized nature of the PBE model also seemed to have a positive impact on relationships between teachers and students, which meant that students were more engaged while receiving support from their teacher. Participants shared strategies that they used to ensure that all students attained proficiency, which led to relationship building with students. A common finding across the focus groups was that participants believed teacher-directed instruction to be quite valuable. They said that it was important for instructional models to continue including this component and not leaving it up to students to learn on their own in an individualized learning structure.

Regarding the hazards of PBE, participants mentioned (1) negative student outcomes, (2) failure to build support for PBE, and (3) the paradigm shift to PBE. The teachers gave multiple examples of less-than-desirable student outcomes resulting from the new learning practices. For instance, students tended to memorize material for the assessment to attain proficiency but were unable to recall what they had learned later. The teachers reported struggles with the individualized component of PBE and a level of apathy that had grown among students due to a perceived lack of accountability.

Through the focus group, it was clear that some participants had not bought into their school's shift to a PBE structure and that this impacted pedagogical practices. Teachers shared that there were differences in opinion between themselves and the administration and that parents did not fully understand the PBE system. Some educators only made surface-level changes to their own practices, as they did not fully support the system.

As the paradigm shift took place within school districts, certain grade levels were asked to complete the work first. This led to differences in the allocation of professional development opportunities and resources. Participants expressed their frustrations about this situation and explained that simple tasks such as reassessing students could impact the structure of their class and the resources that were utilized.

The third theme in the qualitative findings was the continuum of instructional practices and learning opportunities. This continuum of pedagogy was centered around the six categories that comprise the IPI. These were (1) complete disengagement, (2) student work with teacher not engaged, (3) student work with teacher engaged, (4) teacher-led instruction, (5) student learning conversations, and (6) student active engaged learning.

Participants explained that they often saw a lack of engagement from students at the end of class, when they were given time to complete homework or practice problems. Some teachers reported that high school students were often left to their own devices and that many chose to not do their mathematics homework during this time. In some cases, students worked individually with a teacher not engaged during homework time; in other cases, they were allowed to work on their own in an individual pace learning plan.

On many occasions, students worked on their own, but the teacher had a plan to check in with them and ensure that they were supported throughout the class. Educators walked from student to student or called students up to their desk to review problems. They also worked with small groups of students on common practice material. The practice of whole-group direct instruction remained very common at most of the participating schools. Sometimes, the teacher introduced a topic to the class and students had an opportunity to practice it and ask questions to the instructor.

Moreover, participants shared many examples of higher-order thinking in the mathematics classroom. They said that truly meaningful learning conversations could happen with students of all levels through open middle problems that required them to generate multiple responses and explain their thinking. Teachers also spoke about opportunities to undertake projects that helped increase student engagement levels. They explained that not all students typically pursued work that corresponded to Level 4 on the proficiency-based rubric, which exceeds the standard and is typically linked with a greater frequency of higher-order thinking opportunities. These tasks were often left for more advanced students who had already demonstrated their proficiency.

The most important takeaways from this research are two-fold. First, when individual pace learning structures are incorporated into PBE classrooms, students do not have as many opportunities to engage in higher-order thinking. Although these structures allow students to move at their own pace, they often involve students reviewing algorithms on their own, then replicating them on assessments to demonstrate their proficiency. Teachers shared some real concerns when it comes to accountability when it comes to preparation for assessments and pacing when retakes are allowed without accountability measures in this type of structure.

The research also demonstrated a clear disparity in higher-order thinking opportunities for students in proficiency-based mathematics classes. Advanced students tended to have more opportunities to examine the content in greater depth, as teachers believed that it was important for students to develop foundational skills to succeed at a higher-order task. However, participants demonstrated that was not necessarily the case and that students of all levels could be truly engaged in learning if they were empowered to contribute to learning conversations.

Maine has now moved away from proficiency-based graduation requirements, but many traces of the PBE system remain in classrooms throughout the state. Many middle schools are now committed to a 1–4 grading rubric, and high schools have curricula that are still driven by common standards. Although a benefit of this change is a clear increase in clarity for students, teachers, and families, challenges are likely to arise if educators are still teaching in a system that they never fully bought in to. A goal for all students to be proficient in all standards the curriculum is noble. The question is, what does it mean to be proficient? Is proficiency in a mathematics standard tied to the remembering and understanding level of Bloom's taxonomy, or is it tied to the analysis level? If this is where proficiency is rooted, then mathematics students will not have regular opportunities to synthesize, evaluate, and create.

The typical middle school or high school classroom engages in higher-order thinking opportunities 18–25% of the time. In this study, mathematics classrooms scored in the range of 4–24% overall (Valentine, 2007). As mathematics teachers create their unit and lesson plans, they must be mindful of HOTS and not allow their framework dictate available HOTS opportunities for students. Higher-order thinking opportunities can take the form of engaging entry activities for the daily lesson, process thinking, group discussions, or examining the learning content in greater detail; they do not need to be a standalone project. If projects are the only opportunity being provided, then HOTS become an afterthought.

In today's educational environment, teachers frequently express how overburdened they feel and that new initiatives feel like "one more thing." Therefore, it is important for administrators to find time for teachers in the regular professional development calendar to plan and discuss teaching, learning, and higher-order thinking. The need to develop resources for use in the classroom can be a considerable deterrent for teachers, as this requires additional work and planning beyond the foundational practice problems that are traditionally used. Thus, teachers must support and share with each other. The development of HOTS resource banks would be a significant support for teachers who want to incorporate higher-order thinking opportunities in their instruction but feel burdened by time.

As a former high school mathematics teacher and school principal, parents often told me that they were "never good at math." This sentiment was also expressed by countless students. It always hurts to hear a student say that mathematics was their least favorite subject. Mathematics educators and educational leaders alike must work together to dismantle this stigma. Students will enjoy mathematics if they are engaged in their learning and asked to think in ways that connect their learning to topics that matter to them. They will also enjoy mathematics if they are all given the opportunity to participate in a classroom conversation about questions that are accessible to all. It will not be easy at first, but mathematics teachers must think of ways to engage all students in HOTS opportunities in their classes more than a quarter of the time. If this happens, I am certain that we will hear the "never good at math" claim much less and that there will be many more smiles in mathematics classrooms.

REFERENCES

ACT, Inc. (2012). *Raising the bar: A baseline for college and career readiness in our nation's high school core courses*. Iowa City, IA: ACT Inc.

AP Calculus AB The Course. (2016). Retrieved October 1, 2017, from https://apcentral.collegeboard.org/courses/ap-calculus-ab/course?course=ap-calculus-ab

AP Program Participation and Performance Data 2017 – Research – The College Board. (2017). Retrieved October 1, 2017, from https://research.collegeboard.org/programs/ap/data/participation/ap-2017

Akınoğlu, O., & Karsantık, Y. (2016). Pre-Service Teachers' Opinions on Teaching Thinking Skills. *International Journal of Instruction,9*(2), 61-76.

Arbaugh, F., Herbel-Eisenmann, B., Ramirez, N., Knuth, E., Kranendonk, H., & Quander, J. (2008). *Linking Research and Practice* (Rep.). National Council of Teachers of Mathematics.

Bandura, A. (1977). Social Learning Theory. New York: General Learning Press.

Behaviorism. (2016, September 02). Retrieved May 21, 2017, from https://www.learning-theories.com/behaviorism.html

Black, P., & Wiliam, D. (2010). "Kappan Classic": Inside the Black Box--Raising Standards through Classroom Assessment. *Phi Delta Kappan, 92*(1), 81-90.

Barton, P., & Coley, R. (2009). *Parsing the Achievement Gap II* (Rep.). Princeton, NJ: Education Testing Services.

Basari, H., Purwanto, Rahman As'ari, A., & Sisworo. (2019). Investigating Critical Thinking Skill of Junior High School in Solving Mathematical Problem. *International Journal of Instruction, 12*(3), 745-758.

Bay-Williams, J., Duffett, A., & Griffith, D. (2016). *Common Core Math in the K-8 Classroom: Results from a National Teacher Survey* (Rep.). Washington, DC: Thomas B. Fordham Institute.

Blackburn, B. R. (2013). *Rigor is not a four letter word*. Larchmont, NY: Eye On Education.

Bloom, B. S. (1956). *Taxonomy of educational objectives. vol. 1: Cognitive domain*. New York: McKay.

Brodersen, R. M., Yanoski, D., Mason, K., Apthorp, H., and Piscatelli, J. (2016). *Overview of selected state policies and supports related to K–12 competency-based education* (REL 2017– 249). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Central. Retrieved from http://ies.ed.gov/ncee/edlabs.

Brookhart, S.M. (2016). Start with higher-order thinking. *Educational Leadership*, 7(2), 10-15.

Brookhart, S. M., Long, B. A., & Moss, C. M. (2011). Know your learning target. *Educational Leadership,68*(6), 66-69.

Brown, R. E., & Milner, P. M. (2003). The legacy of Donald O. Hebb: more than the Hebb Synapse. *Nature Reviews Neuroscience, 4*(12), 1013-1019.

Burris, C. C., Wiley, E., Welner, K., & Murphy, J. (2008). Accountability, Rigor, and Detracking: Achievement Effects of Embracing a Challenging Curriculum as a Universal Good for All Students. *Teachers College Record,110*(3), 571-607.

Cain, J. S. (2002). An evaluation of the connected mathematics project. *Journal of Educational Research, 95*(4), 224-33.

Cognitivism. (2015, February 04). Retrieved May 21, 2017, from http://www.learning-theories.com/cognitivism.html

Costa, A., & Kallick, B. (2015). Five Strategies for Questioning with Intention. *Educational Leadership, 73*(1).

Creswell, J. W. (2013). *Qualitative inquiry and research design: choosing among five approaches*. Los Angeles: SAGE Publications.

Creswell, J. W., & Plano Clark., V. L. (2011). *Designing and conducting mixed methods research*. Los Angeles: SAGE.

Croft, M., Guffy, G., & Vitale, D. (2014). *Communicating College and Career Readiness through Proficiency Standards* (Issue brief). Iowa City, IA: ACT, Inc.

Daggett, W. R. (2017). The Rigor Relevance Framework. Retrieved June 27, 2017, from http://www.leadered.com/our-philosophy/rigor-relevance-framework.php

Davis, J. D., McDuffie, A. R., Drake, C., & Seiwell, A. L. (2019). Teachers' perceptions of the official curriculum: Problem solving and rigor. *International Journal of Educational Research, 93*, 91-100.

Dean, C. B., Hubbell, E. R., Pitler, H., & Stone, B. (2012). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.

Diploma requirements. (2021). Retrieved February 13, 2021, from https://www.maine.gov/doe/learning/diplomas

Direct instruction (2013, December 20). In S. Abbott (Ed.), The glossary of education reform. Retrieved from http://edglossary.org/direct-instruction

Donaldson, G. A. (2014). *From schoolhouse to schooling system: Maine public education in the 20th century*. Orono, ME: College of Education and Human Development, University of Maine.

Doorey, N., & Polikoff, M. (2016). *Evaluating the Content and Quality of Next Generation Assessments*(Rep.). Washington, DC: Thomas B. Fordham Institute.

Doyle, W., & National Commission on Excellence in Education. (1982). *Academic work.*

Eaton,S.E. (2017). Research assistant training manual: Focus groups. Calgary: University of Calgary.

Education Trust-West. (2011). *Unlocking doors and expanding opportunity: Moving beyond the limiting reality of college and career readiness in california high schools.* Oakland, CA: Education Trust-West.

Eliyasni, R., Kenedi, A. K., & Sayer, I. M. (2019). Blended Learning and Project Based Learning: The Method to Improve Students' Higher Order Thinking Skill (HOTS). *Jurnal Iqra', 4*(2), 231-248.

Engelmann, S. (1999). The Benefits of Direct Instruction: Affirmative Action for At-Risk Students. *Educational Leadership, 57*(1), 77-79.

Evans, C. M. & DeMitchell, T. A. (2018). Mapping out the terrain: Northeast principals' perceptions of the barriers and facilitators to implementing K-12 competency-based education (Policy Brief #18-02). Durham, NH: University of New Hampshire, Education Department, Division of Educational Studies.

Evans, C., Landl, E., & Thompson, J. (2020). Making sense of K‐12 competency‐based education: A systematic literature review of implementation and outcomes research from 2000 to 2019. *The Journal of Competency-Based Education, 5*(4).

Fitzgerald, M. S., & Palincsar, A. S. (2019). Teaching Practices That Support Student Sensemaking Across Grades and Disciplines: A Conceptual Review. *Review of Research in Education*, *43*, 227–248.

Fosnot, C. T. (2005). *Constructivism: Theory, perspectives, and practice*. New York: Teachers College Press.

Gine, R., & Kruse, D. (2007). What if Less Is Just Less? The Role of Depth over Breadth in the Secondary Mathematics Curriculum. *Horace, 23*(2).

Great School Partnership. (2015). *Proficiency-based learning simplified.* Retrieved June 20, 2015, from http://www.greatschoolspartnership.org/proficiency/research/

Guskey, T. R., Townsley, M., & Buckmiller, T. M. (2020). The Impact of Standards-Based Learning: Tracking High School Students' Transition to the University. *NASSP Bulletin, 104*(4), 257-269.

Haciomeroglu, E. S., & Chicken, E. (2012). Visual thinking and gender differences in high school calculus. *International Journal of Mathematical Education in Science and Technology*, *43*(3), 303–313. https://doi.org/10.1080/0020739x.2011.618550

Hadi, S., Retnawati, H., Munadi, S., Apino, E., & Wulandari, N. (2018). The difficulties of high school students in solving higher-order thinking skills problems. *Problems of Education in the 21st Century, 76*(4), 520-532.

Hammond, B. (2014). *Oregon schools' big switch to academics-only grades, with no penalty for late papers or missing homework, is kaput.* Retrieved July 28, 2015, from http://www.oregonlive.com/education/index.ssf/2014/04/oregon_schools_big_switch_to_a.html.

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

Hess, K., & Gong, B. (2014). *Ready for college and career? achieving the common core standards and beyond through deeper, student-centered learning*. Boston, MA: National Center for the Improvement of Educational Assessment.

Hess, K. K., Jones, B. S., Carlock, D., & Walkup, J. R. (2009). *Cognitive rigor: Blending the strengths of bloom's taxonomy and webb's depth of knowledge to enhance classroom-level processes* Online Submission.

Higgins, S., Hall, E., Baumfield, B., & Mosley, D. (2005, October). A meta-analysis of the impact of the implementation of thinking skills approaches on pupils. *Durham Research Online*. *How Many Schools and Districts Have Not Made Adequate Yearly Progress? Four-Year Trends* (Rep.). (2010). Washington, D.C.: Center on Education Policy.

Hunzicker, J., & Lukowiak, T. (2012). Effective Teaching and Student Engagement in the College Classroom: Using the Instructional Practices Inventory (IPI) as a Tool for Peer Observation and Self-Reflection. *Journal on Excellence in College Teaching,23*(1), 99-132.

Hwang, G., Yin, C., & Chu, H. (2019). The era of flipped learning: Promoting active learning and higher order thinking with innovative flipped learning strategies and supporting systems. *Interactive Learning Environments, 27*(8).

Johnson, A. F. (2019). *Maine's Superintendents' Perceptions of Proficiency-Based Education and Proficiency-Based Diploma Systems* (Rep.). Gorham, ME: Center for Education Policy, Applied Research, and Evaluation. University of Southern Maine.

Johnson, A., & Stump, E. (2018). *Proficiency-based High School Diploma Systems in Maine: Educator Perceptions of Implementation* (Rep.). Gorham, ME: Center for Education Policy, Applied Research, and Evaluation. University of Southern Maine.

Johnston, H. (2011). *Proficiency-based education. Research brief.* Education Partnerships, Inc.

Johnstone, S. M., & Soares, L. (2014). Principles for developing competency-based education programs. *Change: The Magazine of Higher Learning, 46*(2), 12-19.

Kelleher, J. (2016, September). Make introductory lessons opportunities for higher-level critical thinking. *Association for Middle Level Education Magazine*.

Kellinger, J. J. (2012). The flipside: Concerns about the "new literacies" paths educators might take. *Educational Forum, 76*(4), 524-536.

Kilpatrick, J. (2001). Understanding mathematical literacy: The contribution of research. *Educational Studies in Mathematics*, *47*, 101–116.

Kjos, R., & Long, K. (1994). Improving critical thinking and problem solving in fifth grade mathematics.

Krathwohl, D. R. (2002). A Revision of Bloom’s Taxonomy: An Overview. *Theory Into Practice, 41*(4), 212-218.

Krawec, J., Huang, J., Montague, M., Kressler, B., & de Alba, A. M. (2013). The effects of cognitive strategy instruction on knowledge of math problem-solving processes of middle school students with learning disabilities. *Learning Disability Quarterly, 36*(2), 80-92.

Larson, K., Guidera, A. R., Smith, N., Business Coalition for, E. R., & National Alliance, o. B. (1998). *The formula for success: A business leader's guide to supporting math and science achievement.*

LaVenia, M., Pineau, K., & Lang, L. (2010). *The Predictive Validity of Critical Thinking Disposition on Middle-Grades Math Achievement* (Rep.). Evanston, IL.

Lewis, D. (2020). Student Anxiety in Standards-based Grading in Mathematics Courses. *Innovative Higher Education, 45*, 153-164.

Linchevski, L., & Kutscher, B. (1998). Tell Me with Whom You're Learning, and I'll Tell You How Much You've Learned: Mixed- Ability versus Same-Ability Grouping in Mathematics. *Journal for Research in Mathematics Education*, *29*(5), 533–554.

Maine cohort for customized learning. (2015). Retrieved March 22, 2015, from http://mainecustomizedlearning.org.

Maine Department of Education. (2015). *What is proficiency-based education?* (2014). Retrieved March 22, 2015, from http://www.maine.gov.prxy4.ursus.maine.edu/doe/proficiency/about/proficiency-based.html.

Maine revised statutes §4722-A. proficiency-based diploma standards. (2014). Retrieved March 22, 2015, from http://www.mainelegislature.org/legis/statutes/20-A/title20-Asec4722-A.html.

Mahiroglu, A. (2007). *Teachers' opinions on students' higher order thinking skills.* Online Submission.

Marshall, J. C., & Horton, R. M. (2011). The relationship of teacher-facilitated, inquiry-based instruction to student higher-order thinking. *School Science and Mathematics, 111*(3), 93-101.

Marzano, R. J., & Kendall, J. S. (2007). *The new taxonomy of educational objectives*. Thousand Oaks, CA: Corwin Press.

Marzano, R. J., Pickering, D., McTighe, J., & Mid-Continent Regional, E. L. (1993). *Assessing student outcomes: Performance assessment using the dimensions of learning model.*

Mathematics Standards. (2016). Retrieved September 30, 2017, from http://www.corestandards.org/Math

Matsumura, L. C., Slater, S. C., & Crosson, A. (2008). Classroom climate, rigorous instruction and curriculum, and students' interactions in urban middle schools. *Elementary School* Journal, 108(4), 293-312.

Maine DOE Videos on Customized Learning « Competency Works. (2012). Retrieved October 30, 2017, from https://www.competencyworks.org/resources/must-watch-maine-doe-videos-on-customized-learning/

Mayer, R. E. (2002). Rote Versus Meaningful Learning. *Theory Into Practice, 41*(4), 226-232.

McGatha, M. B., & Bay-Williams, J. M. (2013). Making Shifts toward Proficiency. *Teaching Children Mathematics, 20*(3), 162.

Miller, S. D. (2003). How high and low-challenge tasks affect motivation and learning: Implications for struggling learners. *Reading & Writing Quarterly*, 19, 39–57.

Moon, T. R. (2009). Myth 16: High-Stakes Tests Are Synonymous With Rigor and Difficulty. *Gifted Child Quarterly,53*(4), 277-279.

Moore, J. (2011). Behaviorism. *The Psychological Method, 61*, 449-464.

Morgan, D. L. (2009). *Focus groups as qualitative research*. Thousand Oaks, CA: Sage Publ.

Mueller, M., Yankelewitz, D., & Maher, C. (2014). Teachers promoting student mathematical reasoning. *Investigations in Mathematics Learning*, *7*(2), 1–20.

O'Neil, J. (1999). Core Knowledge and Standards: A Conversation with E.D. Hirsch, Jr. *Educational Leadership, 56*(6), 28-31.

O'Neil, J. (2011). Hermeneutic Haunting: E. D. Hirsch, Jr. and the Ghost of Interpretive Validity. *Educational Studies, 47*(5), 451-468.

Paige, D. D., Sizemore, J. M., & Neace, W. P. (2013). Working inside the box: Exploring the relationship between student engagement and cognitive rigor. *NASSP Bulletin, 97*(2), 105-123.

Paige, D. D., Smith, G. S., & Sizemore, J. M. (2015). Conceptualizing rigor and its implications for education in the era of the Common Core. *Cogent Education,2*(1).

Piaget, J. (1951). *The Child's Conception of the World*. Lanham, MD: Littlefield Adams.

Phillips, M. (1976). The Application of Gestalt Principles in Classroom Teaching. *Group & Organization Management,1*(1), 82-98.

Porter, A. C. (2002). Measuring the Content of Instruction: Uses in Research and Practice. *Educational Researcher, 31*(7), 3-14.

Pólya, G. (2004). *How to solve it: A new aspect of mathematical method*. Princeton, NJ: Princeton University Press.

Principles to Actions Executive Summary (pp. 1-6, Rep.). (2016). Reston, VA: e National Council of Teachers of Mathematics.

Re-inventing Schools Coalition. (2015). *What is competency-based education?* Retrieved March 22, 2015, from http://www.reinventingschools.org/resources/what-is-competency-based-education/.

Rier, J. (2014). *Extension options ensure proficiency-based diploma done right.* Retrieved July 28, 2015, from http://mainedoenews.net/2014/05/29/extension-options-ensure-proficiency-based-diploma-done-right/.

Ritchhart, R., Church, M., & Morrison, K. (2011). *Making Thinking Visible: How to Promote Engagement, Understanding, and Independence for All Learners*. San Francisco, CA: Jossey-Bass.

Roessingh, H., & Chambers, W. (2011). Project-Based Learning and Pedagogy in Teacher Preparation: Staking out the Theoretical Mid-Ground. *International Journal of Teaching and Learning in Higher Education, 23*(1).

Saldaña, J. (2015). *The coding manual for qualitative researchers*. Los Angeles: SAGE.

Scarlett, M. H. (2018). "Why did I get a C?": Communicating Student Performance Using Standards-Based Grading. *InSight: A Journal of Scholarly Teaching, 13*, 59-75.

Schneider, M., & American Enterprise Institute for Public Policy Research. (2009). *Math in american high schools: The delusion of rigor. education outlook. no. 10* American Enterprise Institute for Public Policy Research.

Schraw, G. J., & Robinson, D. H. (2011). *Assessment of higher order thinking skills*. Charlotte, NC: Information Age Pub.

Schultz, H., & FitzPatrick, B. (2016). Teachers' Understandings of Critical and Higher Order Thinking and What This Means for Their Teaching and Assessments. *Alberta Journal of Educational Research,62*(1), 61-86.

Schwahn, C., & McGarvey, B. (2011). *Inevitable mass customized learning in the age of empowerment* CreateSpace.

Scriven, M., & Paul, R. (1987). Defining Critical Thinking. Retrieved June 23, 2017, from http://www.criticalthinking.org/pages/defining-critical-thinking/766

Shakman, K., Foster, B., Khanani, N., Marcus, J., & Cox, J. (2018). *"In Theory It's a Good Idea": Understanding Implementation of Proficiency-Based Education in Maine* (Rep.). Waltham, MA: Education Development Center Inc.

Silver, E. (1995). *Shuffling the Deck to Ensure Fairness in Dealing: A Commentary on Some Issues of Equity and Mathematics Education from the Perspective of the QUASAR Project* (Tech.).

Silver, E. A., & Smith, M. S. (1996). Building discourse communities in mathematics classrooms: A worthwhile but challenging journey. *Communication in Mathematics, K-12 and Beyond*.

Silver, E., & Stein, M. K. (1996). The QUASAR Project The "Revolution of the Possible" in Mathematics Instructional Reform in Urban Middle Schools. *Urban Education,30*(4), 476-521.

Silvernail, D., Stump, E., McCafferty, A., & Hawes, K. (2014). *Implementation of a proficiency-based diploma system: phase II*. (Rep.). Gorham, ME: Center for Education Policy.

Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching in the Middle School, 12*(2), 88-95.

Smith, M., & Stien, M. K. (1998, February). Selecting and Creating Mathematical Tasks: From Research to Practice. *Mathematics Teaching in the Middle School, 3*(5), 344-350.

Stiggins, R., Arter, J., Chappuis, J., & Chappuis, S. (2007). *Classroom assessment for student learning*. Portland, OR: Educational Testing Service.

Stump, E., Doykos B., Fallona, C. (2016). *Proficiency-based high school diploma systems in Maine: local implementation of state standards-based policy phase IV.* (Rep.). Gorham, ME: Center for Education Policy.

Stump, E., & Silvernail, D. (2015). *Proficiency-Based Diploma Systems in Maine: Implementing District-Level High School Graduation Policies Phase III Technical Report* (Rep.). Gorham, ME: Center for Education Policy.

Stump, E., & Silvernail, D. (2014). *Implementation of a proficiency-based diploma system* (Rep.). Gorham, ME: Center for Education Policy.

Sturgis, C. (2016). *Reaching the Tipping Point Insights on Advancing Competency Education in New England* (Rep.). Vienna, VA: INACOL.

Talebi, K. (2015). John Dewey--Philosopher and Educational Reformer. *European Journal of Education Studies, 1*(1), 1-13.

Tambunan, H. (2018). Impact of Heuristic Strategy on Students' Mathematics Ability in High Order Thinking. *International Electronic Journal of Mathematics Education, 13*(3), 321-328.

Tambunan, H. (2019). The Effectiveness of the Problem Solving Strategy and the Scientific Approach to Students' Mathematical Capabilities in High Order Thinking Skills. *International Electronic Journal of Mathematics Education, 14*(2), 293-302.

Tanujaya, B., & Mumu, J. (2020). Students' misconception of HOTS problems in teaching and learning of mathematics. *Journal of Physics: Conference Series, 1657*.

Thompson, C. J. (2008). Preparation, practice, and performance: An empirical examination of the impact of standards-based instruction on secondary students' math and science achievement. *Research in Education, 81*(1), 53-62.

Tomlinson, C. A. (2016). *The differentiated classroom: responding to the needs of all learners*. Boston: Pearson Education, Inc., by special arrangement with the Association for Supervision and Curriculum Development (ASCD).

Torff, B. (2014). Folk Belief Theory, the Rigor Gap, and the Achievement Gap. *The Educational Forum,78*(2), 174-189.

Torff, B. (2008). Using the Critical Thinking Belief Appraisal to assess the rigor gap. *Learning Inquiry,2*(1), 29-52.

Torff, B., & Sessions, D. (2006). Issues Influencing Teachers' Beliefs about Use of Critical-Thinking Activities with Low-Advantage Learners. *Teacher Education Quarterly,33*(4), 77-91.

Valentine, J. (2005). *Instructional Practices Inventory: A Process for Profiling Student Engaged Learning for School Improvement*. Manuscript, Middle Level Leadership Center.

Valentine, J. (2007). *The instructional practices inventory: Using a student learning assessment to foster organizational learning.* University of Missouri, Columbia, MO: Middle Level Leadership Center.

Van Alten, D. C., Phielix, C., Janssen, J., & Kester, L. (2019). Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. *Educational Research Review, 28*.

Vermont Agency of Education. (2015). *Proficiency based learning.* Retrieved July 28, 2015, from http://education.vermont.gov/proficiency-based-learning.

Vygotsky, L. (1978). Interaction between learning and development. Readings on the development of children, 23(3), 34-41.

Wade, J., Wolanin, N., & McGaughey, T. (2015). *A Study of Critical Thinking Skills in the International Baccalaureate Middle Years Programme* (Rep.). Washington, D.C.

Wagner, T. (2008). Rigor redefined. *Educational Leadership, 66*(2), 20-25.

Walters, K., Smith, T., Ford, J., & Scheopner Torres, A. (2014). (rep.). *Gearing up to teach the Common Core State Standards for Mathematics in the rural Northeast Region*. Washington, DC: Regional Educational Laboratory Northeast & Islands.

Webb, N. L. (2007). Issues related to judging the alignment of curriculum standards and assessments. *Applied Measurement in Education, 20*(1), 7-25.

Wilen, W. W. (2004). Refuting Misconceptions about Classroom Discussion. *The Social Studies,95*(1), 33-39.

Yi, H. S., & Lee, Y. (2016). A latent profile analysis and structural equation modeling of the instructional quality of mathematics classrooms based on the PISA 2012 results of Korea and Singapore. *Asia Pacific Education Review,18*(1), 23-39.

Yuan, K., & Le, V. (2014). *Measuring Deeper Learning Through Cognitively Demanding Test Items: Results from the Analysis of Six National and International Exams* (Rep.). Santa Monica, CA: RAND Corporation.

Zohar, A., & Dori, Y. J. (2003). Higher Order Thinking Skills and Low-Achieving Students: Are They Mutually Exclusive? *Journal of the Learning Sciences,12*(2), 145-181.

APPENDIX A

Math Teacher Focus Group - Guiding Questions

1. Describe how Maine's move to a proficiency based diploma has impacted you in your role as a mathematics teacher.

    a) What has been the most beneficial result?

    b) What has been the greatest resulting challenge?

    c) Based on the move towards proficiency-based instruction, describe the role of formative and summative assessments in your classroom.

2. Describe the role that differentiation of instruction or learning pathways play in your classroom.

    a) How does a student's individual pace drive instruction in your classroom?

    b) If students cover topics at their own pace, what does the structure look like in your classroom?

    c) Once a student has achieved a "3" (or attained proficiency) on a learning target can they move on to the next target?

    d) What role does student choice play in your unit planning?

3. Describe the structure of a "typical" math class.

    a) What role does memorization or recollection of facts play in your classroom?

    b) What roles do identification of patterns and estimation play in your classroom?

    c) What role does student feedback play in your teaching?

4. Describe how you utilize projects that ask students to analyze or synthesize information.

a) Describe the students that typically take part in this.

b) Describe when these activities occur.

c) Describe what happens before and after this type of activity.

5. What are some examples of times that students have been asked to develop a logical argument or formulate a hypothesis?

a) Describe the students that typically take part in this?

b) Describe when these activities occur.

c) Describe what happens before and after this type of activity?

6. Overall, how would you assess student engagement levels in the typical math class?

a) At what times do students seem most engaged? Why do you feel this is the case?

b) At what times do students seem least engaged? Why do you feel this is the case?

c) Describe the role that student voice and choice plays in your classroom

APPENDIX B

Instructional Practices Inventory Categories

Instructional Practices Inventory Categories

Broad Categories	Coding Categories	Common Observer "Look-Fors"
Student-Engaged Instruction	Student Active Engaged Learning (6)	Students are engaged in higher-order learning. Common examples include authentic project work, cooperative learning projects, hands-on learning, problem-based learning, demonstrations, and research.
	Student Learning Conversations (5)	Students are engaged in higher-order learning conversations. They are constructing knowledge or deeper understanding as a result of the conversations. Common examples are cooperative learning, work teams, discussion groups, and whole-class discussions. Conversations may be teacher stimulated but are not teacher dominated.
Teacher-Directed Instruction	Teacher-Led Instruction (4)	Students are attentive to teacher-led learning experiences such as lecture, question and answer, teacher giving directions, and media instruction with teacher interaction. Discussion may occur, but instruction and ideas come primarily from the teacher. Higher order learning is not evident.
	Student Work with Teacher Engaged (3)	Students are doing seatwork, working on worksheets, book work, tests, video with teacher viewing the video with the students, etc. Teacher assistance, support, or attentiveness to the students is evident. Higher-order learning is not evident.
Disengagement	Student Work with Teacher not Engaged (2)	Students are doing seatwork, working on worksheets, book work, tests, video without teacher support, etc. Teacher assistance, support, or attentiveness to the students is not evident. Higher-order learning is not evident.
	Complete Disengagement (1)	Students are not engaged in learning directly related to the curriculum.

Note. From "The instructional practices inventory: Using a student learning assessment to foster organizational learning" by J. Valentine, 2007, University of Missouri, Columbia, MO: Middle Level Leadership Center, p. 5.

APPENDIX C

IPI Reliability Rating

IPI Reliability Assessment Results Inbox x

Valentine, Jerry W. (Emeritus) <ValentineJ@missouri.edu> Sep 27
to me

Email Address: qdonahue@rsu3.org

Name: Quinton Donahue

Location: Univ. of Maine

Date: 9/21/2017

Listed below are your results for the Reliability Assessment you completed at the conclusion of your Instructional Practice Inventory (IPI) Workshop.

Your IPI Reliability rating was: 1.00

* As indicated during the IPI Workshop, a reliability score of .90 or higher is necessary for permission to use the Instructional Practices Inventory Process for research purposes.
* A reliability score of .80 or higher is necessary for permission to use the IP Process for internal use within a school or district to collect data for faculty study for school improvement.
* A reliability score .70 - .79 indicates that you may use the IPI for personal or informal purposes only—not for research or to use in school improvement. If your score falls in this range and you desire to use the IPI for research or school improvement purposes, you should work further to hone the observational and coding skills relevant to this process and undertake a reliability assessment at a future date. One of the best methods to develop a higher reliability rating is to shadow a colleague during a data collection who has a reliability rating of .80 or higher.
* A reliability score below .70 indicates that you should not use the IPI for data collection. You should participate in another IPI workshop and score .80 or higher before using the IPI. In the meantime, you are encouraged to shadow a colleague during a data collection who has a reliability rating of .80 or higher to practice your coding skills.

All participants in the IPI Workshop are on the honor system regarding their reliability scores. Your scores are sent directly to you and are not provided to your colleagues or supervisor. Therefore, if you did not score .80 or higher, then you should not serve your school as a data collector for the IPI process. As noted above, you may wish to attend another workshop to improve your coding skills so you can assist with the data collection process. And, if you wish to collect IPI data for the purposes of research and did not score .90 or higher, you should attend another workshop and take another assessment to improve your reliability score.

Finally, please remember that the IPI process was not designed for, and should never be used for, personnel evaluation. All IPI data should be aggregated and not be used in any manner that reflects directly upon an individual or specific classroom. The IPI was designed to provide data for faculty to study how students across the school are engaged and should be used for those school improvement purposes only. And when you serve as a data collector for the process, you should honor the privacy of all teachers and classrooms and keep confidential what you observe.

Thank you for your participation in the IPI workshop, and we certainly hope you will find the IPI Process of use for both yourself and your school.

For more information about the IPI Process, visit the IPI website at www.ipistudentengagement.com.

Sincerely,

Jerry Valentine, Ph.D.
Professor, Emeritus
University of Missouri
ValentineJ@missouri.edu
www.ipistudentengagement.com

APPENDIX D

Email Recruitment for Participants

Dear X,

I am conducting a study to examine the role of higher order thinking skills instruction in secondary and middle-level mathematics classrooms in the light of Maine's move to proficiency-based education.

As a secondary or middle-level mathematics educator in Maine, you are invited to take part in my research. Participation will include short classroom observations on two days, along with two follow up focus group discussions. A full description of this study is attached in the "Informed Consent" form.

The focus groups will be about an hour long and will be recorded for accuracy purposes. It can be held at a location of convenience for you and the other participants. Please feel free to respond by email and we can discuss a time that will work for the focus group. I would be happy to answer any questions that you have. I hope you agree to participate and I look forward to discussing it with you soon.

Sincerely,

Quinton Donahue
Ph.D. Student in Educational Leadership
University of Maine

APPENDIX E

Informed Consent for Participants

EXAMINING OPPORTUNITIES FOR HIGHER-ORDER THINKING IN SECONDARY AND MIDDLE-LEVEL PROFICIENCY-BASED MATHEMATICS CLASSROOMS

You are invited to participate in a research project being conducted by Quinton Donahue, Ph.D. student in the department of Educational Leadership at the University of Maine. Dr. Ian Mette, Assistant Professor of Educational Leadership at the University of Maine, is the faculty advisor for this study. This research will examine the frequency of engaged higher-order classroom tasks in proficiency-based mathematics classrooms. In addition, this research will aim to better understand the instructional practices associated with higher-order thinking skills integration in secondary and middle-level proficiency-based mathematics classrooms.

What Will You Be Asked to Do?

If you decide to participate, you will open your mathematics classroom to two days of short walkthrough style observations using the Instructional Practices Inventory (IPI), an observational tool that focuses on the level of thinking associated with engaged instructional practices. These observations will take place in the months of October and December.

In addition, participants will be asked to partake in two focus groups with fellow mathematics teachers in your school. IPI observational results will be shared with the group and conversation will be focused around what structures or practices helped to produce the observed results.

Risks

Your identity will be protected by keeping observational data and focus group conversations confidential and providing you with a pseudonym upon transcription. Another risk will be sharing information in front of your peers. Time and inconvenience could also be considered as risks.

Benefits

The first benefit that could be received from your participation is the opportunity for professional discussion and reflection around the topic of higher order thinking skills in the classroom. An additional benefit will be that information provided will be used to better understand higher-order instructional practices in the middle school mathematics classroom.

Confidentiality

All attempts will be made to protect your identity with a pseudonym. The key linking your name to the pseudonym will be kept in a file on a single password protected computer. Upon transcription of focus group conversations, this information be deleted. All transcribed conversations and observational data will be kept on a password protected computer for an indefinite period of time.

Voluntary
Your participation in this study is voluntary. If you choose to take part in this study, you may stop at any time. You may skip any focus group questions that you do not wish to answer.

Contact Information

If you have any questions about this study, please contact the researcher at the information listed below:

Quinton S. Donahue
Ph.D. Student in Educational Leadership
University of Maine
(207) 356-8449
quinton.donahue@maine.edu

If you have any questions about your rights as a research participant, please contact Gayle Jones, Assistant to the University of Maine's Protection of Human Subjects Review Board, at 581-1498 (or email gayle.jones@umait.maine.edu).

Your signature below indicates that you have read the above information and agree to participate. You will receive a copy of this form.

____________________________________ __________________
Signature Date

APPENDIX F

IPI Data Collection Sheet

INSTRUCTIONAL PRACTICES INVENTORY DATA RECORDING FORM 8-11-11

PERIOD OR TIME ____________________ PAGE NUMBER________

Sub-Groups If Requested											
Course/ Content	Grade Level	Other	Student Disengage-ment	Student Work with Teacher Not Engaged	Student Work with Teacher Engaged	Teacher-Led Instruction	Student Verbal Learning Conversa-tions.	Student Active Engaged Learning	Anecdotal Notes	Core	Non Core
			①	②	③	④	⑤	⑥		○	○
			①	②	③	④	⑤	⑥		○	○
			①	②	③	④	⑤	⑥		○	○
			①	②	③	④	⑤	⑥		○	○
			①	②	③	④	⑤	⑥		○	○
			①	②	③	④	⑤	⑥		○	○
			①	②	③	④	⑤	⑥		○	○
			①	②	③	④	⑤	⑥		○	○
			①	②	③	④	⑤	⑥		○	○
			①	②	③	④	⑤	⑥		○	○
			①	②	③	④	⑤	⑥		○	○
			①	②	③	④	⑤	⑥		○	○

SCHOOL___________________________ DATE OF OBSERVATION________________ OBSERVER_______________

Jerry Valentine 8-11-11

The IPI was not designed for personnel evaluation and should not be used for that purpose. Sub-group profiling should be by faculty request only.

BIOGRAPHY OF AUTHOR

Quinton Donahue grew up in Fairfield, Maine and graduated from Lawrence High School in 1997. He attended the University of Maine in Orono and received his Bachelor of Science degree in Secondary Mathematics Education in 2004. He also received a Master of Education degree in Educational Leadership from the University of Maine in 2010. He worked as a high school math teacher at Hampden Academy in Hampden, Maine for 10 years. During his time as a teacher, he served as Mathematics Department Chair for four years. Following this, Quinton was the middle school principal of Mount View Middle School in Thorndike, Maine for five years. He is currently the principal of Gorham Middle School in Gorham, Maine, where he has worked for three years. Quinton is a candidate for the Doctor of Philosophy degree in Educational Leadership from the University of Maine in December 2021.