IMPROVING STUDENTS' SCORES ON THE MAINE EDUCATIONAL ASSESSMENT FOR SCIENCE

by

Suzanna Brawn

Liberty University

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University

2020

IMPROVING STUDENTS' SCORES ON THE MAINE EDUCATIONAL ASSESSMENT

FOR SCIENCE

by Suzanna Brawn

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University, Lynchburg, VA

2019

APPROVED BY:

Russell Claxton, Ed. D., Committee Chair

Grania Gothard Holman, Ed. D., Committee Member

Bunnie Claxton, Ed. D., Committee Member

ABSTRACT

The purpose of this applied study was to solve the problem of low test scores on the Maine Educational Assessment for Science at Hartman High School, a suburban public high school in southeastern Maine, using a multimethod approach and to design a solution to the problem. The central research question was, “How can the problem of low test scores on the Maine Educational Assessment for Science be solved at Hartman High School?” Data were collected using both qualitative and quantitative approaches, including interviews with teachers and administrators familiar with students’ performance on the assessment at Hartman High School, a survey of all science teachers and administrators at Hartman High School, and review of documents, including archival data regarding Hartman High School students’ performance on the assessment from the Maine Department of Education. Data were analyzed for codes and themes, from which the solution to solve the problem of low test scores on the Maine Educational Assessment for Science at Hartman High School were derived.

Keywords: standardized test, Maine Educational Assessment, high school science, instructional strategies, test scores, professional development

Dedication

This dissertation is dedicated to my two favorite human beings—my husband and my son.

Matt—your willingness to put up with my crazy ideas (including pursuing this degree) is beyond measure, and I am forever grateful for your patience and understanding as I make this dream come true! I love you more than words can ever say.

Braedyn—thank you for always offering to help me with my homework and to sit with me as I work. I hope that I have modeled for you what dedication looks like and that this degree serves as a reminder of what hard work and perseverance can become, and I hope that you have an eternal love of learning. You are my entire world—I love you.

Acknowledgments

First, I thank my parents for their unending love and support and for encouraging a love of learning and teaching in me from an early age. Without the countless hours of dedication that have been put into helping me be the best I can be, I would not be where I am today. I love you.

I also want to thank my dissertation chair, Dr. Russ Claxton. He is not only a wonderful chair, but he has been a fantastic source of wisdom and knowledge when I have faced both personal and professional challenges. He has dealt the hard blows, picked me up, dusted me off, and sent me on my way several times, and I will forever be grateful for his guidance throughout the years. He is a profound reminder of God's work in our lives.

Dr. Bunnie Claxton has been my biggest cheerleader at Liberty University throughout this journey, and God has surely blessed me with her wisdom, insight, and feedback. She has challenged me to better myself and my work, pushing me forward and helping me to grow as a scholar and individual. Not only has she been an incredible mentor, but she is an invaluable friend, and I am forever indebted to her for the graciousness she has given me.

Above all, I thank God for His grace and mercy, for His gift of talent and perseverance, and for His unending love. Isaiah 40:31 (New International Version) says "but those who hope in the Lord will renew their strength. They will soar on wings like eagles; they will run and not grow weary, they will walk and not be faint." Through His will alone have I endured the dissertation process, "running" without becoming weary, renewing my strength time and time again, and now, soaring like eagles. I pray that my work and my life will be a testament to His will and His love for each of us.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

American College Test (ACT)

Adequate Yearly Progress (AYP)

End of Course (EOC)

Every Student Succeeds Act (ESSA)

Elementary and Secondary Education Act (ESEA)

General Education Diploma (GED)

Maine Assessment of Educational Progress (MAEP)

Maine Department of Education (MDOE)

Maine School Administrative District (MSAD)

Massachusetts Department of Education (MDOE)

Maine Educational Assessment (MEA)

Next Generation Science Standards (NGSS)

New Teacher Center (NTC)

No Child Left Behind Act of 2001 (NCLB)

Professional Learning Community (PLC)

Scholastic Aptitude Test (SAT)

Standards of Learning (SOL)

State of Texas Assessments of Academic Readiness (STAAR)

Teaching and Learning Cycle (TLC)

United States Department of Education (USDOE)

Virginia Department of Education (VDOE)

CHAPTER ONE: INTRODUCTION

Overview

This applied research study sought to solve the problem of low test scores on the Maine Educational Assessment (MEA) for Science at a mid-size suburban public high school in southeastern Maine. Historically, approximately half of the students at Hartman High School have failed to meet the state's level of satisfactory performance on the assessment, (Maine Department of Education, n.d), which is defined as work that, "demonstrates an adequate understanding of essential concepts in science, including the ability to make connections among central ideas," (MDOE, 2018b, para. 2), and an "ability to analyze and solve routine problems and explain central concepts with sufficient clarity and accuracy to demonstrate general understanding" (MDOE, 2018b, para. 2).

To identify how the low test scores on the MEA for Science at Hartman High School can be improved, a central question and three sub-questions were developed. The central question was, "how can the problem of low test scores on the Maine Educational Assessment for Science be solved at Hartman High School?" Following the central question, three sub-questions were developed as follows: (1) How would teachers and administrators in an interview solve the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine? (2) How would science educators participating in a survey solve the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine? And (3) How does a review of documents inform the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine?

While there are a variety of factors that have a negative impact on students' scores that the school cannot change, such as students' socioeconomic status, race, gender, and community resources (Greenwald, Hedges, & Laine, 1996), there are several factors that the school may change that may have a positive impact on students' performance. These factors include teacher efficacy, class size, curriculum, and instructional strategies used. Identifying the factors that have the most impact on low test scores on the MEA for Science at Hartman High School is particularly important to the stakeholders, as the implications of the data can have significant effects on the community. First and foremost, improving students' scores on the MEA for Science can provide increased academic and scholarship opportunities for students, as well as improved student and teacher efficacy. Likewise, increased scores can lead to higher overall ratings of the school, making the community more desirable for families, businesses, and companies exploring relocation options.

Chapter One provides the background of this study, including the historical, theoretical, and social contexts. The problem to be addressed and the purpose of the study are discussed. The chapter continues to introduce the research questions and the significance of the study. The chapter closes with a list of key terms used in the study with definitions relevant to the context of the research.

Background

Standardized testing has become a staple of the American public education system over the past 150 years (National Education Association, 2019a). Because of accreditation and accountability measures put in place by various government agencies, today's students are tested at each level of their K-12 educational journey (Fletcher, 2009), which provides data necessary to satisfy requirements set forth by both state and federal governments. Failure to have an overall satisfactory performance as a school can result in a variety of sanctions (Ladd, 2017), including a loss of funding (Maine School Administrative District 75, 2019) and reconstitution if progress is not made (Lynch, 2016b). Therefore, schools strive to improve their students' performance on standardized assessments to ensure that they continue to receive the benefits associated with high scores on standardized tests such as funding for a variety of programs, including Title I monies to improve curriculum and programming (MSAD 75, 2019).

In Maine, accreditation and accountability are measured, in part, by the MEA. The MEA is a series of tests that "measure the progress of students who live in Maine in the areas of English, Language Arts, and Literacy, Mathematics, and Science" (MDOE, 2018c, para. 2). The MEA for Science is administered to students during the spring of their third year of high school. The assessment is a comprehensive exam that includes topics from both life and physical sciences. Historically, students at Hartman High School have demonstrated a subpar performance with as many as 53.42% of students falling below the state's expectations of students in the last three years (MDOE, n.d.).

Historical Context

Standardized tests have been a staple of the American education system since the nineteenth century. At the time, Horace Mann suggested written assessments as an alternative to the traditional oral exams for students (Gallagher, 2003). Since then, standardized tests have grown in their use and application, especially in the United States. As the population grew in the second part of the nineteenth century, universal education became a need (United States Congress, Office of Technology Assessment, 1992). Standardized assessments have been intended to provide educators with data needed to offer students an appropriate education based on their abilities rather than by students' age (United States Congress, Office of Technology Assessment, 1992). However, in most public schools, students remain segregated by age rather than ability, though some improvements have been made in this area, such as the implementation of gifted and special education programs.

In 1965, the Elementary and Secondary Education Act (ESEA) was enacted by President Nixon, intended to bring consistency to educational programs nationwide, especially for students considered to be socioeconomically disadvantaged (United States, 1965). To monitor this, standardized testing was heavily relied upon to ensure that opportunities were equitable and that the return-on-investment of federal funds was, in fact, a wise investment. Congress renewed ESEA every five years, making only minor revisions along the way (Tillman, 2006).

A Nation at Risk was published in 1983, in which it was noted that students' performance on standardized tests were falling short of expectations, schools were not adequately challenging students, and the country as a whole was falling behind other countries in terms of producing a "literate and educated society" (Tillman, 2006, p. 1). State and local governments were identified as being incapable of providing an appropriate education for students; as a result, the federal government would have greater oversight of public education programs (Tillman, 2016). Standardized tests became the standard through which all schools would be monitored (Tillman, 2016).

While standardized tests were originally used to determine students' performance, the tests failed to provide much more information, especially with regard to various demographic subgroups' performance on the assessment (Education Post, 2019). As a result, two reauthorizations of ESEA, the *No Child Left Behind Act* (NCLB) and the *Every Student Succeeds Act* (ESSA), were enacted in the first part of the twenty-first century.

Identification of students' affiliation with a particular subgroup is important when considering the data acquired from standardized tests. These data are often used to determine a school's accreditation status, which may impact the amount of funding a school receives from any one of a number of sources. For instance, if a school in Maine is not in compliance with testing requirements, funding can be reduced or eliminated (MSAD 75, 2019). Title I funding, a result of the 1965 ESEA, can be withheld until a school is, or has made sufficient efforts to be, in compliance with testing requirements (MSAD 75, 2019). NCLB put additional requirements on schools to demonstrate adequate yearly progress (AYP); failure to demonstrate such progress often results in sanctions for the school (Lynch, 2016b).

Social Context

Several case studies were reviewed and used to identify causes of students' low performance on standardized assessments. Analysis of these studies suggest that a variety of factors exist, a considerable amount of which are not factors that schools and their administrators and teachers can directly change. These factors include socioeconomic status, race, gender, and resources provided by the community (Greenwald, Hedges, & Laine, 1996). For instance, Caldas and Bankston (1997) found that students from socioeconomically disadvantaged homes often lack intellectually stimulating materials that help them to advance their education, even prior to beginning school. While these factors are significant, and their impact cannot and should not be ignored, there are other factors that can help to improve students' standardized test scores that can be influenced by decisions made at the school and classroom levels.

Though some researchers say that the factors that cannot be controlled by the school have little, if any influence on students' performance on standardized tests (Berliner, 2001; Betts, Reuben, & Danenberg, 2000; Tizard, Blatchford, Burke, Farquhar, & Plewis, 2017), much remains that schools *can* do to increase students' scores on these assessments (Grissom, Kalogrides, & Loeb, 2015; Maxwell, 2016). Factors within the control of the school, administrators, and teachers, such as teacher preparation (Darling-Hammond, 2015), teacher efficacy (Ronfeldt, Loeb, & Wyckoff, 2013), class size (Bosworth, 2014), curriculum specifications (Tarr et al., 2008), classroom environment (Tarr et al., 2008), and instructional strategies (Day, Gu, & Sammons, 2016), can have a direct impact on students' success on standardized assessments. In fact, a correlation has been determined by Ronfeldt et al. (2013) between a teacher's educational background and professional development acquired and students' performance. Ronfeld et al. (2013) also identified the correlation between teacher collaboration and positive effects on test scores. Because teacher performance improves when teachers have the opportunity to effectively collaborate, students' performance on standardized assessments in math and reading tends to improve, as well (Ronfeldt et al., 2013).

Theoretical Context

The first theory applied in this research study is Piaget's Theory of Cognitive Development (Piaget, 1951, 1952, 1964). This theory proposes that children develop their intelligence over time and that they think differently than adults (Piaget, 1951, 1952, 1964). Children's intelligence is developed through a four-stage process, which includes the sensorimotor stage, the preoperational stage, the concrete operational stage, and the formal operational stage (Piaget, 1951, 1952, 1964). Each of these occurs over a set window of time, though the time spans are not hard-and-fast rules, but rather guidelines for expected developmental changes. This theoretical framework is well-aligned with this study, as it helps to provide a foundational understanding of how students learn, which is important when considering an appropriate standardize assessment.

The second theory considered in this study is the Measurement of Intelligence Theory by Binet and Simon (1916). This theory is the one upon which modern standardized tests are based. The theory proposes that intelligence can only be measured accurately when compared to others of a similar background (Binet & Simon, 1916). This theory has served as the backbone for many standardized assessments, including the Stanford-Binet Intelligence Scale (United States Congress, Office of Technology Assessment, 1992), which remains a widely-used assessment in present practice.

Problem Statement

The problem is that students at Hartman High School are not meeting the Maine Department of Education's (MDOE) expectations for performance on the MEA for Science. In fact, over the last three academic years, approximately 50% of students at Hartman High School have failed to meet the MDOE's expectations for performance on the MEA for Science (MDOE, n.d.). Specifically, low test scores for this study are defined as when a students' performance "demonstrates an incomplete understanding of essential concepts in science and inconsistent connections among central ideas" (MDOE, 2018b). Further, "the student's responses demonstrate some ability to analyze and solve problems, but the quality of responses is inconsistent" (MDOE, 2018b). Students scoring below a 50% on the assessment, earning 39 out of the 80 possible points or less, are considered to have failed to meet the state's expectations (MDOE, 2018b). Low scores on the assessment can have detrimental effects for the stakeholders at Hartman High School, who include students, parents, teachers, staff, and administrators, as well as the school board, town governments, and businesses and companies within the region.

Current research provides a multitude of information regarding factors that have a negative impact on students' performance on standardized tests. Specifically, socioeconomic status (Caro, McDonald, & Willms, 2009), race (Quinn & Cooc, 2015), and gender (Quinn & Cooc, 2015) tend to produce considerable achievement gaps between demographic groups. Further, class size, curriculum, and teacher effectiveness, if not managed well by a school and its administration, can have a negative impact on students' performance. Regardless of the existence of these factors that vary from school to school, the MDOE has the same expectations for all students in the state (MDOE, 2018b). One expectation is that students will perform at or above a level described by the MDOE as demonstrating "an adequate understanding of essential concepts in science, including the ability to make connections among central ideas" (MDOE, 2018b, p. 1). Further, students must be able to construct a written response that indicates an ability to analyze and solve problems and explain central ideas in a clear and accurate manner (MDOE, 2018b). These expectations, of course, parallel and fulfill the requirements set forth by the United States Department of Education (USDOE) (United States, 2015). Research also shows that a variety of interventions and strategies can be effective in helping students to perform better on standardized tests such as the MEA for Science (Akiba & Liang, 2016; Gehlbach, Brinkworth, Hsu, King, McIntyre, & Rogers, 2016). What this research is lacking, however, is application to science-specific standardized assessments.

Purpose Statement

The purpose of this applied study is to solve the problem of low test scores on the MEA for Science at Hartman High School and to formulate a solution to address the problem. A multimethod design was used consisting of both qualitative and quantitative approaches. The first approach used structured interviews with a total of five participants from Hartman High School. Four science teachers and one administrator were interviewed. Each of these participants were familiar with the MEA for Science and students' historical performance on the assessment at Hartman High School. The second approach employed a survey of seven teachers in the science department, two special education collaborative teachers for the science department, the school's testing coordinator and counselor, a retired department chair with 30 years of experience at the school, and four administrators at Hartman High School using questions grounded in literature. This survey was administered using Google Forms, a web-based platform hosted by Google. The third approach utilized the review of documents and archival data, including specific data about students' performance from Hartman High School the MEA for Science from the MDOE.

Significance of the Study

The significance of this study lies not only in improving students' test scores on the MEA for Science, but also in the general advancements that can be made in the instructional practices at Hartman High School. The benefits of improving students' performance on the MEA for Science include extrinsic aspects, such as scholarship and college placement opportunities for students (Ellis, 2018), as well as more intrinsic benefits, including an increased sense of pride and ability. The instructional benefits resulting from this study include identifying practices that can improve student learning in ways not measured by this standardized assessment; improving instructional practices can have great benefit for students who may demonstrate below-standard performance, as well as students who may meet or exceed expectations set forth by the MDOE on the MEA for Science. For teachers and administrators, increased scores on the MEA for Science may lead to increased teacher efficacy, which studies have shown lead to even greater instructional practices (Ware, 2002). In areas where test scores are a component in the teacher evaluation system, increased test scores may be seen as an indicator of teacher effectiveness. Administrators benefit from increased test scores, as such scores allow their focus to shift to other initiatives within the school and may lead to increased funding and community support.

Improving scores on the MEA for Science can also have a positive impact for the communities served by Hartman High School. When students' achievement is high, the school is likely to boast a higher rating among other schools. This can translate to a more desirable community in which to live, increasing property values and attracting businesses and companies to move into the area (Lynch, 2015). This, then, leads to increased local revenue and employment opportunities, making the community an ideal location for graduates to live in upon exiting Hartman High School or college learning experiences. Likewise, the school may become a desirable place to work, leading to less turnover and an ability to attract more highly-qualified teachers.

Research Questions

Central Question: How can the problem of low test scores on the Maine Educational Assessment for Science be solved at Hartman High School?

Sub-question 1: How would teachers and administrators in an interview solve the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine?

Sub-question 2: How would science educators participating in a survey solve the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine?

Sub-question 3: How does a review of documents inform the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine?

Definitions

1. *Accountability* – "the process of evaluating school performance on the basis of student performance measures" (Loeb & Figlio, 2011, para. 1).

2. *Accreditation* – "a process by which recognized authorities validate that an institution meets minimal professional standards and accountability based on its mission" (Greenberg, 2014, p. 2).
3. *Assessment* – "can refer to the process faculty use to grade student course assignments, to standardized testing imposed on institutions as part of increased pressure for external accountability, or to any activity designed to collect information on the success of a program, course, or University curriculum" (Stassen, Doherty, & Poe, 2001, p. 5)
4. *Bubble Student* -- "students who might otherwise perform just below the proficiency threshold" (Springer, 2008, para. 1).
5. *High-Stakes Testing* – "tests used to make important decisions about students" (The Education Alliance, 2019, para. 1).
6. *Intervention* – "provide students with support needed to acquire the skills being taught by the educational system and should address functional skills, academic, cognitive, behavioral, and social skills that directly affect the child's ability to access an education (Lestrund, 2013, para. 1).
7. *Standardized Test* – provides information as to why a child may be struggling or succeeding with specific elements of their grade-level standards; results are used to inform the next step of learning (O'Malley, 2012)
8. *Teacher Effectiveness* – the act of consistently producing higher academic gains among students (Johnson & Semmelroth, 2014)

Summary

Because students' scores on the MEA for Science at Hartman High School have historically shown that approximately 50% of students do not meet the state's expectations, this study, which aimed to identify the causes for their poor performance and strategies to increase student performance, was necessary. Past research has identified causes of poor performance, some of which are not able to be changed, such as race, gender, and socioeconomic status (Greenwald, Hedges, & Laine, 1996), as well as other factors, including teacher preparation (Darling-Hammond, 2015), teacher efficacy (Ronfeldt, Loeb, & Wyckoff, 2013), class size (Bosworth, 2014), curriculum specifications (Tarr et al., 2008), classroom environment (Tarr et al., 2008), and instructional strategies (Day, Gu, & Sammons, 2016), that can be changed by the school and its teachers. However, research has fallen short of identifying the aspects specific to Hartman High School and science assessments as an individual consideration.

Chapter One provided the background of the study, including the historical, theoretical, and social context. The problem to be addressed was introduced, as well as the purpose of the study and its significance. The chapter then continued with the research questions of the study, as well as list of key terms in the study with definitions relevant to the context of the study. This information provides the support necessary for the study and the intentions of the proposed solution to the problem.

CHAPTER TWO: LITERATURE REVIEW

Overview

This chapter begins by introducing the theoretical framework upon which the research is based. Two theories are used to guide this study—Piaget's Theory of Cognitive Development, which focuses on students' learning by connecting previous experiences to make new meaning (Ultanir, 2012), and the Measurement of Intelligence Theory by Binet and Simon (1916), the foundation upon which modern standardized tests are built. The combination of these theories allows students to build on prior experiences, while acquiring new knowledge and applying skills and concepts to various situations (Zhang, 2016), while also considering the purpose, structure, and relevance of standardized testing. Extensive discussions of how these theories and their origins are included, along with the ways in which these theories apply to this study.

A review of literature related to this research and associated research questions is provided. The historical significance of standardized testing is discussed, along with its evolution into modern educational practices. Specifically, the models used for standardized testing in science, the purposes of these assessments, the implications of data, and their alignment to state and federal standards are described. Further, the MEA for Science is discussed in detail, particularly with regard to how this test was designed and developed, as well as its manner of fulfilling federal accountability requirements. Finally, factors that are associated with students' performance on standardized tests are discussed.

Theoretical Framework

Because a theory describes behavior through "a set of related concepts, assumptions, and generalizations" (Joyner, Rouse, & Glatthorn, 2013, p. 57), it is necessary to identify the theoretical foundation upon which this multimethod applied research study was based. Doing so allows the related literature, research questions, and data to be viewed through lenses that solve a particular problem while remaining grounded in sound educational theories. The theories upon which this study were based are Piaget's Theory of Cognitive Development (1951, 1952, 1964) and the Measurement of Intelligence Theory by Binet and Simon (1916).

Piaget's Theory of Cognitive Development

Piaget's Theory of Cognitive Development (Piaget, 1951, 1952, 1964) proposes that children are not little adults, but rather that intelligence is developed over time. That is, Piaget proposes that children are not less intelligent than adults; children simply think differently (Piaget, 1951, 1952, 1964). According to Piaget's theory, children's intelligence is developed through a four-stage process. These stages include the sensorimotor stage, the preoperational stage, the concrete operational stage, and the formal operational stage.

The first stage, sensorimotor, occurs from birth to age 2 (Piaget, 1952). During this stage, children "acquire knowledge through sensory experiences and manipulating objects" (Cherry, 2019b, para. 10). As children grow and develop during this stage, interacting with their environment, they are becoming aware of how the world around them works, developing a sense of object permanence, and acquiring and building vocabulary and communication skills (Brown & Desforges, 2006). During the second stage, the preoperational stage, language continues to develop, and pretend play and associated skills begin to emerge. This stage occurs between ages 2 and 7 (Piaget, 1951). From ages 7 to 11 (Piaget, 1964), children are in the concrete operational stage, which is the third stages in Piaget's theory. During this stage, children's thinking remains concrete and literal, but more logical (Cherry, 2019b). Egocentrism diminishes and empathy begins to develop, but hypothetical and abstract concepts remain challenging for children (Cherry, 2019b). The fourth and final stage, the formal operational stage, occurs at age 12

(Cherry, 2019b) and beyond and is characterized by an "increase in logic, the ability to use deductive reasoning, and an understanding of abstract ideas" (Cherry, 2019b, para. 22). Children are able to start thinking more scientifically about their surroundings and situations in the world around them (Cherry, 2019b). Children develop the skill set necessary to make plans and propose solutions to hypothetical situations during this stage (Cherry, 2019b).

Piaget's Theory of Cognitive Development aligns with this study, as it explains the process of how children acquire knowledge from birth through adolescence and beyond. This information is pertinent to the study because understanding how students gain knowledge and build upon their experiences is necessary to understanding the ways in which they learn in the present moment, as well as how they may best demonstrate this learning. Piaget and Inhelder (1958) asserted that problem-solving strategies and skills are not taught, but rather discovered throughout a child's life (as cited in McLeod, 2018). The MEA for Science as a measure of students' knowledge and understandings is one way they demonstrate their learning and is aligned with Piaget's Theory of Cognitive Development; the scores on the assessment provide evidence of student learning and students' ability to use a variety of problem-solving strategies.

Measurement of Intelligence

The Measurement of Intelligence Theory by Binet and Simon (1916) provides a framework and foundation for modern standardized tests. The theory posits that intelligence cannot be measured through a singular tool; rather, the broad nature of one's intelligence can only be fairly measured when compared to others of similar backgrounds (Binet & Simon, 1916). In France, Binet was tasked with identifying students that needed additional assistance in their educational journey as compulsory attendance became mandatory (Cherry, 2019a). This led to the creation of the Binet-Simon scale, which later became known as the Stanford-Binet

Intelligence Scale when it was standardized by Stanford University psychologist, Lewis Terman (United States Congress, Office of Technology Assessment, 1992). Because many modern standardized tests are based on Binet and Simon's Measure of Intelligence Theory (1916) and, likewise, the Stanford-Binet Intelligence Scale (United States Congress, Office of Technology Assessment, 1992), this theory is an appropriate choice for this study which seeks to improve standardized assessment scores.

It is currently known, as found in research-based literature, that several factors impact students' performance on standardized tests (Bertolini, Stremmel, & Thorngren, 2012). Likewise, there are also several strategies that can be implemented at various levels within the structure of a school to help improve students' performance on standardized tests (Education World, 2019; National Association of Elementary School Principals, n.d.; Garcia & Thornton, 2014; Marsh, Pane, & Hamilton, 2006; Herman, Osmundson, & Dietel, 2010, WiseWire, 2016; Case, 2005). This study seeks to identify the themes related to why students' performance on the MEA for Science at Hartman High School are low and how teachers and administrators can help to improve these scores.

Related Literature

A literature review serves the purpose of providing details and focus regarding the research topic (Yin, 2014), marrying the existing research to the proposed study. Understanding the parallels between the MEA for Science and other state-based assessments used for federal accountability purposes, including the structure and scoring of these assessments, is critical to understanding the MEA for Science and its implications. Thorough research has been conducted regarding ways to improve student achievement on high-stakes assessments, but no research has been conducted specifically related to the MEA for Science. This literature review was conducted to provide an understanding of the history of high-stakes testing, the structure of standardized tests, the use of standardized testing in science, the MEA for Science, and strategies and interventions that can improve students' performance on standardized tests.

Historical Overview

Standardized testing has a long-standing history in American public education. First suggested by Horace Mann in the first half of the nineteenth century as an alternative to oral exams for students, written examinations were intended to identify the "quality of teaching and learning in urban schools, monitor the quality of instruction, and compare schools and teachers within each school" (Gallagher, 2003, pp. 83-84). Since their introduction, standardized tests have grown in popularity, especially as the population of the United States has grown rapidly through the latter half of the nineteenth century and along with it, the need to provide universal education (United States Congress, Office of Technology Assessment, 1992). These assessments provided educators with information necessary to place students according to ability, rather than the previous practice of grouping students by age (United States Congress, Office of Technology Assessment, 1992).

Fast-forward to 1965 when President Nixon enacted the Elementary and Secondary Education Act (ESEA), and evidence of the continued need for standardized testing becomes abundant. ESEA sought to bring consistency to the educational opportunities offered to socioeconomically disadvantaged students when compared to those from more affluent upbringings (United States, 1965). Standardized testing became a way to measure the quality of schools and their programs, as well as a way to monitor the return-on-investment of government funds into schools with underserved populations. ESEA was renewed every five years in congress, often with minor revisions made at each renewal (Tillman, 2006).

In 1983, *A Nation at Risk* was published, which outlined concerns about the progress and achievements of American society, with education as the foundation upon which all progress and achievement is built (United States National Commission on Excellence, 1983). Specifically, "the commission noted that student test scores were falling, schools were requiring less rigor (fewer required courses in math, science, and advanced placement classes), and the United States fared poorly with other countries in producing a literate and educated society" (Tillman, 2006, p. 1). The report implied that local and state education entities were not sufficient in monitoring students' and schools' progress; instead, the federal government would have an increased role in education (Tillman, 2016). The report also called for a greater emphasis on science and math education and standardized testing in K-12 education (Tillman, 2016).

Since *A Nation at Risk* was published, several iterations of the original ESEA have been implemented (Klein, 2015). While the ESEA was renewed every five years since its inception in 1965, significant changes warranted a new name for the initiative, including the *No Child Left Behind* during George W. Bush's presidency (Brenchley, 2015) and Obama's *Every Student Succeeds Act* (Brenchley, 2015). Part of the reason that changes to the original guidelines of ESEA were necessary lies in the standardized tests themselves. The tests were intended to monitor students' and schools' performance, but the original guidelines failed to identify specific subgroups within the tests' results (Education Post, 2019). Thus, testing parameters now require the reporting of demographic information for each test-taker, allowing stakeholders at all levels to access information related to the performance of a specific subgroup, which has remained a focus of each of the iterations of ESEA since 1965.

Accreditation and Accountability

Standardized testing is often tied to the accreditation and accountability policies of school districts, as well as the state and federal governments. For instance, in Maine, students' performance on MEA assessments, administered at various times throughout a students' K-12 experience, contribute to a schools' report card by measuring the overall performance of students in English, math, and science, as well as by comparing achievement gaps among subgroups in English and math (MSAD 75, 2019).

A school's performance on these assessments can determine its accreditation status, which can ultimately impact the availability of funding available from various entities. For example, in Maine, if a school fails to comply with the testing requirements for student participation, funding can be reduced or eliminated (MSAD 75, 2019). Title I funds, those which are used to support programs that specifically enhance the educational experience of socioeconomically disadvantaged students, can be withheld until reasonable efforts to be compliant with participation standards are made (MSAD 75, 2019). NCLB puts further requirements in place based on standardized testing scores, requiring students to make adequate yearly progress (AYP), and schools that failed to do so are subject to sanctions (Lynch, 2016b). These sanctions include developing an improvement plan as a minimum with reconstitution of the school if a school fails to make AYP for five years (Lynch, 2016b).

Characteristics of Standardized Tests

Because of accountability requirements in the United States per ESSA (United States, 2015), standardized tests are prominent in every school nationwide. While classroom assessments often include several similarities to standardized tests, a test can only be characterized as standardized if it meets a set of specific criteria (Christensen, 2018). These criteria include being consistent, norm-referenced, criterion-referenced, reliable, and valid (Christensen, 2018). Further, standardized tests follow standard procedures for administration and scoring and include the same content in all versions (Cerezo, n.d.).

The consistency of a test can be determined in a number of ways, including test-retest reliability, parallel forms reliability, inter-rater reliability, and internal consistency reliability (Phelan & Wren, 2006). Test-retest reliability is "a measure of reliability obtained by administering the same test twice over a period of time to a group of individuals" (Phelan & Wren, 2006, para. 2). Using this method, the reliability of a test is measured by administering the same test to the same group of individuals at two different times and determining the correlation coefficient of the scores (Lavrakas, 2008). Parallel forms reliability is determined by giving two similar tests that measure the same skills and knowledge, but do so with different questions (Salkind, 2010); comparison of participants' responses provides evidence of the reliability of the test. Inter-relater reliability is best used for tests that require human scoring of answers (Phelan & Wren, 2006). Finally, internal consistency is a useful measure of reliability for tests that probe for the same information with two or more questions on the same test (Frey, 2018). The correlation coefficient between each set of questions is determined, and the reliability is established from this measurement.

Norm-referenced standardized tests compare students' performance on a standardized test to other students of the same age, where criterion-referenced assessments compare students' performance to a set of pre-determined skills and standards (Williams, 2008). Depending on the assessment's purpose, either one of these references may be appropriate. For example, a criterion-referenced test is an appropriate approach for a driver's test, where a growth chart for children as used at a pediatrician's office would be an appropriate norm-referenced assessment (Hawker Brownlow Education, n.d.). In education, school-based assessments, such as those at the end of a unit or grading period are prime examples of criterion-based assessments, where more cumulative, large-scale assessments such as the Scholastic Aptitude Test (SAT) and the American College Test (ACT) are norm-referenced (The Glossary of Education Reform, 2015). Both approaches are widely used in standardized assessments, though tests used for accountability purposes are likely to be criterion-based, as their purpose is to measure students' mastery of course standards.

The validity of a test "refers to what characteristic the test measures and how well the test measures that characteristic," (United States Department of Labor Employment and Training Administration, 1999, para. 19). It is important to note that "validity is not a property of the tool itself, but rather of the interpretation or specific purpose of the assessment tool with particular settings and learners" (Sullivan, 2011, para. 2). That is, a test may be considered valid for a particular group of students in a certain setting, but the validity may not hold true for all students in all settings. Standardized tests, especially when intended to be used in a variety of schools with the majority population, strive to be deemed valid across all settings and all students.

Ensuring that a test is reliable and valid allows interested parties to use the data for a variety of purposes. Most importantly in education, when a test is considered standardized, and therefore reliable and valid, educators can use the results to inform future instructional practices (Mertler, 2007). Additionally, depending on the test and its reference as either norm-based or criterion-based, standardized tests serve as an indicator of students' academic performance and mastery of skills (Jacob & Rothstein, 2016). This information, of course, is valuable to students as they prepare for future studies, though not all tests can serve each of these purposes. Therefore, it is necessary to use a standardized test that will provide the data desired by stakeholders.

Standardized Tests in Science

Due to the accountability requirements of ESSA, state departments of education nationwide have a battery of tests that are required of students at various intervals throughout their K-12 educational experience (United States, 2015). These tests are required in math, reading, and science (United States, 2015), but states are free to choose which assessments will be used in their schools that meet the federal requirements. As a result, different states use different tests to measure students' proficiency. Some states choose to use nationally-normed tests, such as the SAT and ACT for math and reading, while other states choose to create their own criterion-referenced assessments.

Science, despite being required for accountability purposes, is not a required part of school quality reporting (Klein, 2018). Because of this, states have even more freedom in creating their own assessments for science. States such as Virginia with its SOL assessment, New York with its Regents Exam, and Texas with the State of Texas Assessments of Academic Readiness (STAAR), have constructed their own tests based on the standards that students are expected to master for science. These tests, while created by different states and each with their own unique qualities, have several shared characteristics. Understanding the similarities and differences between these standardized science tests, as a sample of those administered nationwide on a state-by-state basis, is necessary to fully understand the details and implications of the MEA for Science; as such, each of these tests is discussed in the following sections.

Virginia Standards of Learning Assessment. In Virginia, students must take at least one standardized science assessment between grades three and five, grades six and nine, and grades 10 and 12 in order to meet federal accountability (USDOE, 2016; Lane, 2018). This test will likely be a biology-based assessment based on enrollment, though students who have previously earned a proficient score on the Earth Science end-of-course (EOC) SOL will not be required to take the Biology SOL unless the proficient Earth Science score was earned prior to enrollment in 10^{th} grade (Lane, 2018). The test consists of 60 multiple-choice, short answer, drag-and-drop, or hot spot questions, 10 of which are considered field questions (Virginia Department of Education, n.d.). Field questions do not count for or against a student's score, but rather, these questions provide information as to the strength of the question for future use (VDOE, n.d.). Test questions are organized into five reporting categories, from which educators can determine students' performance based on themes of each course (VDOE, n.d.). Depending on the course, students must earn a 50% to a 54% in order to receive a proficient score on the exam (VDOE, 2016).

New York Regent's Exam. In New York, the Regents Exam is administered to students in an effort to "align with and assess the knowledge and skills set forth in the NYS learning standards" (University of the State of New York – New York State Education Department, 2018, para. 1). Students are required to be assessed only in particular subject areas. For science, this includes one life science assessment and one physical science assessment during their high school experience (New York State Higher Education Services Corporation, n.d.). These exams, when combined with other assessments in math and reading, satisfy the accountability requirements of ESSA (USNY – NYSED, 2018). Each test varies in format depending on the content, but all tests contain multiple-choice and open-ended questions (New York City

Department of Education, 2019). In order for students to earn a satisfactory score on the Regents Exam for either of the sciences, a score of 65 or higher out of 100 possible points, determined from a conversion scale, must be obtained (USNY – NYSED, 2009, para. 5-6). However, this score is only necessary for students who are working towards earning an Advanced Regents Diploma in addition to the five core subject assessments required for the standard Regents diploma.

State of Texas Assessments of Academic Readiness. The STAAR test in Texas is similar to those offered in Virginia and New York. Students are required to take a standardized science assessment in Grades five and eight, as well as an EOC assessment in biology (Texas Education Agency, 2019). The biology assessment includes 50 multiple-choice questions, of which 40% include some component of scientific process skills (Texas Education Agency, 2018b). According to the most recently reported score conversions, students must answer 33 of the 50 multiple-choice questions correctly in order to meet the state's expectations of performance (Texas Education Agency, 2018a).

While these states serve as examples as to how different states implement the testing requirements of ESSA for accountability purposes, several other assessment models exist based on each state's needs and requirements. Regardless of the differences among the parameters of testing, each of these tests have similarity in their function and design. For example, each test is constructed in a way that measures students' mastery of standards selected by the state. These standards may vary from state to state, but the criterion-referenced nature of the assessment remains consistent throughout. Likewise, each state administers these tests to satisfy the requirements of ESSA, though not all schools use the results of these assessments for other reporting factors, such as school quality profiles or reports.

Maine Educational Assessment for Science

The Maine Educational Assessment for Science is part of a battery of tests intended to satisfy the federal requirements for accountability and to inform stakeholders of the progress and proficiency of local schools (MDOE, 2018c). The evolution from its inception to its current status is one of rocky starts and stops, however, adding several layers to the challenge of understanding the assessment and its implications for students, teachers, and schools.

In 1972, when the Maine Assessment of Educational Progress (MAEP) used a series of nationally-designed tests to measure a cross-section of 4th and 8th-grade students' in the state (Donaldson, 2014). These tests, when compared to national averages, indicated that Maine students were performing at high levels (Donaldson, 2014). This success continued through the 1980s and 1990s, with Grades four and eight students in Maine scoring higher than the national average in math and reading from 1992 until 2000 (Donaldson, 2014).

In the late 1990's, Maine was considered a leader in educational interests, implementing Maine's Common Core of Learning (1990) at the start of the decade and Maine Learning Results in 1997, making the state one of the first to develop state-wide standards for students (Stone, 2018). Students' progress towards these goals was assessed with MEA, in addition to local districts' self-developed assessments, though local assessments were abandoned in 2007 after concerns about the development of these assessments and their implementation arose (Stone, 2018).

In 2009, the MEA was eliminated in favor of the New England Common Assessment Program (NECAP), an assessment used both other states in New England, including New Hampshire, Rhode Island, and Vermont (Stone, 2018). The change in assessment also meant a change in learning standards; this, combined with the federal Race to the Top initiative to secure more funding, resulted in the adoption of the Common Core Standards and Next Generation Science Standards (NGSS) in 2011 (Stone, 2018). The state adopted the Smarter Balanced assessment for the 2014-15 academic year, along with 18 other states, but this test was only used for one year due to rampant objection from students and parents (Gallagher, 2015). It was at this time that the SAT became the assessment for math and reading skills for students and the MEA for Science came back into play (Gallagher, 2015), as there is no SAT component that assesses students' scientific proficiency.

The MEA for Science is administered to students during their third year of high school. Because ESSA dictates that 95% of students in Maine must take accountability assessments during the specified year (MDOE, n.d.b), all students are required to take the test, though no significant penalty is in place for students who do not take the assessment. At Hartman High School, the SAT is given in the spring during the school day rather than on a weekend, which is the traditional SAT administration format (College Board, 2019a), and the MEA for Science is given the day immediately following the administration of the SAT. This is partly in an effort to keep all testing for accountability purposes confined to the same week and under as similar conditions as possible for all students. This, however, increases the likelihood of student burn-out with regard to testing, and inhibits teachers' abilities to review material with students immediately prior to the assessment, which is a common practice with other assessments, especially those such as Virginia's SOL test, New York's Regents exam, and Texas's STAAR test.

The MEA for Science includes content from both life and physical sciences (MDOE, 2018d), assessed through a series of multiple-choice and constructed response items (MDOE, 2018a). Students will answer 48 multiple-choice questions and five constructed response items, though eight of the multiple-choice questions and one of the constructed response items do not count towards or against students' final scores, as these questions are field-test items (MDOE, 2018a). Each multiple-choice question is worth one point each, while constructed response questions are worth four points each. This gives an overall possible score of 56 points on the assessment. Scores, however, are reported between 1100 and 1180 points. Students must earn at least 1142 out of the 1180 points to meet the state's expectations for performance (MDOE, 2018b).

Similar to other states' science tests, the MEA for Science is required for federal accountability purposes (USDOE, 2016). It is not, however, required for school quality profiles, nor is it a consideration for graduation requirements at Hartman High School (Hartman High School, 2019). Instead, the data collected from this assessment are used to make comparison between Maine students and students in other states, helping stakeholders to assess the quality of Maine's schools and curriculum and to identify where further supports are needed (MDOE, 2018c).

Strategies to Improve Standardized Test Scores

Because high-stakes standardized testing is prevalent in nearly all public education systems nationwide, bringing with it a myriad of potential benefits and consequences for schools, teachers, and students, it is no surprise that educators seek to identify the factors that have a negative impact on students' performance on standardized assessments. Several factors that have a negative impact on students' performance are a much larger problem than a school and its teachers can solve, including, but not limited to, socioeconomic status, race, child abuse and neglect, and healthy life of individual students (Bertolini, Stremmel, & Thorngren, 2012). Other factors, however, can have a significantly negative impact on students' achievement, and schools can make changes that help to eliminate these negative influences, sometimes by making changes to current practice and other times by abandoning part of current practice in favor of another strategy that will provide greater gains (Teaching Tolerance, 2019). For instance, analyzing test data, involving parents in their students' education, focusing on bubble students, constant monitoring of success, examining student work, and practice tests and benchmark assessments throughout the academic year can all lead to increased student achievement as measured by a standardized test, such as the MEA for Science (Education World, 2019). However, expecting educators to take on these tasks, especially if they've never experienced such expectations in the past, can be a daunting task. Instead of expecting these changes to be made overnight and to full implementation immediately, it is important to examine ways in which multiple of these strategies can be implemented with a single approach, as well as how to support teachers while they begin the processes. Regardless, each of these approaches, whether combined with another strategy or not, has considerable value in improving students' scores.

Analyzing test data "allows teachers to identify the strengths and weaknesses of an entire class as well as individual students" (National Association of Elementary School Principals, n.d., p. 3). Test data to be analyzed should include all high-stakes testing reports, as well as formative and summative assessments within the classroom. By examining these data, teachers begin to develop a sense of how and why students are performing the way they are and can collaboratively plan interventions for each student, whether remediation or extension is needed to meet the student's needs. As more data are collected throughout the year, no matter how big or small, the data must be analyzed for changes in students' performance, and adjustments must be made to help students to continue to grow.

Parental involvement in the school community "improves student achievement, reduces absenteeism, and restores parents' confidence in their children's education" (Garcia & Thornton, 2014, para. 5). Students whose families remain actively engaged in the school tend to earn higher grades, perform better on assessments, and exhibit better social skills and classroom behavior (Garcia & Thornton, 2014). Parental involvement can take many forms, including the active, in-person participation at the school, but it may also include monitoring of online grades and reports, setting goals with children, developing a relationship with teachers, and advocating on behalf of the school to the greater community (Garcia & Thornton, 2014).

Making a concerted effort with bubble students, those who are performing just at or below the threshold for proficiency, can help to boost their scores and the school's overall performance on an assessment by moving these students whose scores are approaching expectations to over the mark (Marsh et al., 2006). Because these students are close to meeting the criteria, they are likely to benefit from short-term, but intense remediation of missing skills and earn a qualifying score the next time they take the assessment. However, focusing on these students at the exclusion of students who are far below the threshold for success is a dangerous practice, as well, as these students could also benefit from remedial efforts, potentially enough to also boost these students over the mark necessary to meet the state's expectations. Instead of focusing on one group exclusively, educators would be wise to include all students, even those that are high-achieving, moving all students along the continuum of progress, though working diligently and intentionally so that those close to meeting expectations will, in fact, meet them (Marshall, Kane, & Wilson, 2015).

Monitoring students' progress through the use of benchmark assessments can provide additional indicators as to how students will perform on a high-stakes assessment. Because benchmark assessments are "assessments administered periodically throughout the school year, at specified times during a curriculum sequence, to evaluate students' knowledge and skills relative to an explicit set of longer-term learning goals" (Herman et al., 2010, p. 1), teachers can make note of students' progress towards the end-of-course goal, which may be a high-stakes test, such as the MEA for Science. When these tests are constructed in a way that not only mimics the structure of the questions on an assessment such as the MEA for Science with regard to the questioning stems, but also the rigor of the questions, students are likely to perform better on the actual assessment. These benchmark assessments provide information about students' ability to retain information, as the assessments are typically cumulative in nature. Likewise, when authentic questions are used, students become familiar with the format of the test and are more relaxed when they take the high-stakes assessment later on. If constructed properly with clear intentions with regard to the purpose of the test and the users, benchmark assessments are useful for informing instructional practice and policy, as well as decision-making at all levels (Herman et al., 2010). Constant monitoring of student work is also critical to the success of students on high-stakes assessments. Rather than being reactive to students' shortcomings and offering remedial work, teachers should be pro-active and offer supports and interventions throughout lessons so that students don't first fall behind and then be expected to catch up.

One of the most critical steps that educators should take to ensure that students are prepared for high-stakes testing is to ensure that instructional practices are well-aligned to the standards assigned to the course. A common way to accomplish this is by "unpacking" a standard— "the process of taking the text of each standard and translating it into actual teaching strategies" (WiseWire, 2016, para. 1). By doing this, teachers become aware of the knowledge and understandings students must possess in order to be deemed "proficient" with the standard, but the associated skills can also be identified, which is critical in understanding how students will demonstrate their mastery of the content. Once it is understood what is to be taught and understood by students, teachers are better able to match activities and instructional practices to meet students' needs.

Another common way that teachers can improve instructional practices is to understand how the standards are vertically aligned from one course to another throughout a student's educational experiences (Case, 2005). In the case of the MEA for Science, this is particularly important when examining student's Grade 9-11 course offerings, as the standards assessed on the MEA for Science are to be taught during these three academic years. Theoretically, however, this should be happening at all levels, including with the elementary and middle school curricula, as well. If, and when, that is done, the district as a whole can ensure a vertically-aligned curriculum, creating a cohesive and thorough science education experience for students, which will aid them in achieving higher scores on high-stakes assessments.

Student Motivation

As noted in previous sections, the MEA for Science is considered a high-stakes standardized tests with regard to its necessity for federal accountability (MDOE, 2018c). However, because the MEA for Science is not linked to students' grades for a particular course or graduation requirements, it is a reasonable assumption that some students may not feel as though the test is high-stakes as other assessments, such as the SAT which is used for college admission and scholarship purposes. As such, this reduced student motivation to perform well on the MEA for Science may have a negative impact on students' scores (Wolf & Smith, 1995).

Early in their academic career, students tend to be much more motivated and engaged in their education than they are at the end of it (Hulleman & Hulleman, 2018). That is, elementary students tend to have a higher level of motivation to perform well than their high school counterparts. Mathewson (2019) asserts that the reason for students' declining motivation as they progress though public education occurs because of changes in students' excitement about what they are learning. Early on, students are excited to be learning new things, but as they advance, they are required to learn things about which they are not passionate, changing motivation from learning for the sake of learning, an intrinsic motivation, to performing for the sake of earning a grade or credit, an extrinsic motivation (Mathewson, 2019). Wolf and Smith (1995) further state that, "when a test is of direct consequence to an examinee, that person may be more motivated to put forth a strong effort than under nonconsequential conditions" (p. 239), meaning that extrinsic motivations, especially at the high school level, can provide more motivation for students to do well than other motivators. What consequences are tied to the MEA for Science, however, whether positive or negative, are minimal, all but eliminating extrinsic motivation and most certainly reducing the intrinsic motivation, as well.

Short of changing graduation requirements and the direct relationship between students' course grades and their performance on the MEA for Science, ways to increase students' motivation lies in the hands of science teachers at Hartman High School. However, several practices within the classroom can positively impact students' motivation (Martinez, 2018; McKay, 2015). These practices include encouraging students to get plenty of rest prior to the test and to engage in stress-reducing activities, setting high expectations, building a culture of success around testing, building positive student-teacher relationships, and positively influencing student mindsets (Martinez, 2018; McKay, 2015). The key, however, to increasing students' motivation on standardized tests is to determine what is causing students to be unmotivated in the first place and to then develop a long-term plan of action so that students are motivated not only when they take the test, but throughout the entire school year (Clay, 2016).

Teacher Effectiveness

Teachers' effectiveness in the classroom has a significant impact on student performance on a variety of assessments, including standardized tests such as the MEA for Science (Darling-Hammond & Youngs, 2002). For this study, the definition provided by Johnson and Semmelroth (2014), that effectiveness is the act of consistently producing higher academic gains among students, will be applied. Though standardized tests can provide a significant amount of data regarding students' academic performance, it cannot and should not be the only source of information in determining a teacher's effectiveness (Piro, Dunlap, & Shutt, 2014). In fact, several factors contribute to teachers' effectiveness, including a strong content knowledge (DeMonte, 2015), college degrees earned (Stronge, 2018), and teaching credentials obtained by educators (Darling-Hammond, 2015). Professional development, both for new teachers and on-going support for veteran teachers, plays a significant role in teachers' effectiveness (American University, 2019). Further, factors such as the size of the class and the overall size of the school play a critical role in student achievement (Blatchford, Chan, Galton, Lai, & Lee, 2016), as well as teachers' attendance patterns and the qualifications of substitute teachers (Okeke, Shumba, Rembe, & Sotuku, 2015). Determining how these factors contribute to students' scores on the MEA for Science is a critical step in determining how scores can be improved.

Teacher training. The instructional environment at a school, directly linked to students' academic performance, is a determining factor of the effectiveness of classroom instruction (Heck & Hallinger, 2014). When teachers have an appropriate education prior to beginning their teaching careers, are able to demonstrate content-based competency, and possess teaching credentials required for the course they have been assigned to teach, they are more likely to construct and deliver high-quality lessons for their students (Curry, Reeves, McIntyre, & Capps, 2018; Rice, 2003), increasing students' understanding of skills and concepts measured by high-stakes tests, such as the MEA for Science.

Studies have indicated that students are more likely to be successful in a given field if their teacher possesses a college degree in that area of study rather than a generalized teaching degree (University of Missouri-Columbia, 2018). This is particularly true for high school-level courses, which are typically highly-specialized and focus on discrepant skills that are more likely to be used by students in future endeavors. That is, skills and understandings garnered in these courses are very specific as compared to the generalized topics of the elementary and middle school experiences. Because of this, the more aligned a teacher's degree is to the content to which they are assigned to teach, the better the academic outcome for their students.

Often, especially at the high school level, teachers are either career-switchers or do not have a formal degree in education. These teachers have a degree in their discipline followed by education courses that meet the requirements for teaching in a public school system. Many of these teachers have experience in industry prior to teaching; that is, they have used their degree for employment outside of the world of education (Wilcox & Samaras, 2009). Still, others have a degree in a field unrelated to the subject they teach, but they have acquired the credentials necessary to be employed as a teacher. Alternative paths of certification have grown in popularity, as a critical shortage of teachers in certain content areas and in urban settings has remained an issue, along with the high attrition and retirement rate of teachers (USDOE Office of Innovation and Improvement, 2004).

Students whose teachers who have taken an alternative path to licensure are not necessarily at a loss when it comes to the instruction they receive. In fact, the pre-service experiences their teachers have had can greatly enrich their educational experiences in the classroom (Lynch, 2016a). The difference, however, lies in the content delivery methods utilized by the teacher. Career switchers often start their teaching careers without a full gamut of education courses, limiting the range of the instructional models with which their teachers may be familiar. As such, instructional delivery can be limited to only a few strategies; targeted interventions may be lacking for students who are struggling to learn content and enrichment opportunities may be lost for advanced students. The overall instructional planning is often limited for teachers that enter the profession through alternative routes (Lynch, 2016a), and the outcome of poor planning lies in the poor execution of lessons, greatly impacting students' performance on assessments such as the MEA for Science.

Regardless of the method by which teachers enter the profession, all teachers are required to be highly-qualified (USDOE, 2004). In order for a teacher to be considered highly qualified, he or she must have earned a bachelor's degree and full state certification, and he or she must prove him or herself competent in the subject which he or she teaches (USDOE 2004). For middle and high school teachers, competency must be demonstrated via a subject-matter test, a major in the subject or credits equivalent to a major, an advanced certification from the state, or a graduate degree prior to starting their careers (USDOE, 2004). For teachers already in the profession assigned to teach a new subject, additional requirements are necessary and can include prior teaching experience, professional development, and subject-matter knowledge (USDOE, 2004).

When a teacher is employed by a school and a district, it is the responsibility of both the school and the district to ensure that professional growth and development occurs throughout the teacher's employment. This professional development is the key to ensuring that best practices are used in the classroom, which can lead to improved scores on standardized assessments, such as the MEA for Science (Fischer et al., 2018). In order for professional development to be effective, however, it must be crafted and encouraged in a way that makes teachers value their experience and seek their own professional development in addition to what is offered at the school, rather than feeling that the development is forced and simply a requirement rather than an opportunity for growth. Developing a culture that supports a growth mindset in a school is imperative to the growth of the individuals within the organization (Dweck, 2000), and because intelligence is malleable, it is critical that the expectations for growth is cultivated. Administrators must empower teachers, and teachers must encourage one another to grow if students' academic achievement is to improve.

Professional development. Professional development, whether in-house, workshop-based, online, or full-term courses, helps teachers to stay abreast of current trends in education and empowers them to use a variety of strategies to help their students learn. When teachers are able to correctly implement a variety of instructional practices, they are better able to increase student learning, which translates to improved performance on standardized assessments. As such, professional development must be targeted in order to most effectively and efficiently meet the needs of teachers; school- and district-level administrators must be able to clearly identify areas for improvement of both individuals and teaching faculties as a whole in order to plan for appropriate opportunities.

Professional development is considered effective if it "is content focused…incorporates active learning…supports collaboration…uses models of effective practice…provides coaching and expert support…offers feedback and reflection…is of sustained duration" (Darling-Hammond, Hyler, & Gardner, 2017, para. 4). Administrators must also ensure that professional development is executed in a way that is most beneficial for teachers. This includes a variety of methods that are implemented in a frequent and ongoing manner, focused on both general instructional practices as well as those that are course- and content-specific, intensive and continuous, and monitored for success (Kosanovich & Rodriguez, 2019). Failure to ensure that these considerations are made when developing and implementing professional development can lead to ineffective practice of teachers, resulting in a negative impact on students' scores on assessments such as the MEA for Science.

New teacher support. Regardless of the path by which teachers enter the profession, those who are new to their role in the classroom certainly need support as they become familiar with their new charge. Teaching is no longer simply delivering instruction; rather, teachers must also be well-versed in mediation, discipline, customer service, documentation, differentiation, accommodation, data analysis, committee work, technology, and a seemingly endless list of other duties that are otherwise undefined in most job descriptions. Learning how to manage these responsibilities, on top of the expectation to provide an education to students, linking prior knowledge to new material in an effort to prepare them for their futures, is difficult for even the best and most prepared teachers. A purposeful new teacher support program can be the key to minimizing the stress caused by these responsibilities, helping teachers to build efficacy and competency, while minimizing the burnout that is often experienced by teachers overwhelmed as they enter the profession (Ingersoll, 2012).

In fact, research has shown that new teacher support programs, or new teacher induction programs, can be the key in retaining high-quality teachers. New teachers that received no support as they entered the profession have a predicted turnover rate of 41%, while those with basic support, which includes mentoring and a supportive administrator, have a 39% probability of turnover, and those who participated in a more comprehensive support program has a predicted turnover rate of 18% (Johnson, Berg, & Donaldson, 2005, as cited in Solomon, 2016). Additionally, new teacher support programs not only help to improve retention rates, but such programs have also resulted in a marked improvement in the quality of instructional practices used by new teachers (California County Superintendents Educational Services Association, 2016); transferring skills learned during teachers' college experiences can be challenging without the support structures of a formal induction program (Kielwitz, 2014). The combination of teacher retention and the use of high-quality instructional methods often produces higher levels of student achievement (California County Superintendents Educational Services Association, 2016).

New teacher support programs come in a variety of shapes and sizes, as well as differences in which teachers are invited to participate. While it is critical that each program matches the school, district, and teachers that it intended to serve, there are common characteristics that should be shared among all programs. The Massachusetts Department of Education has set forth standards by which induction programs in the state must abide (MDOE, 2002). From these standards, recommendations as to what components induction programs should contain were developed. These recommendations include a new teacher orientation that provides pertinent information about the school and district, mentoring relationships that can involve observation and collaborative teaching and planning, support teams upon which new teachers can rely, workshops specifically targeted to the needs and concerns of new teachers, and evaluation so that new teachers can develop self-reflection skills in their efforts to grow as professionals as outlined by expectations, standards, and processes (MDOE, 2002). According to Huling-Austin (1990, as cited in Stansbury & Zimmerman, 2000), the goal of such programs is

> improving teacher performance, increasing the retention of promising beginning teachers, promoting the personal and professional wellbeing of beginning teachers, satisfying mandated requirements for induction and/or licensure, and transmitting the culture of the system to beginning teachers.

Though the goals of new teacher support programs are universally applicable, the manner in which the goals are met depends on the needs and resources of individual schools and districts. The most comprehensive new teacher support programs are those that not only introduce new teachers to the ins and outs of their new teaching assignment, but ones that also continue throughout the year and are sustainable and responsive to teachers' needs (Haver, n.d.). Some programs are lengthy, with new teachers participating for up to their first three years, while others are very short, occurring at the beginning of the school year, sometimes prior to the arrival of returning staff.

One of the most notable induction programs is the New Teacher Center based in California. The New Teacher Center (NTC) is a non-profit organization that strives to provide support structures and resources to schools and new teachers to help improve instructional practices in four key areas: "student learning, educator effectiveness, leadership development, and optimal learning environments" (New Teacher Center, 2019, para. 2). The organization focuses on cycles of support, during which observations are conducted and feedback is provided to new teachers by a mentor teacher that has received training from the organization. These observations are available in an online portal (New Teacher Center, 2016c) and include selective scripting, seating chart: movement, interaction, and behavior patterns, and content, strategies, and alignment (New Teacher Center, 2016b). Additionally, forms and protocols are available for new teachers that focus on analyzing student work, lesson plans, and parent communication (New Teacher Center, 2016a).

The New Teacher Center recommends that all new teachers participate in an induction program for the first two to three years of their careers at a minimum (Jacobson, 2018). Other programs, such as Greenville Public School District in Mississippi, recommend three to four years of support, with the expectations and experiences changing from year-to-year based on teachers' individual development (Greenville Public School District, n.d.). Still, other programs offer formal support for teachers only during their first year of employment with a district, though the relationships formed during the process can foster further development for years to come.

While full-fledged induction programs are not feasible in all situations, all programs should, at a minimum, include a mentoring component. Mentors should be paired intentionally, using the best teachers to help guide new teachers as they navigate their first year. These pairings can be determined by proximity of classrooms, similarity in content, and commonality in educational background, among other characteristics. However, according to Steelman (2018), the best partnerships result when the mentor and mentee teach at least one common course. This allows the partnership to focus on instructional strategies and classroom management techniques, both in general and for specific lessons (Steelman, 2018). Mentoring relationships also thrive when partners are able to meet frequently, are reflective, and encourage positivity, despite the vulnerability of such a relationship (Steelman, 2018).

Teachers that feel supported are more likely to remain in the field, and studies have shown that these programs help to greatly reduce the attrition rate of new teachers, which can be as high as 40-50% of new teachers (Grissmer & Kirby, 1987, 1992, 1997; Hafner & Owings, 1991; Ingersoll, 2003; Murnane, Singer, Willett, Kemple, & Olsen, 1991, as cited in Ingersoll & Strong, 2011). Because the correlation between new teacher support programs, teacher retention, and student performance is strong, it is necessary to determine the level to which teachers at Hartman High School feel supported, not only currently, but also the level of support that they received upon starting at the school in order to solve the problem of low assessment scores on the MEA for Science.

Class and school size. The size of the class in which students are enrolled can have a direct impact on the instructional experience students have. When relationships between students and teachers flourish, opportunities for remedial and enrichment opportunities become available, behavioral challenges subside, and student engagement increases (Higgins, 2014). Unfortunately, staffing is limited in many schools due to budget constraints, causing class sizes to grow and students' learning and academic achievement to be compromised, as measured in a number of ways, including standardized assessments such as the MEA for Science.

Students' perceptions and sense of belonging within a classroom can be a determining factor of their academic achievement. When meaningful relationships exist between students and their teacher, students are more likely to want to please their teacher, resulting in improved behavior (Boynton & Boynton, 2005) and a likeliness to engage in challenging activities that are outside of their comfort zones (Thompson, 1998). Students feel more confident taking risks when trying new skills and activities, while being more engaged in their learning, out of a desire to please their teacher and avoid negative interactions (Thijs & Fleischmann, 2015). Because of this, students are more likely to deepen their understandings, which can translate to higher assessment scores.

In schools where student-to-teacher ratios are small, students are more likely to receive one-on-one instruction, which can benefit struggling learners as they receive remedial support, as well as high-achieving students who engage in enrichment activities (Cuseo, 2007). These individual interactions with the teacher provide the opportunity for relationships to develop between students and the teacher; positive relationships as such can foster an intrinsic motivation in students that develops from the extrinsic motivation to impress the teacher (Kalenze, 2016). Therefore, positive relationships between teachers and students, when given the time to develop and be cultivated, can increase students' achievement; determining the extent to which teachers feel relationships exist at Hartman High School informed the recommendations for improving the scores on the MEA for Science at the school.

Students' behavior and level of engagement both tend to improve in smaller classes (Finn, Pannozzo, & Achilles, 2003). This decrease of behaviors that take away from students' learning and the increase in student engagement allows students to better interact and understand new learning. This is particularly true for low-achieving students; with its results shared at a conference in March of 2008, a British study found that smaller class sizes are particularly beneficial in supporting academic success for these students (Viadero, 2008). Further, research has shown that students in smaller classes can be as many as two months ahead of their peers that are in larger classes (National Council of Teachers of English, 2014); these same students were shown to perform better on assessments. However, Osborne (2018) cautions against putting too much stock in the impact of smaller class sizes on performance, as studies of the impact of smaller class size on achievement by both Hoxby (1998) and Chingos (2010) show that smaller class sizes do not necessarily correlate to better performance on standardized assessments.

Teachers with small class sizes are more likely to feel less overwhelmed by their teaching responsibilities, spending less time grading for the sake of grading and devoting more time to offering meaningful and growth-minded feedback to students (EF Academy, 2019). Meaningful and timely feedback provides students with information as to whether or not they are appropriately learning the new skill or task at hand; in order to be effective, however, it must be as specific as possible, given immediately, address the student's work towards a specific goal, be presented purposefully, and involve the student (Stenger, 2014). In addition to providing the opportunity for teachers to provide this sort of feedback, smaller class sizes also foster a community in which students are open to receiving feedback from their peers, and the management of this process is much more successful with smaller groups of students rather than larger groups. This feedback can help students to grow in their academic achievement as measured by a variety of indicators, including the MEA for Science.

Teachers with smaller classes are also more likely to reflect on their lessons in meaningful ways, seeking to improve their craft, not only for the next time they teach the same lesson, but also for the next time they teach any lesson. This introspective approach to improvement allows the teacher to focus on areas that she feels will create the best learning opportunities for her students (Shandomo, 2010). This can translate to more buy-in with presented with professional development or other opportunities that allow for personal growth that will be transferred into practice. This practice can greatly enhance a teacher's effectiveness, which is then evident in students' learning.

Small class sizes can also lead to an increase in teacher effectiveness and efficacy because teachers are better able to form relationships with their students. Logically, the fewer students teachers have in their classroom, the more one-on-one time they can effectively spend with each student, fostering a relationship that empowers students to challenge themselves academically (Thompson, 1998; Thijs & Fleischmann, 2015). These relationships are able to bolster students' confidence when it is waning in the face of an obstacle and are part of the celebration when students find success. When these correlations are recognized by teachers, their own confidence soars and enthusiasm for teaching climbs, both of which translate to better instruction and better academic achievement for students (Ware, 2002).

Teacher attendance. Just as students' attendance in imperative to their success at school, so is a teacher's attendance necessary for the success of her students. When these absences occur with enough frequency, students' learning is not only disrupted, but their overall academic performance is at risk (Okeke et al., 2015). Most notably, when teachers are absent in the weeks and months immediately prior to testing, students' scores tend to suffer considerably. Though substitute teachers are hired in the absence of a full-time educator, lessons are far less effective with a substitute teacher present than when the regular classroom teacher is in attendance (Okeke et al., 2015). Therefore, this study seeks to determine if teachers' attendance has a negative impact on the MEA for Science at Hartman High School.

Research currently demonstrates that teacher absenteeism does, in fact, have a negative impact on students' performance on student achievement. In fact, Miller, Murnane, and Willet (2007) found that as few as 10 teacher absences can have a detrimental effect on students' performance on standardized assessments; what is alarming about this finding is that each teacher misses an average of 10 school days per academic year. Though these 10 days do not necessarily occur consecutively, the impact it has on student learning often exceeds a 10-day deficit (Whelan, 2008). While some studies, such as that of Ehrenberg and Brewer (1995), noted that teachers' attendance did not have a considerable impact on students' test scores, other studies have noted the significantly negative impact teachers' absenteeism has on students' assessment performance (Aucejo & Romano, 2016; Pianta & Ansari, 2018). Because of the conflicting findings in these prior studies, the current study is warranted to determine the impact of teachers' attendance patterns on students' performance on the MEA for Science at Hartman High School.

Substitute teacher qualifications. When teachers are absent, substitute teachers are typically hired on a day-by-day basis to ensure that students continue to receive their regular classroom education. While the intentions of this practice are sound, several issues lie within the hiring of substitute teachers. These issues generally lie in the lack of qualifications of substitute teachers due to inexperience with both content and classroom management techniques.

In a study by Westrick, Le, Robins, Radunzel, and Schmidt (2015), it was found that students are more likely to exhibit challenging behaviors when a substitute teacher is present than when their regular classroom teacher is in attendance. This is partly due to the change in the daily routine, the lesson plans provided, which often incorporate less-engaging activities, and a lack of connection to the substitute teacher in a way that mimics the relationship between the regular classroom teacher and students (Davies, 2019). Because a substitute teacher is usually employed one day at a time, he is less likely to form these relationships, ones that would otherwise lead to an understanding of students' academic needs in a way that can help them to be successful. This, in turn, can have a negative impact on students' learning and their later performance on standardized assessments, such as the MEA for Science.

Perhaps the biggest concern with substitute teachers is that they seldom have the full teaching credentials that regular classroom educators are required to possess. Though some substitute teachers are, in fact, retired full-time educators or individuals with full teaching credentials, but not employed in a full-time position, many possess the minimum qualifications necessary. In Maine, a substitute teacher can be employed up to 10 days in any position with only a General Education Diploma (GED) or high school diploma (National Education Association, 2019b). A more alarming detail of the substitute policy in Maine is the trend for schools to hire substitute teachers who fail to possess the appropriate content knowledge and classroom skills as long-term substitutes for various vacancies (National Education Association, 2019). Add to this detail the fact that schools do not need to notify parents and guardians of the hiring of an underqualified teacher for a position per the NCLB Act (2002), as the substitute is not hired as a full-time teacher, and students' success becomes even more at risk, as parents are unaware of the need to advocate for their children differently than if a full-time teacher were hired.

Instructional Resources

The instructional resources available for use in the classroom can strongly support students' learning in the science classroom. Textbooks, electronic resources, and consumables for hands-on learning, among other tools, are available for teachers to use in their classrooms, but without the assurance that these resources are closely aligned with the content, rigor, and abilities of students in the classroom, these tools can be ineffective (Matthews, 2012). However, when classroom instructional resources are appropriate, the resources can have as significant of an impact on students' learning as would reducing class size by 10 students (Koedel & Polikoff, 2017).

While factors that should be considered when selecting appropriate resources for classroom instructional purposes vary from district to district, many similarities exist between selection criteria. These include supporting the goals and objectives of the course, school, and district, be appropriately matched to students' age, social, and emotional development, and offer diverse applications in terms of difficulty, appeal, and points of view (Urbandale Community School District, 2019; Committee on Undergraduate Science Education, 1997). In order to understand if and how resources are matched to the goals of the course, school, and district, individuals tasked with choosing resources, whether as a committee on behalf of a larger group or individual teachers selecting tools for their classrooms, a thorough understanding of the underlying beliefs of the school and district must be present in order to ensure that resources are not misaligned. Likewise, a deep knowledge and understanding of all course standards and the way in which these standards are interwoven throughout the course will help to ensure appropriate materials are selected.

Because students have unique learning needs and desires, selecting a variety of appropriate learning tools is critical to their success in learning science. Understanding how the variety students and their abilities in all academic and social areas impact and are impacted by the use of different instructional resources provides the foundation upon which choices for instructional resources can be made. While textbooks have been the long-standing choice as the primary instructional resource, a variety of options now exist, available from both from publishing companies and in piecewise format with a careful search by educators. These resources can include the textbooks themselves, along with workbooks, worksheets, and manipulatives, such as flash cards, games, models, and activities, to help support classroom instruction. Electronic resources are also available in a variety of formats, including apps, websites, movies, podcasts, and online activities. These tools can be used in several different combinations to best meet the learning needs of students, but first, teachers must have a sound understanding of students' learning needs.

Summary

This literature provided theoretical frameworks associated with the study, including Piaget's Theory of Cognitive Development regarding how children learn in stages and the Binet and Simon's Intelligence Theory that has long served as the framework for standardized tests. Additionally, the history of standardized testing, the structure of standardized tests, the use of standardized testing in science nationwide, the history and evolution of the MEA testing process, including the use of the MEA for Science, underlying factors that impact students' performance on assessments, and strategies and interventions that can improve student achievement as measured by standardized tests was provided. These strategies include analyzing test data, involving parents in their students' education, focusing on bubble students, constant monitoring of success, examining student work, and practice tests and benchmark assessments throughout the academic year, as well as unpacking standards and ensuring a vertically-aligned curriculum. The literature shows that strategies exist that can improve student performance on standardized assessments, which serve as one of several indicators of students' proficiency as related to standards. This then beckons the question, what strategies and interventions can be implemented at Hartman High School to improve students' performance on the MEA for Science? More research is necessary to determine what strategies are already in place and which ones would most benefit students in both the short- and long-term application. The literature emphasizes the importance of research in this area; therefore, this applied research study sought to contribute to the field of knowledge and to narrow the gap in the research and literature concerning the strategies that can best aid students and teachers in achieving higher scores on standardized science assessments, such as the MEA for Science.

CHAPTER THREE: PROPOSED METHODS

Overview

The purpose of this applied research study is to identify factors that impact students' performance on the Maine Educational Assessment (MEA) for Science and to devise a plan to improve students' performance on the assessment. The MEA is part of a series of tests that "measure the progress of Maine students in the areas of English, Language Arts, and Literacy, Mathematics, and Science" (MDOE, 2018c, para. 2). This test, administered to students during their third year of high school, is a comprehensive exam that includes topics from both life and physical sciences. Historically, the scores on the MEA for Science at Hartman High School have revealed that 51.85% of students failed the MEA for Science in the last three years (MDOE, n.d.).

This chapter defines the study and its purpose, including the research design and questions, as well as the setting in which the study will occur. Participants are described, as well, along with the researcher's role within the school and in the study. Data collection procedures are identified and are followed by the data analysis protocols to be employed. The chapter concludes with a discussion of the ethical considerations of the study.

Design

A multimethod research design will be used for this applied study, incorporating both qualitative and quantitative methods. An applied research design is an appropriate approach for this study, since "applied research uses scientific methodology to develop information to help solve an immediate, yet usually persistent, societal problem" (Bickman & Rog, 2009, p. x). This design uses both qualitative and quantitative data, the combination of which "together, provide a better understanding of your research problem than either type by itself" (Creswell, 2015, p.

537). The multimethod approach to this study will be “used to gain complementary views about the same phenomenon or relationship” (Bickman & Rog, 287), allowing the researcher to triangulate data gathered from three different sources and through three different means. The qualitative component of this study includes an interview, while the quantitative components include a Likert-scale survey, for which questions are developed based on the Literature Review, and the review of archival data, including documents from the MDOE regarding students’ performance on the MEA for Science at Hartman High School over the context of several years.

Research Questions

Central Question: How can students’ scores on the Maine Educational Assessment for Science be improved?

Sub-question 1: How would teachers and administrators in an interview solve the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine?

Sub-question 2: How would science educators participating in a survey solve the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine?

Sub-question 3: How does a review of documents inform the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine?

Setting

The site selected for this study is a suburban public high school in southeastern Maine. The school serves four towns and 739 students, of which are 52% male and 48% female (US News, 2018). The minority enrollment is 6%, and 32% are considered economically disadvantaged (US News, 2018). A total of 64 teachers serve the school, resulting in a 12:1

student-to-teacher ratio (US News, 2018). School administrators include a principal and two assistant principals, as well as a Director of Student Services.

For this study, the school's science department will be the focus. The science department is comprised of 10 teachers and offers a total of 17 courses. These courses include:

- General, Academic, and Honors level 9th Grade Science
- General, Academic, and Honors Biology
- General, Academic and Honors Chemistry
- Chemistry/Physics combination course
- Academic and Honors Physics
- Advanced Placement (AP) Biology
- AP Environmental Science

Within the department, leadership is shared among four teachers who, when responsibilities are combined, act as the department chairs. Instructional decisions, including which courses are taught and the standards associated within these courses, are generally determined within the department, though administrators have input and oversight, and graduation and students' enrollment needs are also considered.

The site is an ideal selection for two significant reasons. First, the school's poor performance on the MEA for Science in recent years makes it an ideal setting for the study. It is critical to focus on these scores and the efforts necessary to raise students' performance, as improving students' performance on this assessment may ultimately lead to increased funding and programming for the school (CNN, 2012). Doing so may also boost the community's confidence in the school and its ability to prepare students to be contributing members of society upon graduation. Secondly, because the researcher is an employee at the school, ease of access and support for this research within the school makes it an ideal site.

To protect the identity of the school and participants in this study, pseudonyms will be used. These pseudonyms include Hartman High School, Mrs. Harrison, Mr. Gilmer, Ms. Marion, Mr. Preston, and Mrs. Lewis.

Participants

For this study, nonprobability sampling will be used. Nonprobability sampling includes the selection of individuals as participants because they are “available and convenient and represent some characteristic the investigator seeks to study” (Creswell, 2015, p. 144). Such sampling is “best used to provide information about specific cases or members of the study population that are intrinsically interesting or important for the study” (Bickman & Rog, 2009, p. 79). That is, purposeful sampling, the selection of participants and ties in order to better understand a central phenomenon (Creswell, 2015), will be used to select the five participants for the interview component of this study from administration and the science department at Hartman High School based on their familiarity with the school’s historical performance on the MEA for Science. In this sample, three participants are males, aged 48, 58, and 60, and the two other participants are females aged 37 and 59.

The survey will be sent to a total of 15 participants. This sample includes the five interview participants, along with three additional teachers in the department, each of which has been teaching at the school for three years or less, two special education collaborative teachers for the science department, the school’s testing coordinator, who is also a school counselor, a retired science department chair with 30 years of experience at the school, and three additional administrators, two of which have been at the school for three years or less. The additional participants in this sample include four males and seven females. Three participants’ ages were in the 21-29 range, four each in the 30-39 and 40-49 ranges, one in the 50-59 range, and three in the 60 and older range.

These added teachers will be suitable candidates for the survey portion of the study, as they have a direct influence on students' performance on the assessment, but they are not ideal participants for the interview, as their familiarity with and influence of prior years' scores is minimal.

The Researcher's Role

Because I, the researcher, am a science teacher at Hartman High School, the motivation for this study lies in an effort to improve my students' education as evidenced by their performance on the MEA for Science. Improved scores not only provide evidence of the students' learning, but the increase in scores may lead to additional funding for the school and its programs from various entities. Improved scores also build the community's confidence in the school's ability to provide an education to students that will equip them for their future endeavors.

The participants in this study are my colleagues. The participants will be made aware of my intentions of this study, as well as the importance the research has with regard to future instructional efforts to improve MEA for Science scores at Hartman High School.

Bias and assumptions that will be brought to this study by myself are a result of my position within the school and previous employment experiences. As a science teacher at Hartman High School, I am aware of the need to improve students' performance on the MEA for Science, as well as the positive implications this can have for the school and my position within the school. Because of the researcher's position within the science department, bracketing will be used to remove the researcher's own perspective from the findings of the study (Creswell &

Poth, 2018). The researcher assumes, based on casual conversations and interactions prior to this study, that a systematic approach to improving students' performance on the MEA for Science has not been implemented.

I have also been employed as an instructional coach for science with a previous school district. This position allowed me to offer suggestions to science teachers, science departments, and administrators regarding daily instruction, assessment, and data analysis; this experience honed my skills in identifying areas for improvement and implementation processes.

In this study, my role will be multifaceted, just as the study itself uses a multimethod approach. I will be the interviewer, responsible for developing the interview questions, transcribing the interviews, and noting the tone and inflections throughout each interview. I will also be required to ask follow-up questions in a timely manner during the interview. During the qualitative component of the study, participants will be at ease in my presence due to my employment within and the camaraderie among the department. For the quantitative components of the study, I will review archival data in the form of test scores and analyze the results of the Likert-scale survey to be administered to all science teachers in the science department at Hartman High School.

Procedures

Prior to beginning the process of data collection and analysis, the researcher will secure approval for the study from the Institutional Review Board (IRB) at Liberty University (see Appendix A for IRB approval). This process will include the approval and defense of the research proposal, as well as a review of the research by the IRB (Liberty University, n.d.). The IRB review process requires the researcher to complete the General IRB Application and Signature Page, which will then be sent to the researcher's Dissertation Chair. Once the documents are approved by the Chair, they will be sent to the IRB by the researcher (Liberty University, 2019b). Any revisions suggested by the IRB will be made to the study prior to gathering data (Liberty University, 2019b).

Permission will also be secured from the principal of Hartman High School, as well as the superintendent of the district, to conduct the research at the school and to utilize information available regarding the school's performance on the MEA for Science. See Appendix B for permission request letter and permissions. Participants will be elicited via personal communication with science teachers and administrators at Hartman High School, and consent forms will be secured from each participant prior to beginning data collection. See Appendix C for consent forms.

Data Collection and Analysis

An applied research study requires the use of a combination of both qualitative and quantitative approaches, with at least one of these data collection methods being an interview (Liberty University, 2019a). This combination is beneficial, as each approach brings different value. For instance, qualitative approaches allow participants' thoughts, feelings, and ideas to be considered in the context of the study (Denzin, 1989, as cited in Rahman, 2017). In contrast, quantitative approaches tend to reduce bias when collecting and analyzing data, producing reliable, valid, and generalizable results (Dowd, 2019). For this study, an interview, Likert-scale survey, and review of documents will be used as the primary data collection methods. Each of these methods will address one of the research sub-questions, and the data collected will be triangulated to develop a solution for the problem of low test scores on the MEA for Science at Hartman High School. The triangulation of data will allow the researcher to find common themes among data, including disconfirming evidence, helping to limit the researcher's bias in the study (Creswell & Poth, 2018).

Interview

The first sub-question for this study will explore how teachers and administrators would solve the problem of low scores on the MEA for Science at Hartman High School located in southeastern Maine. To answer this question, data will be collected in a qualitative manner via a 16-question semi-structured interview. This approach is appropriate for this study, as it allows for the researcher to construct an interview guide based on specific topics to be investigated, and it allows the researcher to format questions in a way that incorporates previous answers and each participant's experiences (Bickman & Rog, 2009). This interview will be conducted with five participants at Hartman High School, all of which are either administrators or teachers in the science department. The participants will be selected based on their historical familiarity with the MEA for Science and the results for Hartman High School.

The interviews will be conducted off-campus in a one-on-one, face-to-face format. Each interview will last approximately one hour and will be recorded and immediately transcribed for data analysis. Throughout the interview, the researcher will make note of body language and tone demonstrated during the interviews to further define the participants' thoughts and feelings towards each topic included in the interview.

Initially, the purpose of the study will be explained to the participants to ensure that they clearly understand the risks that may be associated with the study, as well as the intention of the results of the study. After participants consent to participate in the study, the following 16 questions will be asked:

1. What skills are assessed with the Maine Educational Assessment (MEA) for Science?

This question aimed to determine educators' awareness of how the MEA for Science is aligned with Next Generation Science Standards (NGSS). Understanding the alignment of an assessment with standards in terms of rigor and content is critical to ensuring the assessment truly measures students' abilities with the material and skills (Carnegie Mellon University, 2019). By asking this question, the researcher will learn how familiar teachers are with the test and standards, particularly which concepts are assessed with the test.

2. How are these skills incorporated through students' 9-12 curriculum map?

   This question aimed to determine educators' familiarity with the standards included on the MEA for Science and how they are incorporated into the science course offerings at Hartman High School. Ensuring that all concepts to be assessed are taught at an appropriate rigor level throughout the courses that all students are required to take at Hartman High School is necessary to provide students the experiences necessary to be successful on the test (Drake & Burns, 2004). Likewise, understanding the vertical alignment of science course offerings at Hartman High School creates a cohesive experience for students prior to sitting for the assessment (Case, 2005).

3. How do the day-to-day instructional strategies in your classroom prepare students for the assessment?

   The purpose of this question is to uncover the extent to which teachers work to unpack standards within their course. This practice helps to ensure that the instructional practices are aligned with both the content and the rigor required of the standards (Wiggins & McTighe, 2005).

4. How do the instructional resources available, including textbooks and other in-class materials, as well as electronic tools, support daily instructional practices?

   This question is intended to determine the extent to which the current instructional materials support daily instruction. Appropriately-aligned resources are imperative for student success, but must be aligned to the course content, rigor of the standards, and abilities of the students in class (Koedel & Polikoff, 2017).

5. How do students demonstrate mastery of these skills and concepts in your class?

   The purpose of this question is to uncover how formative and summative assessments are used in each science course. The use of such assessments can serve as a predictor of students' performance on high-stakes assessments such as the MEA for Science (Herman et al., 2010). When assessments are used in a formative manner, teachers are able to offer interventions to students to improve their learning (Wiliam, 2018), which is measured by the MEA for Science.

6. The MEA is administered during a student's 11th grade year of high school and is a cumulative exam. How are skills needed to be successful on the MEA learned during the 9th and 10th grade years reflected in the curriculum map?

   This question seeks to determine what review strategies, if any, are used with students to help them prepare for the MEA for Science. Doing so can help to reduce test anxiety, gain confidence, and study more effectively (Nest, 2019). It also helps students call to mind information that was learned in prior courses.

7. What specific preparation is offered to students prior to taking the MEA?

   This question seeks to determine what review strategies, if any, are used with students to help them prepare for the MEA for Science. Doing so can help to reduce test anxiety, gain confidence, and study more effectively (Nest, 2019). It also helps students call to mind information that was learned in prior courses.

8. For the past three years, our students have scored in the bottom half (and sometimes the bottom quarter) of all schools in the state. What factors do you believe contribute to such low test scores?

   This question called for educators to consider what factors might be negatively impacting students' performance on the MEA for Science. While some of these factors are beyond the control of the school, several may be addressed by making small adjustments to instructional practices and mindsets (Teaching Tolerance, 2019).

9. When the assessment results are shared with the department, what do you think should be the next steps for teachers?

   This question aims to identify ways that teachers have used the test data to make instructional decisions. When teachers work collaboratively to analyze test data, they can better understand the strengths and weaknesses of students overall, which can lead to an adjustment in instructional practices to help improve students' performance future (National Association of Elementary School Principals, n.d.).

10. What strategies have been explored to increase students' performance on the test?

    This question seeks to identify what strategies have been considered in the past to improve students' performance on the MEA for Science and why they were or were not chosen for implementation. Several strategies exist (Teaching Tolerance, 2019), but not all strategies are the correct fit for all students, making the selection of strategies a critical practice to ensure students' learning needs are being met.

11. How were these strategies decided upon?

The purpose of this question is to determine how teachers work together to make instructional decisions, whether by content or as a department as a whole. Working collaboratively provides teachers the opportunity to further enhance their practice and increase student performance (DuFour, 2004).

12. How were they implemented?

This question also seeks to identify whether work is completed independently or collaboratively among teachers and with what fidelity. It is not enough to simply choose appropriate interventions for students to improve their learning and; therefore, their test score; educators must make sure that the strategies are implemented in a way that is effective (National Center on Response to Intervention, n.d.).

13. How do you think students feel about taking this assessment and what evidence do you have to support your opinion?

This question is intended to collect data regarding educators' perceptions of students' motivation to perform well on the MEA for Science. Because the MEA for Science is not directly linked to any consequence, whether positive or negative, for students, they may not be as motivated to do well (Wolf & Smith, 1995). Further, this question sought educators' perceptions of how students' motivation impacts their performance on the assessment. Students that are motivated would likely be disappointed by their poor performance, but if students are not motivated to do well on the assessment, they will likely be indifferent towards their scores (Tyner, 2018).

14. In what ways do you feel that teacher effectiveness impacts students' performance on the MEA for Science?

This question allows participants to reflect on their practice, as well as that of their peers, with regard to students' performance on the MEA for Science. Such reflection will reveal what characteristics each participant feels define effective instruction, as well as what can be immediately changed to improve instruction (Sierra, 2015). Reflective practice and defining expectations and desirable qualities in teachers is critical to the growth of an individual and the growth of the science department as a whole, which translates to better instruction for students and possibly higher scores on the MEA for Science.

15. What factors positively impact students' performance on this test?

    This question will allow educators to pinpoint what they feel is working well at Hartman High School with regard to students' performance on the MEA for Science. These are likely to be things that will remain unchanged in the future and bringing to light the positive components already in place is key to improving the areas in which instructional practices and student performance are falling short. These successes should be celebrated and maintained as the more challenging work of changing factors that have a negative impact on students' performance begins (Battelle for Kids, 2011).

16. What other thoughts or feedback about how students' scores on the MEA could be increased at our school could you add?

    This final question allows participants to provide any other thoughts that they have about the MEA for Science and why scores are as low as they are. Using this open-ended format for this question allows for "richer and more extensive" data to be collected than the close-ended questions (Bickman & Rog, p. 264). The information provided by participants in response to this question may create additional follow-up questions, particularly if the topic has not been addressed otherwise during the interview.

Once interviews are transcribed, the transcripts will be reviewed, and coding will be used to determine the categories and themes present in the participants' dialogue. Coding and categorization are an appropriate data analysis method, as it allows the researcher to align the participants' responses with the literature related to the study (Creswell & Poth, 2018). This data analysis method also lends itself to the creation of a codebook, from which other researchers can determine what constitutes an entry under a specific theme, along with what does not (Creswell & Poth, 2018). Generally speaking, coding involves reading the transcript from each interview and identifying the various themes that are represented throughout. First, the researcher will make note of the various words or phrases throughout the interview related to the study, identifying specific quotes that support the codes (Creswell, 2015). When each transcript has been coded, the codes from each transcript will be combined and categorized into themes based on similarity, reducing the codes into a smaller number of categories to be analyzed. The frequency of each of these categories will be reported on a question-by-question format in tabular form in Chapter Four.

Survey

The second sub-question in this study explores how science educators participating in a survey would solve the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine. To collect data to answer the second sub-question, a close-ended survey will be administered electronically using Google Forms, an internet-based program. This is an appropriate approach, as it provides participants with the opportunity to have one and only one answer to each question, while ensuring that there is, in fact, an answer to every question (Bickman & Rog, 2009). The survey will include 15 statements developed from the literature review, to which survey participants will respond using a 5-point Likert scale rating. The statements include:

1. Before teaching a new concept, the associated Next Generation Science Standards (NGSS) are intentionally built into instruction.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

This question is intended to identify teachers' efforts to implement standards and understand the underlying skills, concepts, and understandings that students must have prior to learning the new material, as well as what competencies should be present when students have successfully mastered the idea (Wiggins & McTighe, 2005).

2. Classroom instruction is delivered with the same rigor specified in the NGSS standards.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

This question seeks to identify how well-aligned instruction is with the standards and the rigor of each. Ensuring that students are learning material, albeit scaffolded at the onset, at the same level at which they are expected to demonstrate mastery, leads to better standardized test performance (Drake & Burns, 2004).

3. All students are assessed in the same manner across grade levels.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

Common assessments allow teachers to have meaningful discussions regarding the data collected and foster collaborative efforts to improve all students' academic performance

(Guskey, 2003). This question seeks to identify the frequency with which this type of assessment is used.

4. Time is provided during contract hours for collaboration between teachers of the *same* courses.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

Meeting with other teachers of the same course helps teachers to disaggregate data and provide insight into students' performance. This question is intended to identify the frequency of these meetings that occur during contract hours, when teachers are most likely to analyze data effectively and complete the task. Further, during this time, instructional decisions can be made, and the collaborative planning that occurs during these meetings can lead to improved lessons for all teachers (DuFour, 2004).

5. Data collected from assessments are used to plan future instruction.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

This question helps to determine the manner in which assessment data are used to plan instruction. When formative, as well as summative, assessments are used to determine students' understanding of concepts, instruction can be tailored to students' needs, increasing performance on standardized assessments (Wiliam, 2018).

6. Content taught in each course is clearly communicated throughout the department.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

This question seeks to identify the level of vertical alignment within the science department at Hartman High School. Vertical alignment within the department ensures that time is allocated in an appropriate manner to each standard throughout a student's science progression. When too much or too little time is spent on a concept depending on a students' prior knowledge of the material, performance on assessments can be negatively impacted. Understanding how and when concepts are taught throughout the science sequence at Hartman High School creates a comprehensive educational experience for students, improving their performance on the MEA for Science (Case, 2005).

7. Professional development focuses on new instructional strategies has been offered to science teachers.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

This question is intended to pinpoint the frequency of introduction of new instructional strategies for teachers. While a number of options exist for instructional methods, it is important that teachers are well-versed in a variety of approaches, including those that are less common in the district, as well as those that are new in the world of education (Fischer et al., 2018).

8. Teachers participate on a regular basis in meaningful data analysis conversations.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

This question seeks to identify the level of engagement between teachers specifically related to data analysis. When teachers discuss the data gathered from common lessons and assessments, the conversations can provide insight as to what instructional strategies have been effective, allowing teachers to learn from one another in an effort to improve their own practice (DuFour, 2004). This collaborative aspect of teacher interaction can improve the instructional quality of courses, leading to improved test scores on standardized assessments.

9. Students are excited and ready to learn in my class each day.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

Students' motivation can be highly influential on their performance on a standardized assessment (Wolf & Smith, 1995). As such, when students are excited to learn and engage in class, they are more likely to perform better on standardized assessments. This question seeks to identify the level of excitement, as perceived by teachers, of students at Hartman High School.

10. The science department at Hartman High School is made of teachers that are highly-qualified educators.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

This question is intended to determine the overall quality of the science department at Hartman High School. When teachers are highly-qualified, the quality of their instructional practices is better, leading to better student achievement (Darling-Hammond, 2015).

11. On-going support is provided to new teachers.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

When new teachers are part of an on-going induction program, they are more likely to remain in the profession, improve their practice, and have higher-achieving students (Ingersoll, 2012), thus this question seeks to identify whether or not teachers believe that new teachers are provided with on-going support.

12. Professional development that serves the instructional needs of teachers is provided.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

This question is intended to pinpoint the alignment of professional development with the efforts of teachers in the science department at Hartman High School. Responsive professional development can help to build teachers' effectiveness by delivering timely and targeted training to build their skillset and improve instructional practices (American University, 2019).

13. Teachers in the science department at Hartman High School have high attendance rates.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

This question seeks to determine the attendance rate of teachers within the science Department of Hartman High School. When teachers frequently miss school, the achievement of students is severely compromised (Okeke et al., 2015).

14. When teachers are absent from school, high-quality substitute teachers are hired.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

This question is intended to determine the quality of the substitutes hired at Hartman High School should a teacher need to be absent. When substitutes are not well-qualified, the instructional component of a substitute lesson plan is often compromised, having a negative impact on student achievement (Davies, 2019).

15. Class sizes in the science department at Hartman High School are conducive to student learning.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

This final question seeks to identify teachers' perceptions of the appropriateness of class sizes as it relates to student learning. When class sizes are small, students are more likely to have positive experiences and increased performance on a variety of assessments (Cuseo, 2007).

All nine science teachers, two special education collaborative teachers, and a retired science department chair, along with four administrators, at Hartman High School will receive an email with instructions as to how to complete the survey. Participants will be given a two-week window in which to complete the survey; if more time is needed, arrangements will be made.

This survey serves as the quantitative component of this applied study. As such, results will be determined by calculating the frequency of each number reported on the Likert scale on a question-by-question basis, as well as the average score reported by all participants for each question. This data will be displayed in tabular format in Chapter Four.

Document Review

The third sub-question for this study explores how a review of documents can inform the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine. This review will produce quantitative results, providing evidence of students' historical performance on the MEA for Science at Hartman High School. Archival data will be retrieved from the MDOE website, where test results from 2015-16 through 2018-2019 school years' results are available.

The review of these documents will be particularly helpful in identifying similarities and differences in students' performance according to socioeconomic status, race and ethnicity, gender, and special education services received. Data gathered from this approach will be presented in a tabular format in Chapter Four according to these categories.

Ethical Considerations

Because "ethical research practice entails skillful planning and effective communication, reduction of risks, and creation of benefits," (Bickman & Rog, 2009, p. 107), the researcher will ensure that the research is well-planned and communicated and that participants will not endure risk, especially as compared to the benefit of the study. That is, the research design and procedure will be communicated to all participants, and these participants will be made aware of what risks, though minimal, may be present. Because the researcher is among the faculty at Hartman High School, objectivity is removed, as the benefits of this study apply not only to the students of the participants, but to all of the students at Hartman High School, including those of the researcher. Further, the researcher is among those capable of making changes within the school, providing a native interest in the research. However, in an effort to reduce bias, the researcher has ensured that questions are phrased in a neutral manner, avoiding guiding words or phrases and eliminating jargon (Creswell, 2015), and that participants are not coached to answer questions in any one way. The researcher will use pseudonyms to protect the identity of participants, and identifying information will not be collected during the survey process. Interviews will be conducted off-campus, which will provide an additional level of confidentiality for participants. All materials will be stored electronically with password protection. These ethical considerations are incorporated into this study to preserve the integrity of the process, collected data, and results of the study.

Summary

At the time of the study, it is known that the MEA for Science scores at Hartman High School have been in the lower third of all scores in the state (MDOE, n.d.). In fact, in the last three years over 40% of students performed "below or well-below state expectations" (MDOE, n.d.). Teachers and administrators at Hartman High School are aware of these scores, but little has been done to improve these scores. What is not known is precisely what factors have the greatest impact on students' performance on the MEA for Science.

This applied research study is intended to identify the factors that have the greatest impact on students' performance and to determine what interventions and strategies can be implemented to improve students' scores. The researcher will analyze interview data to determine what factors participants communicate as having the greatest impact on students' performance, survey all teachers in the science department, along with additional administrators and former teachers in the school, to identify the areas in which changes can be made to have a positive effect on scores, and review archival documents from the MDOE website. The researcher will make recommendations of these changes and provide a rationale for their implementation, such as providing time for teachers to analyze students' data in a group format and focusing on bubble students (Education World, 2019), allowing teachers to identify areas of strength and weakness among instructional practices, groups of students, and topics and to develop strategies to improve weak areas and to continue to build upon successes, while also identifying students whose abilities are approaching the rigor level specified by the standard who would likely benefit from targeted practices with specific skills and understandings to meet the expectations of student performance.

CHAPTER FOUR: FINDINGS

Overview

The purpose of this multimethod study was to solve the problem of low test scores on the MEA for Science at Hartman High School, a suburban school located in southeastern Maine. This method was chosen in an effort to gather input from teachers and administrators familiar with the assessment at Hartman High School, while also considering the results of the assessment available from the MDOE. Chapter Four addresses the findings of the study. The detailed research methodology was provided in Chapter Three, where the three data collection methods were outlined along with their role in answering the research questions. These methods included interviews, surveys, and a review of documents. In this chapter, the findings of the study are presented. The findings of the data analysis, which was conducted through coding and the identification of emerging themes from interviews, including field notes, results of the Likert-scale survey, and documents, are illustrated in the forthcoming tables. Data analysis was sequential and iterative, occurring over a six-week period of time. Themes emerging from this data analysis include:

1. Data Analysis
2. Instruction
3. Risk Factors
4. Standards

This multimethod research study was driven by the following research questions:

Central Question: How can the problem of low test scores on the Maine Educational Assessment for Science be solved at Hartman High School?

Sub-question 1: How would teachers and administrators in an interview solve the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine?

Sub-question 2: How would science educators participating in a survey solve the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine?

Sub-question 3: How does a review of documents inform the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine?

Participants

In this study, purposeful sampling was used due to the need to select participants that are familiar with the MEA for Science at Hartman High School. A total of 15 participants were included in the study, with the 53% of the participants being female and 67% being teachers, either general science instructors or special educators. All participants met the criteria for inclusion in at least one, if not both, qualitative components of the study. All participants were assigned a pseudonym to protect their identity.

Interview Participants

A total of five participants took part in the fact-to-face interviews. The participant criteria for this portion of the study was a tenure of three years or more at Hartman High School and a familiarity with the MEA for Science. Because of their tenure at the school, these participants have been directly involved in science instruction, analysis of the MEA scores, or both.

Mrs. Harrison. The first interview candidate was Mrs. Harrison, the building principal, a 59-year-old female with 31 years of experience in education, eight of which were spent as an administrator at Hartman High School. Prior to entering administration, she was an elementary teacher, teaching both math and English, and her administrative experience started at the middle school level. Because of her experiences at all three levels of education, Mrs. Harrison brings a unique and interesting perspective to the study.

Mrs. Harrison believe that teachers are the biggest contributing factor to student test scores. She cites teacher buy-in to testing and the data acquired through assessment has a significant impact on students' performance. Because students at Hartman High School have great respect for their teachers and are more likely to put forth substantial effort to perform well on the assessment in an effort to please their teachers. During the interview, Mrs. Harrison states

> Some kids just really don't like science, so getting them to engage is a challenge daily. But if teachers value it and they put value behind it and they talk to their kids about it, I think the investment in the assessment could be different. Can we create value to it or to the kid who perhaps failed somewhere along the line in science that it could be a credit recovery opportunity? I think there's a way to promote engagement (in a way) that you would have student buy-in as juniors. (Interview, November 5, 2019)

Mrs. Harrison recognizes the need to improve performance on the assessment, and mentioned several factors that she feels contribute to the bigger picture of student performance on the MEA for Science, including literacy skills, teacher buy-in, and relevance for students.

Mr. Gilmer. Mr. Gilmer, a 48-year-old male, is currently in his second year as a co-department chair and has been teaching at Hartman High School for 15 years, with three years as a teacher elsewhere prior. While he teaches mostly 12^{th} grade students who have already taken the MEA for Science prior to enrolling in his course, Mr. Gilmer's leadership within the science department, as well as his keen eye for data analysis, has resulted in a strong familiarity with the MEA for Science and its results at Hartman High School, along with experience discussing these results with others.

Mr. Gilmer expressed during his interview that in taking over the department head position, he has hoped to improve the preparation available to students for assessments, including the MEA for Science, and he indicates that this will continue to remain a priority (Interview, November 26, 2019).

Ms. Marion. Ms. Marion, the other co-department chair, is a 37-year-old female and has been a science teacher for 13 years, all of which have been spent at Hartman High School. Historically, Ms. Marion has taught 9th grade science and biology, both of which are courses that students enroll in prior to their junior year, during which they take the MEA for Science.

Ms. Marion recalls several iterations of the class schedule that have been in place at Hartman High School throughout the years. While only one arrangement of the schedule has been used during the three years examined in this study, Ms. Marion credits the poor scores observed during this time to, at least in part, the reduction in learning time available for students in science. In fact, Ms. Marion states, "because we've talked about how our test scores always used to be better and now they're not, I feel like an obvious thing that hasn't been looked at is [test scores] before and after the schedule change" (Interview, November 14, 2019).

Mr. Preston. Mr. Taylor, is a 58-year-old male in his fourth year as a science teacher at Hartman High School, with six years a science teacher elsewhere. Mr. Preston's current teaching assignment includes only a small number of 11th grade students that will be taking the MEA while they are enrolled in his class; most others are seniors and have taken the test prior to being in one of Mr. Preston's classes. Prior to his current assignment, he taught 9th grade science at Hartman High School.

While Mr. Preston has experience teaching both 9th and 11th grade students, he admits that his familiarity at an in-depth level with the MEA for Science is minimal. He understands that the standards assessed with the MEA are those that are taught throughout students' 9-12 science education at Hartman High School, but exactly how these standards are assessed is understood by him on a minimal level. He also believes that teachers' valuation of the assessment is directly related to students' buy-in to the assessment and its process (Interview, November 18, 2019).

Mr. Lewis. Mr. Lewis is a 60-year-old male with 39 years of experience in education, 23 of which have been at Hartman High School. He teaches biology, the course in which students are enrolled in during their 10th grade year, a year prior to the academic term in which they take the MEA.

Like Mr. Preston, Mr. Lewis admits to having a limited familiarity with the MEA for Science itself, but he does have an understanding of the material that is assessed on the test. Mr. Lewis strives to provide students with authentic assessment experiences, both in the traditional and non-traditional sense. He is particularly appreciative of the common assessments used by all teachers that teach the same course that he does and the expectation that all students achieve a minimum of 70% mastery, whether on their first attempt or through remediation efforts (Interview, November 19, 2019).

Survey Participants

Survey participants included science teachers at Hartman High School, as well as special education teachers that work or have worked directly with science, and administrators, including principals, special education administrators, and testing coordinators. Of the 16 participants, three participants are in the 21-29 age range, four are in the 30-39 range, four are in the 40-49 range, two are in the 50-59 range, and three are in the 60 or older range. Five administrators participated, along with two special educators, and nine science teachers, with one being a retired department head, and two currently serving as department head, participated in the survey. Administrators' average years of service was 10.6 in their current role, while teachers, both special educators and science teachers, had an average tenure of 10.5 years, though it should be noted that two of the participants included in this statistic are in their first year at Hartman High School. Seven of the participants were male, while nine were female.

Table 1

Survey Participant Data

Participant	**Gender**	**Age Range**	**Role**
Mrs. Harrison	Female	50-59	Administrator
Mr. Gilmer	Male	40-49	Science Teacher/Department Head
Ms. Marion	Female	30-39	Science Teacher/Department Head
Mr. Preston	Male	50-59	Science Teacher
Mr. Lewis	Male	60 or older	Science Teacher
Mr. Greene	Male	21-29	Science Teacher
Mr. Moreland	Male	30-39	Science Teacher
Ms. Cameron	Female	21-29	Science Teacher
Mr. Thomas	Male	21-29	Science Teacher
Mrs. Grey	Female	30-39	Special Educator
Ms. Temple	Female	40-49	Special Educator
Mr. Webster	Male	30-39	Administrator
Mrs. Taylor	Female	40-49	Administrator
Mrs. Campbell	Female	40-49	Administrator

Ms. Fullen	Female	60 or older	Administrator
Ms. Clark	Female	60 or older	Retired Science Teacher/Department Head; Long-Term Substitute

Results

The purpose of this study was to identify factors that impact students' performance on the MEA for Science at Hartman High School. For this applied research study, a combination of both qualitative and quantitative data collection techniques were used; interview and survey questions (Appendices D and E) were developed to align with the central research question of the study. For this study, participants were purposefully selected based on their position within the science department or as administrators or testing coordinators at Hartman High School. A total of five interview participants and sixteen survey participants took part in the study. Prior to the collection of any data, informed consent was acquired from all participants.

Semi-structured interviews consisting of 16 questions were conducted with each of the five interview participants on an individual basis. The purpose of these interviews was to focus on the factors that impact low test scores on the MEA for Science at Hartman High School. Prior to beginning each interview, participants were provided with a brief summary of the purpose of the study in addition to the information they received when invited to participate in the study. Interviews were conducted off-site, either at the local public library or at a restaurant, with one participant's contributions being via written response per the participant's request. IRB approval was obtained prior to data collection.

The survey was administered via Google Forms and was structured in a Likert-scale format. For this component, sixteen participants responded to 15 questions related to the MEA for Science and contributing factors related to students' performance on the assessment. To reduce bias and potential identification of individual participants' responses, participants chose a unique numeric identifier during the pre-survey screening. This numeric identifier allowed the researcher to align responses to the pre-survey screening, ensuring that demographic data had been collected and that consent had been given without risking a breach of anonymity of participants and their responses.

Field notes were recording during and after each interview, along with a transcription of each interview. Documents were also gathered from the MDOE website, including those that are publicly available related to students' performance on the MEA for Science at Hartman High School. These documents provided further supporting evidence for the themes developed. All participants remained in the study until its completion.

Sub-question 1

Sub-question one for this study was, "How would teachers and administrators in an interview solve the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine?" Interviews were conducted with one administrator and five science teachers at Hartman High School in order to find themes related to the low test scores on the MEA for Science at Hartman High School. Table 2 shows the codes, properties, and examples of participants' words relevant to each code, while Table 3 shows the themes and frequency codes among interview data.

Table 2

Open Codes, Properties, and Examples of Participants' Words from Interview Data

Open Code	Properties	Examples of Participants' Words
Data	Common assessment	Looking at our assessment practices and common assessment and comparing results to what we look at collectively internally and look at results that we're getting from external elements—I think that's sort of the next natural step.
	Data analysis	Systems should be in place within the district to mine individual results, aggregating and

		generalize performance to identify strengths and weaknesses.
	Item analysis	I think it would be interesting to look at the distribution of questions on the MEA and how we're covering those things in curriculum and in what years we're doing it.
	Value of Data	I think the biggest contributing factor is the teacher's perspective on standardized assessments and not seeing the relevance of the testing process.
Instruction	Lesson planning	You're shifting and adjusting your instruction every day based on what you know they learned yesterday and what they still need to continue to work independently or collectively or whatever it may be.
	Released Items	Grade 11 teachers have been provided with the bank of released questions and scoring guides published by the Maine DOE.
	Hands-on learning	All teachers are doing labs and having students generate hypotheses and analyze data.
	Instructional strategies	When teaching lower-level students for a long time, we used to joke about how we had to trick our students into learning. And I think really effective teachers are good at that, and they make learning fun enough that they can get kids involved.
	Instructional Resources	We have our computers as electronic tools, and we use them for graphing and spreadsheets and putting together reports, research on topics that we need to supplement what's in the textbook because what's in the textbook—a lot of stuff has happened since the textbook was written
	Literacy	I think the other thing that is really critical to a student's success or score ultimately on this assessment is not only their scientific understandings and vocabulary knowledge or scientific processes, but literacy skills are critical.
Risk Factors	Home support	It also depends on basic supports and attitudes at home.
	Socioeconomic status	Show me the zip code, and we'll show you how they're going to do on the exam.
	Special education	We have a large number of special education students, which is part of the population I think I've always been really interested in looking at how our test scores compare.

	Disability, homelessness, substance abuse	We have a high percentage of students with performance risk factors such as disability, poverty, homelessness, substance abuse, etc.
Standards	Lack of familiarity	I'm not as clear as I should be on this.
	Scientific processes	A majority of it how one reads and thinks or write from a scientific viewpoint and analysis of data.
	Standards shift	I think part of the problem, too, is that whenever you make a change at the state level in requirements, it's going to take years for schools to implement those changes.
	Alignment	Whenever you make a change at the state level to the requirements, it's going take years for schools to implement these changes, and there's been changes so often at the state level, that it's like schools can't keep up.
	Standards overlap	There is a significant overlap between many NGSS and MLR performance indicators.
	Common pacing/sequence/ resources	Even if we're not all doing evolution at the same time, there's still that assessment and those common assignments available in both places that teachers can use.
	Vertical alignment	The scientific processes that is taught rom ninth grade through the senior year in students' courses of study have a natural connection to the assessment.

Table 3

Themes and Frequency Codes Across Interview Data

Themes	**Code Word**	**Occurrences Across Data**
Data Analysis	Analysis	5
	Assessments	2
	Item Analysis	9
	Teacher Buy-In	9
Instruction	Assessments/Common Assessments	23
	Common Assignments	4
	Discussion/Dialogue/Conversation	6
	Hands-on Learning	6
	Literacy	13
	Practice	5
	Progression/Continuum of Learning	7
	Released Questions/Test Banks	7
	Remediation	2
	Resources	19

	Schedule	4
	Scientific Method	16
	Staffing	4
	Teacher Effectiveness	15
	Time	9
Risk Factors	Lower-level Kids/Students	5
	Low Socioeconomic Status	6
	Special Education	2
	Motivation	5
Standards	Learning Goals/Results	7
	Scientific Practices/Processes	9
	Alignment/Rigor	3
	Common Core/NGSS	7
	Changes	6
	Curriculum	7

Theme 1: Data analysis. The first of the four themes that became evident from the interviews was data analysis in the science department at Hartman High School. Several participants reported that there is limited time allotted for data analysis among teachers, but that a desire to conduct such analysis exists. Participants also reported a recognition of item analysis and the benefit such practice could have on students' performance on the MEA for Science.

Ms. Marion expressed an interested at looking at the "distribution of questions on the MEA" and "how we are covering those things in the curriculum and in what years that we're doing it" (Marion, personal communication, 2019). Likewise, Mrs. Harrison also feels that an item analysis should be conducted, despite its lack of availability from the MDOE in recent years; she adds, however, that while a desire to perform such data analysis exists, albeit in a cumbersome manner, time has not been allotted to teachers to work professionally on such items in favor of other initiatives put forth by state law (Harrison, personal communication, 2019).

Common among participants' thoughts towards data analysis was the idea of common assessments. Participants agreed that common assessments are necessary, and several interview participants indicated that common assessments are being used in their courses. This commonality among teachers has led to item analysis within the specific courses, as explained by Ms. Marion, who stated

> we're getting things together in terms of our learning goals and our common assessments having better curriculum outlines that are going to be standardized by each grade level, and looking at common assessments to make sure that our assessments that are being used among everybody reflect the things that we really want them to know according to the learning results and what we know they're going to be test on (Marion, personal communication, 2019).

Theme 2: Instruction. The second theme, with an abundance of data to support it, is instructional practices at Hartman High School. While teachers within the science department are well-trained, innovative, and remarkably passionate about teaching science to their students, the interview data clearly indicated that teachers feel more attention should be focused on instructional practices, especially with regard to how these practices can positively impact students' scores on the MEA for Science at Hartman High School.

Interview participants indicated that students' literacy is a concern with regard to instructional practices and assessments. While students may understand the content about which they are asked, many lack the literacy skills necessary to understand the question itself (Harrison, personal communication, 2019). Of interesting note, the students who take the MEA for Science also take the SAT just prior to the science assessment. Documentation from the SAT shows that students perform significantly better on history and English passages on the verbal portion of this test, but they do not fare well on the science passages (College Board, 2019b). By this measure, students lack the technical reading skills that are necessary to perform well on science-related tests, including the MEA for Science.

Instructional practices are also impacted by staffing, which was heavily represented in the interview data. Participants indicated that there has been a lack of consistency in the science department staff, often producing an unsettled feeling for the staff. The lack of consistency in the staff, along with the reduction in staff over time due to a decrease in student population, has put a greater responsibility on those that remain in the department, which has an impact on instructional practices. Likewise, because new teachers are frequently joining the staff, it is difficult to maintain forward momentum with consistency in test preparation instructional practices, in both content and test-taking strategies.

Several interview participants spoke to the use of released questions to help students prepare for the MEA. While participants were aware of the existence of these questions, there was no concrete evidence as to the intentional or consistent use of these questions just prior to the assessment, nor in the years leading up to the assessment. There was, however, an understanding that students at the junior level would review these questions with their teachers prior to the assessment (Gilmer, personal communication, 2019), but the fidelity with which this happened cannot be accurately reported. Instead of using these released questions simply to prepare for the MEA, teachers were interested in having a communal test bank from which to choose questions for assessments, including common assessments throughout the academic year among all teachers of a particular course.

Perhaps the most frequently mentioned instructional component in the interviews was the use of hands-on learning in science classes at Hartman High School. All five interview participants spoke about the use of laboratory-based activities in which students develop and practice skills associated with scientific processes, inquiry, and analysis. This commonality amongst participants indicates that the practice is, in fact, common among science teachers at

Hartman High School. The skills assessed with such practices are aligned with the standards assessed with the MEA for Science, seemingly providing a strong foundation upon which students further their scientific understandings.

Theme 3: Risk factors. Two overarching characteristics were mentioned with regard to risk factors associated with the MEA for Science. The socioeconomic status of students, as well as the special education services that students receive, were both cited as risk factors by interview participants. For example, one participant indicated that the zip code of a student is a pre-determining factor of the student's success. That is, the student will likely only reach a certain level of success based on the neighborhood from which he or she hails, regardless of the student's true potential. While there are certainly exceptions to this, many participants referred to students from socioeconomically disadvantaged backgrounds as having a lack of home support of their education (Gilmer, personal communication, 2019) or a lack of buy-in for the assessment (Harrison, personal communication, 2019; Marion, personal communication, 2019), which may have its roots at home due to a lack of investment in education by their parents or guardians.

Theme 4: Standards. Based on the interviews, participants are well aware of the fact that scientific processes are a critical component of the standards assessed on the MEA. In fact, each participant spoke to this, citing the use of hands-on learning, interpretation of data, a continuum of development of these skills throughout students' 9-12 experience, and the way that these skills are incorporated into assessments throughout the academic year. What was not mentioned, however, is a strong certainty of what other standards are incorporated on the assessment.

Some participants described their familiarity by discussing the general topics that are included on the assessment. For instance, one participant stated, “I’ve never seen an MEA exam. (I know) there’s an earth science portion. There’s the biology portion. There’s a chemistry portion. There’s a physics portion. There’s a science (principles) portion,” (Lewis, personal communication, 2019). Another spoke solely of the scientific practices that are included, referencing how the course that they teach would be incorporated—“But then they’re given a whole reading about the peppered moths in England, and they’re given data, and they have to make a graph, and they have to try to explain what the graph is showing and why” (Marion, personal communication, 2019). Only one participant, Mr. Gilmer, specifically outlined which NGSS standards were included on the assessment and at which grade levels they are taught (personal communication, 2019). Notably, participants were aware that of standards associated with each course and their responsibility to ensure that students are learning this material while enrolled in the course; likewise, most participants seem to have a strong understanding of how their subject is incorporated in questions regarding scientific skills and practices on the assessment.

Sub-question 2

Sub-question two for this study was, “How would science educators in a survey solve the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine?” A total of sixteen participants completed a 15-question Likert scale survey in order to collect quantitative data related to students’ scores on the MEA for Science at Hartman High School. Table 4 shows the frequency and average of responses to each survey question.

Table 4

Frequency and Average of Survey Responses

Question	1	2	3	4	5	Average
1. Before teaching a new concept, the associated Next Generation Science Standards (NGSS) are intentionally built into instruction.	1	0	4	8	3	3.75
2. Classroom instruction is delivered with the same rigor specified in the NGSS standards.	1	2	4	7	2	3.44
3. All students are assessed in the same manner across grade levels.	0	2	7	6	1	3.38
4. Time is provided during contract hours for collaboration between teachers of the *same* courses.	0	4	4	6	2	3.38
5. Data collected from both assessments are used to plan future instruction.	0	2	4	9	1	3.56
6. Content taught in each course is clearly communicated throughout the department.	0	3	3	4	6	3.81
7. Professional development focuses on new instructional strategies has been offered to science teachers.	1	5	7	2	1	2.81
8. Teachers regularly participate in meaningful data analysis conversations.	1	7	6	2	0	2.56
9. Students are excited and ready to learn in my class each day.	0	2	7	7	0	3.31
10. The science department at Hartman High School is made of teachers that are highly-qualified educators.	0	0	0	4	12	4.75
11. On-going support is provided to new teachers.	0	2	6	5	3	3.56
12. Professional development that serves the instructional needs of teachers is provided on a regular basis.	1	9	5	1	0	2.38
13. Teachers in the science department at Hartman High School have high attendance rates.	0	0	2	8	6	4.25
14. When teachers are absent from school, high-quality substitute teachers are hired.	0	6	9	1	0	2.69
15. Class sizes in the science department at Hartman High School are conducive to student learning.	0	1	2	9	4	4.00

Theme 1: Data Analysis. From the survey results, it is evident that data analysis remains a theme among participants. The results indicate that participants do not feel that enough time is allocated to collaborative discussions related to meaningful data analysis at Hartman High School. This data analysis includes the use of common assessments and the data collected from the assessments to plan future instruction. Interestingly, however, participants indicated that data from assessments is used to plan future instruction. This discrepancy in data may be due to the use of *assessment* to refer to both the MEA and common assessments that occur in the classroom throughout the academic year.

Theme 2: Instruction. Survey participants indicated that professional development specifically related to instructional practices does not occur on a regular basis. In fact, this question scored an average of 2.38 out of 5—the lowest score of any question on the survey. Because there is a lack of professional development to support instructional practices, it’s reasonable to consider that a focus on more frequent professional development on the topic would strengthen instruction within the classroom, which would likely translate to improved scores on the MEA for Science at Hartman High School. Additionally, participants’ responses show that, in the event of a teacher absence, high-quality substitute teachers do not seem to be hired to fill the vacancies.

Theme 3: Risk Factors. While questions specifically targeting risk factors of students were not asked in the survey, participants indicated with a score of 3.31 out of 5 that students are excited and ready to learn in class each day. While student excitement does not necessarily directly correlate to risk factors, the level of student excitement often translates to engagement, and students’ level of engagement in their learning can certainly impact their performance on assessments and classification within various risk factor categories. Further, a student’s readiness to learn may not be related only to their mental state of preparedness, but also to the resources to which they have access for their learning. The 3.31 out of 5 score is not necessarily a *low* score, it is just about the midway point of the middle score; this is not a strongly-positive score, indicating that, when coupled with the data from the two other methods, risk factors remain a theme.

Theme 4: Standards. Survey participants responded, on average, in a positive way to the alignment of standards in courses to the NGSS standards, as well as the rigor level at which concepts and skills are taught in their classes. Much like the question about students' excitement and preparedness referenced in Theme #3, these scores are positive, but not overwhelmingly so. That is, the average scores for standards incorporation into instruction and rigor are 3.75 and 3.44, respectively, leaving further room for improvement.

Sub-question 3

Sub-question three for this study was, "How does a review of documents inform the problem of low test scores on the MEA for Science at a suburban public high school in southeastern Maine?" A review of data was conducted using results from the MDOE regarding the performance of students at Hartman High School on the MEA for Science. The results of students' performance on the MEA for Science for the three most recent years available are presented in Table 5.

Table 5

Percent of Students Failing to Meet State Expectations on the MEA for Science at Hartman High School

Year	Overall	Male	Female	Socioeconomically Disadvantaged	Special Services
2015-2016	51.1	48.5	54.1	70.6	Suppressed
2016-2017	40.9	38.7	43.0	46.7	Suppressed
2017-2018	51.9	62.5	36.4	63.9	82.8
Mean	48.0	49.9	44.5	60.4	n/a

According to the MDOE, data that is suppressed has been removed from reports in an effort to protect students' privacy (MDOE, 2019). Such suppression was enacted in reporting assessments results for African American, Hispanic, and Asian student groups, as well as for those of two or more races and English language learners.

Theme 1: Data analysis. Understanding the significance of data analysis on students' scores on the MEA for Science from the score reports alone is challenging, as data analysis is not a reportable category. However, the availability of the reports themselves speak to the need for ample data analysis and the information that can be gleaned from the data.

Of particular note, the reports that are publicly available from the MDOE do not provide information with regard to reporting categories, item analysis, or other pieces of information that could help to inform instruction or provide starting points for educators in terms of how to help students perform better on the assessment. Likewise, from the password-protected database, the only feasible way to acquire an item analysis is to assess each student's individual score report and then correlate responses to questions. The challenge with this, however, is not only the resource-consuming aspect of the task, but also the lack of availability to all teachers in the school. Because these reports are not readily available to the appropriate stakeholders within the school, the forward progress that is possible as a result of data analysis is compromised.

Theme 2: Instruction. Similar to the data analysis theme, instructional decisions are also limited by what information is and is not available to teachers. What is evident, however, is the need for changes in instruction to be made, as nearly 50% of all students tested during the three years studied have failed to meet the state's expectations for performance on the assessment. A change in instructional practices may certainly lead to improved scores on the assessment in future years, but these changes must be guided by information that is more informative in nature. The same information that would improve the data analysis component would also serve the instructional theme.

Theme 3: Risk factors. It is apparent from the review of documents that there is a large population of students that are not succeeding at meeting the state's standards for performance on the MEA for Science at Hartman High School. Most notable among this group is the number of socioeconomically disadvantaged students. On average, 60.4% of these students fail to meet the expectations set forth by the state, with the highest percentage occurring during the 2015-2016 academic year, in which 70.6% of this population did not meet the expectations. Likewise, 82.8% of the students receiving special education services failed to meet the state's expectations for student performance on the assessment. It should be noted that data from two of the three years included in this study was suppressed, the data available remains alarming.

Theme 4: Standards. Because the MEA for Science is designed to measure students' abilities to demonstrate mastery of standards, the document review indicates in its entirety that students are not achieving the standards. Unfortunately, because of limitations of the reporting database, a summative item analysis is not available for this data, making the identification of specific strands as strengths or weaknesses is nearly impossible. Aside from the evidence that students' are not meeting the state's expectations for mastery of the science standards, the lack of information available regarding standard-specific performance brings this to the forefront as a part of this theme.

Discussion

The discussion portion of this chapter provides a discussion of the findings of the research in relation to the literature review in Chapter Two. The triangulation of data from each of the methods provides support for each of the themes that emerged, which are further explained in the forthcoming sections.

Theme 1: Data Analysis

When test scores are made available following the administration of an assessment, the analysis of data is a critical component to improving instruction (Lewis, Madison-Harris, Muoneke, & Times, 2010). When instructional practices are revised in response to students' academic needs as indicated by assessment scores, students' learning is likely to improve, which results not only in improved test scores, but in a better education for the students (Lewis, et al., 2010). When data analysis takes place among an educational team, the identification of students who are very nearly meeting expectations becomes easier (National Research Council, 2014), in-class assessment techniques tend to be better aligned with the high-stakes assessment (National Research Council, 2014), general classroom instruction is more aligned with the rigor of the assessed standards (National Association of Elementary School Principals, 2006), and vertical alignment among courses improves so that students have a comprehensive science education experience.

The participants in both the interview and survey components of this study indicated that data analysis is recognized as an important practice within the science department at Hartman High School, but several obstacles make this task difficult. In the survey, participants reported that adequate time is not provided to teachers to analyze the data as a group. Several interview participants confirmed this, indicating that efforts are made to review data, but that this is not a historically prevalent practice within the department.

Adding to the challenges of analyzing data is the lack of reports that are readily available to teachers and administrators via the MDOE. The reports available publicly provide a breakdown of information with regard to gender, race, socioeconomic status, and special services, but little information exists beyond this. A considerably greater amount of information is available to administrators via a password-protected site; however, these data remain unrefined in some ways. Most notably, the only way to acquire an item analysis, according to one of the survey participants with access to this information, is to review each individual student's report, compiling a question-by-question tally for the entire testing pool. Once this is done, however, the availability of analyzing these data with regard to standard alignment, rigor level, or any other variety of indicators is quite limited.

Further, because the test is administered during a student's 11th grade year, but scores are not available until the fall of their 12th grade year, several opportunities to revise practices in a way that benefits the tested students are missed. Also, the 11th grade teachers are more likely to be interested in this information since the students were assessed during the academic year in which they were enrolled in the teachers' classes; the timing of the assessment and score reporting does not facilitate meaningful conversations among the 9th and 10th grade teachers regarding the skills and standards taught while students are enrolled in their courses. The limited amount of data analysis among teachers also lends itself to a lack of parental involvement and focus on helping those students who are nearing a score that meets the state's expectations.

Theme 2: Instruction

The second theme that emerged from the three methods was instruction and instructional practices at Hartman High School. Prior research indicates that teacher effectiveness is a critical component in the learning that occurs in the classroom (Tucker & Strong, 2005). Such effectiveness is often a result of college education (Stronge, 2018), professional development (American University, 2019), school and class sizes (Higgins, 2014), new teacher support (American University, 2019), teacher attendance patterns (Okeke et al., 2015), and the quality of substitute teachers in the absence of the regular classroom teacher (Westrick, et al., 2015; Davies, 2019).

Interview participants agreed that teacher effectiveness definitely has an impact on students' learning. In fact, participants mentioned several characteristics of being an effective teacher—building relationships with students, valuing assessment scores and information to be gathered from data, passion for teaching science—as a part of their discussion; this is a strong indicator that effectiveness is something considered to be a significantly contributing factor to student success among interview participants.

Interview participants indicated that professional development needs remain strong with regard to topics that will help them to improve their instructional practices in the classroom. Providing opportunities for teachers to learn new instructional strategies or deepen their understanding of and efficacy with current instructional strategies can strengthen the classroom instruction which, of course, can have positive effects on students' performance on the MEA for Science. For example, several teachers in the science department at Hartman High School use some variation of the Claim-Evidence-Reasoning approach to develop scientific conclusions based on data collected during various laboratory activities. Providing additional instruction for teachers on how to deepen their own comfort with and understanding of this approach and would be beneficial for students, whose goal would be to students become better with expressing their thoughts and ideas using this structure. However, new strategies, such as learning how to teach literacy skills in the science classroom is critical to ensure that students are learning new content in a variety of ways, while strengthening the skill set necessary to perform well on the MEA for Science.

Theme 3: Risk Factors

In all three data collection procedures, it was quite apparent that students' risk factors have a great effect on their performance on the MEA for Science. Interview participants referenced these risk factors in a variety of ways, such as a lack of motivation, perhaps due to testing burn out, and socioeconomic status. Survey results led to the identification of motivation factors, as well, while a review of documents clearly indicates that students with a low socioeconomic status, receive special education services, or are a part of one of the minority populations at Hartman High School also perform below the state's expectations and lower than their peers who do not identify with any of these risk factors.

While some of these factors are beyond the control of the school and its faculty and staff, such as students' socioeconomic status, race, healthy lifestyle, and so on, other factors, especially student motivation, are. Student motivation, or a lack thereof, can apply to all student groups regardless of other identifying qualities. In the interview, several participants referenced a lack of accountability as one of the reasons that students are not motivated to perform well; the same may also be true for teachers, as the scores from the MEA for Science are not tied to evaluation or continuing contract status. As Wolf and Smith (1995) noted, the less at risk in these assessments, the less engagement students, and perhaps teachers, as well, have in the process. Though the school cannot control most of the other factors that can be considered to make students at risk for low performance, this does not free the school and its teachers from identifying students with these risk factors and utilizing strategies for intervention to reduce the impact that these risk factors have on students' performance, both in the classroom and on high-stakes assessments such as the MEA for Science.

Theme 4: Standards

Overall, teachers and administrators seem to be aware of the standards that are assessed by the MEA for Science, but the depth and breadth of their familiarity with these standards are wide-spread. Some participants indicated that they are very familiar with the standards and what it is that is expected for students to demonstrate mastery of these standards, while others said that they are aware that such standards exist, but they are not as familiar with them as they feel they should be. Being familiar with the standards is the first step in ensuring that classroom instruction and assessments are aligned to both the content and rigor of the standards.

This information was also supported by the survey data. Several participants indicated that standards are intentionally incorporated into instructional plans, but the results do not indicate that this is an overwhelmingly high confidence with this fact. Ensuring that both the content and rigor of the standards are evident in instructional and assessment plans is imperative to equipping students with the skills and strategies for answering questions about the topics and applying scientific problem-solving skills to more complex questions. Based on the information from both the survey and interviews, however, it is reasonable to think that this is being incorporated in certain courses with more intentionality than in other courses.

Finally, the review of documents provides clear evidence that students are not, in fact, meeting the state's standards when it comes to demonstrating a mastery of the specified NGSS standards. While the reports available do not specify in which areas exactly that students are failing to perform well, their overall performance indicates that there is still much work to be done if the school wishes to prepare students to meet the state's standards of performance. Becoming familiar with the standards that are assessed with the test and formulating strategies for teaching and assessing these standards are key to moving students forward.

Summary

For this study, data were collected from participants, including science teachers, special educators, administrators, and a testing coordinator at Hartman High School. Qualitative data from interviews provided substantial text from which themes were identified; these themes were further supported by the quantitative data from the Likert-scale survey data and review of documents, including score reports available from the MDOE. The findings of this study revealed four themes related to factors that impact students' performance on the MEA for Science at Hartman High School. After Chapter Four, Chapter Five will provide a proposed solution, along with necessary resources, funds, roles and responsibilities, implications, and the evaluation of the solution.

CHAPTER FIVE: CONCLUSION

Overview

This multimethod research study sought to identify the factors that impact students' performance on the Maine Educational Assessment for Science at Hartman High School, a suburban public school in southeastern Maine. In this chapter, the problem is re-explained and a solution to the problem, along with the necessary resources, funding, roles and responsibilities, timeline for implementation, and evaluation of the solution are presented. Associated documents in the Appendix, referenced throughout the chapter, also provide more concise, easy-to-follow lists of strategies and sequences, as well as graphics, to help implement the recommended solution.

Restatement of the Problem

The problem addressed with this study was the low test scores on the MEA for Science at Hartman High School. The past three years of testing data available indicate that, on average, just under 50% of students at Hartman High School are meeting the state's expectations for performance on the assessment. A multimethod approach was used for data collection, including a series of interviews, a Likert-scale survey, and a review of documents. The four themes emerging from the data were data analysis, instruction, risk factors, and standards, all of which aligned with the theoretical and empirical evidence provided by prior research studies.

Proposed Solution

After careful analysis of the data collected in this study, the two most effective solutions to propose are that of establishing professional learning communities (PLCs) at Hartman High School, as well as providing targeted professional development to teachers to enhance professional practice. Not only will both of these solutions propel the teaching and learning at

Hartman High School to the next level, but teachers will become empowered to spearhead positive changes on their own as they continue to develop skills and knowledge related to standards-aligned instruction and data analysis practices.

Professional Learning Communities

Professional learning communities afford the flexibility and fluidity to tackle several challenges through a single vehicle. However, the implementation of PLCs within a school must be executed in an effective, yet efficient manner, so as to ensure the fidelity with which they operate and the timeliness of the work and reflection that takes place within the group; failure to do so may not only hinder the current intentions of PLC work, but can also prevent the buy-in from teachers for such a strategy. For this reason, it is imperative that school leaders are well-versed in the various functions and purposes of a PLC, as well as how to create buy-in from teachers and how to troubleshoot unexpected obstacles to the process.

Implementing any new initiative is sure to be met with some skepticism, especially in the case of Hartman High School where interview participants indicated that a flurry of district-based initiatives have come and gone in recent years, many without reaching fruition. As such, teachers may be reluctant to begin the on-boarding process to a new strategy for improving students' success. However, Jessie (2012) recommends the following

Jessie (2012) recommends the following steps for creating buy-in among faculty and staff:

1. Allow staff to see success with the process from other schools with similar demographics.
2. Share a personal vision that is genuine and contagious regarding the benefit of PLCs.
3. Find the connection between short-term goals and progress and the school's mission statement and vision.

4. Allow teachers that share the same vision to take owner- and authorship of the work done in PLCs.

Allowing staff to see the success of other schools using PLCs can provide the foundation for conversations as to how PLCs could operate at Hartman High School. Organizing a series of site visits to other schools with similar demographics that are finding success with PLCs will serve as the catalyst for buy-in by providing a real-life example of the function and success of the strategy. Once teachers recognize the benefit of PLCs, attending a PLC institute can provide the research-based information necessary to start the facilitation of PLCs while keeping all participants focused on the same goal with the same underlying understandings. For those that are unable to attend site visits or the institutes, videos and book studies can offer the foundation upon which successful PLCs can be built.

Once teachers and school leaders are on board with the use of PLCs to improve students' learning, seeking professional development from an outside source, such as that of Solution Tree, the organization founded by Richard DuFour, the developer of the PLC, is recommended. This professional development will allow participants to

> gain an understanding of the PLC at Work process, including the three big ideas of a PLC: focus on student learning, focus on collaboration, and focus on results—and learn how to customize it to meet your needs (Solution Tree, 2020b, para 4).

This professional development can be provided onsite, or participants may travel to one of many workshops held in various locations around the country. Either way, attending one of these formats means that all teachers in the science department at Hartman High School will be able to not only hear the same information and will be provided with time to develop a collaborative setting amongst themselves, establishing norms and expectations from participants and the work to be done. Likewise, guidance will be provided by the presenter to ensure that participants have a precise understanding of what it is that each feature of a PLC is meant to accomplish and how to best go about the process.

Once buy-in is established among teachers and professional development has been conducted, the real work of PLCs begins. The PLCs should be established by course or content; that is, biology teachers should work with other biology teachers, for example. To make this a reality, it is imperative that teachers of the same course have the same planning period. If this would cause too many conflicts with regard to student schedules, time must be protected within teachers' contract hours to allow for PLCs to meet.

To best serve the purpose of a PLC with regard to instructional matters, especially those related to the MEA for Science, the teaching and learning cycle (TLC) should be implemented, as shown in Figure 1. The TLC allows teachers to collaboratively plan instruction—both resources and strategies—that is standards-aligned to both the content and rigor. Within this lesson planning, scaffolding and differentiation strategies can be developed to best serve all students enrolled in the course. After the commonly-planned lessons have been taught, a common formative assessment summative provides information as to students' mastery of the content and understandings. Those students that have demonstrated mastery will proceed to the next lesson, which may be an enrichment opportunity, while those that have not yet mastered the material will be provided with remediation. At this point, teachers may opt to divide students in a way that allows for homogenous grouping as students continue to develop the skills and understandings necessary for mastery. Following this re-teaching, students will again be assessed for mastery, continuing the cycle until all students have performed at an acceptable level. The PLC continues to meet regularly during this time to discuss student performance and to determine the next step in the process. Following this, the PLC reflects on the teaching and learning that took place and begins to plan for the next lessons. A key part to this next step of planning, however is to incorporate spiral review so that students do not lose the recently-acquired skills and understandings.

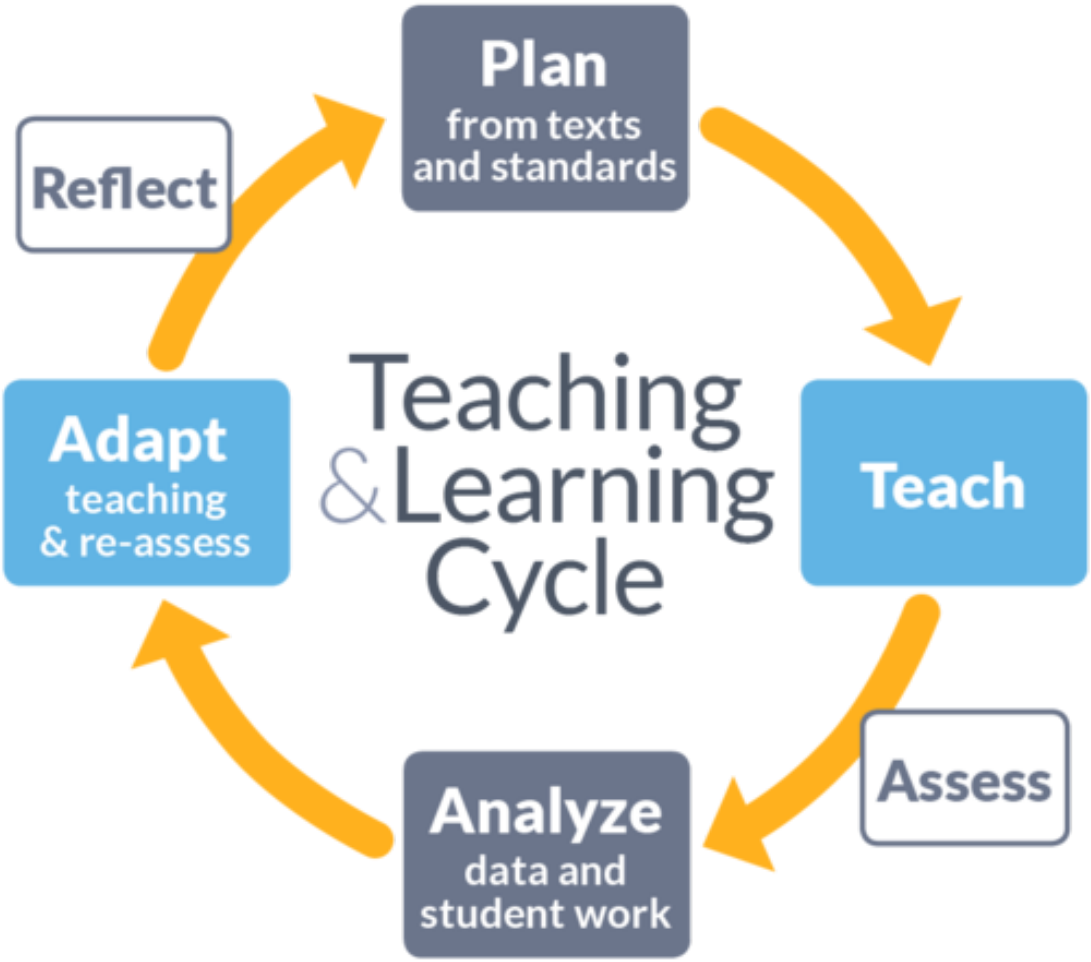

Figure 1. Teaching and Learning Cycle. This figure illustrates the cyclic process of collaborative work within a professional learning community. (Achievement Network, 2020).

It is notable that some of the most significant feedback from participants in this study was the need for data analysis practices among teachers. Using the TLC within a PLC allows for this data analysis to take place in a way that is focused on improving students' learning and understandings of course-related content. The PLC provides a common framework among teachers of various courses within the science department, which can also lead to a discussion of vertical alignment and the sequence in which the NGSS standards are taught at Hartman High School, both of which are critical conversations when attempting to improve students' performance, especially on a standardized assessment such as the MEA for Science.

Professional Development

Professional development, as a whole, should be an on-going and job-embedded experience. Job-embedded learning "means that professional development is a continuous thread that can be found throughout the culture of the school" (Education World, 2012, para. 2). Creating opportunities for such learning is critical to ensuring that the strategies learned are implemented immediately and with fidelity. However, to ensure that this happens, it is important that the learning be relevant to the participants, feedback is encouraged, and facilitates the transfer of newly-acquired skills into practice (Education World, 2012). Choosing professional development topics that are not only interesting to the participants, but also timely, will help to ensure that the time and resources used for such learning are well spent.

Mentioned in the interview portions of this study, literacy skills are lacking among students, which may be impacting their performance on the MEA for Science. Specifically, reading in the content area is an area of concern. Seeking professional development for science teachers to help them learn how to build the literacy skills of their students within the classroom, rather than that of a language arts classroom, is critical. This professional development should not only provide participants with the rationale of such instruction, but also research-based, tried-and-true strategies that can be easily and quickly implemented in the daily instructional plans of teachers. This professional development may come from an in-house source, such as a reading specialist within the school or the district, but it may also come from an outside source. Materials, such as that of Newsela, an online-platform for content-sorted non-fiction passages, would be useful to teachers, especially because Newsela allows for this content to be scaffolded based on students' current reading level; this platform is a feasible option, as all students have access to their own computer, both at school and at home, making the utilization of such a resource relatively low-effort for the teacher.

When planning lessons, whether involving literacy skills or not, teachers must ensure that the lesson is aligned to the appropriate standard with regard to both the content and rigor. Not only should this alignment be obvious in the lesson plans, but also in the daily instruction taking place. Professional development focusing on how to unpack, or breakdown, the NGSS standards into the various components regarding skills, understandings, and rigor, should be provided to teachers. This professional development should include information regarding the scaffolding and differentiation of instruction in a way that builds students' understandings to a mastery level, aligning itself with the goals of the PLC. The professional development should be executed as a workshop in which teachers work collaboratively in their PLC groups to develop a standards-aligned lesson or unit, complete with associated materials, so that a product exists as an exemplar of the expectations and as a usable resource for instruction. The professional development could be provided by an instructional specialist within the district, such as the Assistant Superintendent whose role incorporates curriculum and instruction, or another source, such as specialists from the state or an outside organization. For cost purposes, however, as well as a familiarity with the work to be done, an in-district option is best.

As part of the professional development focusing on standards-based instructional planning, the training should also include strategies for developing standards-aligned assessments. These assessments should, much like the instructional materials, be aligned to both the content and rigor of the standard. These assessments should include a variety of rigor levels, ultimately requiring students to demonstrate mastery at the expected level; however, the use of lower-rigor questions provides information to the teachers as to where, exactly, on the continuum of learning their students are currently performing. Like the lesson planning component of this professional development, a focus on the assessments and their alignment to the standards also fulfills the purpose of the PLC.

Resources Needed

Before any of the professional development can be offered in a meaningful way, it is imperative that common planning is secured for content-related teachers. This may require the reorganization of the current scheduling rubric or the use of late-start times to facilitate the PLC process. Because scheduling, at least in one course, is based upon the availability of a part-time teacher, it is recommended that at least one full-time teacher be hired instead of a part-time position. This change in teaching staff would require the procurement of additional funds and the approval of the school board. If late start times are to be utilized, this would require that other meetings often scheduled during this time, including faculty and department meetings, may need to happen at another time or in another format, such as after school or via electronic means, if PLCs are to meet on a regular basis, preferably weekly.

It is also advisable that teachers have similar schedules beyond that of planning to facilitate shared remediation and enrichment opportunities for students. Ideally, at least two teachers would be teaching the same course, though perhaps at different levels, during the same period. When such a schedule is in place, teachers are able to switch students between classrooms for targeted, homogenous groupings to help students make greater academic gains.

In situations where only two teachers are assigned to a course, such as those at the 11th and 12th grade levels, this task may become more challenging, which, again, leads to the suggestion of a full-time teacher in place of the part-time teacher currently in place.

Time is also needed to host professional development for teachers and administrators. This includes time in August prior to the school year starting, as well as on-going professional development throughout the school year. Because time is a limited resource, especially once the school year begins, some of this on-going learning can take place during the late start days or during PLC time, as the benefit of this professional development is directly applicable to the PLC and student learning. It should be viewed, however, as an extension to the PLC work rather than in place of the typical PLC work.

Additionally, access to online tools specifically related to the work of the PLC and increased literacy skills among students is needed. Solution Tree offers a wide variety of online tools for the facilitation and implementation of PLCs, requiring computer and internet access, as well as password-protect access to information on the website. The same is true for online literacy tools, such as the recommended Newsela. To adequately implement the use of Newsela as a literacy tool, however, information about individual students' reading levels is required. If this information is not readily available from the English or language arts teachers, teachers may still use the platform, as Newsela automatically adjusts the level of the materials provided to students based on their prior performance with assigned resources (The Pennsylvania State University, 2020). This feature allows for students to grow in their literacy-related abilities throughout the academic year while being appropriately challenged with science-related content (Newsela, 2020).

Funds Needed

Converting the current half-time teaching position to a full-time position would require a financial investment of approximately $20,000 depending on the teacher's prior experience. Additionally, because this position would now be full-time, a benefits package, costing up to approximately $7,800 for healthcare for a single individual and (MSAD 75; 2020); this cost would increase if the teacher carries a spouse of children on the policy.

Solution Tree, the company started by Richard DuFour, the pioneer of PLCs, offers a variety of two-day trainings that focus on different aspects of PLCs. To attend one of these workshops costs $689 per person (Solution Tree, 2020a). For the entire science department at Hartman High School, plus the special educator that works most exclusively with the science department, as well as the four school administrators who would provide on-going support for the department, the total cost would be $10,335, plus the cost of travel and accommodations for each participant. This cost includes the two-day workshop, along with opportunities to work within collaborative teams to begin the early work of establishing a PLC. Participants will also leave the seminar with access to a variety of tools to help continue the work of their PLC upon returning to their school.

Because this cost is extensive, it is reasonable to consider sending only key individuals to the two-day session with the charge of bringing information back to the rest of the department and administrators. If this were to be the case, the cost could be reduced to as little as $3,445, plus travel and accommodations. Access to the same materials would be provided for use within the groups, but additional time would be needed to work through the primary steps of establishing a PLC among content-related teachers, as only a limited number of teachers would be attending the training.

To use Newsela as a tool to support literacy within the science department, a school-wide license would necessary. While Newsela offers a free option that provides access to basic features, more advanced resources are available with the purchase of licenses. While Newsela does not publish its rates on its website, the company shared on social media that licenses are available for $18 per student, $2,000 per grade level, or $6,000 for the school (Fordham Institute, 2016). Based on the figures, a school-wide license would be the best option for Hartman High School's science department. This cost, because it is, in fact, a school-wide license, could be distributed among all content areas if administration chooses to utilize Newsela across the curriculum.

Overall, a total cost of at least $37,245 plus travel and accommodations is required to implement these recommendations. This cost is more than the science department's budget at Hartman High School, making these expenses rather significant and requiring the district to provide additional funding for the conversion of the part-time position to a full-time position. Still, the remaining balance is more than half of the science department's budget; if, however, substantial changes are to be made regarding students' learning and their performance on assessments such as the MEA for Science, significant steps must be taken to ensure that all parties involved are operating with the same intention and background understandings moving forward.

Roles and Responsibilities

In order to improve students' scores on the MEA for Science at Hartman High School, it is recommended that all science teachers and administrators, along with special educators that work specifically with the science department, participate in professional development related to professional learning communities and content-based literacy. As previously mentioned, in order to implement PLCs with the desired structure, including common planning periods, at least one additional staff member should be hired. This teacher would continue to teach assigned science courses and would be able and expected to fully participate in all department initiatives along with the other teachers in the department.

Administrators in the building would have the primary role of spearheading the effort to implement PLCs and increase content-based literacy efforts in the classroom. Though administrators do not teach science courses, their participation and support of teachers as PLCs are launched is critical to the long-term success of the intervention. While PLCs are in their infancy, as well as when they have become well-oiled machines, administrators must monitor the progress of PLCs. This should be done in several ways, including active participation in PLC meetings and the review of notes and resources developed during the meetings.

Along with administrators, department heads must also have high expectations for those in the department with regard to PLCs and content-based literacy. Department heads will participate in their own content-specific PLC but must also communicate and visit with other PLCs to ensure that the process is being implemented and maintained with fidelity. Much like the administrators, department heads can review materials from the meetings to monitor progress of each PLC.

Aside from securing funds for professional development, teachers, by far, have the biggest responsibility for PLCs. The work done in the PLC is a direct result of the efforts of the teachers in each group, who must capitalize on common planning opportunities to see PLCs come to full fruition. Teachers will utilize the TLC during their PLC meetings to unpack standards, develop aligned materials, analyze data, and develop interventions to reach students who have not yet mastered the skills and content required of the curriculum. These teachers will also be responsible for communicating this work to other parties in the school, including department heads and administrators.

Timeline

The process of implementing the recommendations of this study should occur over the course of, at minimum, the first semester of the school year, though allowing an entire school year for full implementation is suggested. By extending the amount of time to implement the suggestions, teachers and administrators will be better able to work through the process and develop their skills and expertise in a collaborative manner.

Prior to beginning the implementation of PLCs, school administrators must ensure that teachers of the same course have the same planning period. This should take place in the spring prior to beginning PLC work. Throughout the summer months, creating buy-in from key teachers in the science department will help to facilitate a smooth transition in the fall when all teachers return. It will also provide a foundation upon which PLCs will function in the absence of an administrator.

Though the science department and administrators both consist of teachers who have experience elsewhere, it is important that all teachers receive professional development regarding the function and purpose of PLCs. Doing so will begin the process of building community within the groups and helping all parties to understand the common goal. This training should take place prior to the beginning of the school year so that the PLCs can begin meeting as early as possible.

During the first quarter and beyond, PLC groups will work through the TLC to unpack standards, develop common assignments and assessments, and implement interventions and enrichment opportunities for students. Administrators will participate in PLCs on a regular basis, and on-going professional development will be provided, especially related to assessing students at an appropriate rigor level and implementing literacy components in the science classroom. Outside of PLC time, teachers will utilize the materials developed by the group with their classes, collecting data about student performance for future PLC discussions. Specific to the MEA for Science, PLCs will use meeting time in the spring to identify areas of weakness specific to their courses to share with the 11th grade teachers. This data will help the 11th grade teachers to provide targeted test preparation lessons for students in the month leading up to the MEA assessment date.

Solution Implications

The implications of this study come from the central research question, "How can the problem of low test scores on the Maine Educational Assessment for Science be solved at Hartman High School?" While the primary purpose of the study is to improve students' scores on the MEA for Science, strategies and resources for improving instruction, which would have a positive impact on students' performance on the assessment, are recommended. These positive changes would have an impact, not only for students, but also for the school, including teachers and administrators, the school district, and the greater community.

Students

When equipped with the skills and understandings necessary to perform well on course-related assessments, students benefit in many ways. First and foremost, the knowledge gained as a whole will serve them throughout their lives as they progress through their education and begin their careers. Students may also benefit from college and scholarship opportunities based on their performance on the MEA for Science. Though the scores from the assessment are not required to be submitted with applications for either of these, students may submit them if they so choose.

The implementation of PLCs has a direct benefit for students in that their learning is closely monitored through the use of the TLC. The TLC requires that formative assessment data is analyzed, and students' performance is enhanced, whether they are falling behind or excelling, through remediation and enrichment opportunities. This focus on their specific learning needs will deepen their understandings and applications of skill related to the course content as measured by the MEA for Science, while also strengthening their skill set that they will carry beyond the science classroom. Likewise, focusing on literacy will also boost students' performance on the science assessment, but these skills also transfer to other assessments, such as the SAT and course-embedded tests. Literacy skills are universally applicable and absolutely necessary to thrive in modern society.

Teachers

When students perform well on an assessment, it is not only a reflection of the students' learning, but also of the teacher's teaching. It is reasonable to assume that when students perform well, teachers feel a sense of pride and accomplishment, as well. This increase in efficacy can lead to a continued effort to improve instructional practices, which can also translate to a continued increase in student performance in future years.

Utilizing PLCs as a method to analyze student data, which is then used to make instructional decisions, offer teachers job-embedded professional development opportunities, including informal learning from their peers. The professional development necessary to effectively implement PLCs strengthens teachers' skill set, and the learning that takes place within the PLC from one another, is authentic and often quickly applied.

Similarly, professional development with content-based literacy allows teachers to grow in their practice. Developing and becoming familiar with strategies to support the intentional incorporation of literacy into daily instruction will support the other instructional practices that teachers are utilizing, deepening students' understandings. This small change in instruction can lead to great results in students' performance, essentially supporting learning in a way that makes teachers' jobs a bit easier.

Two potential downfalls to the use of PLCs and literacy-specific strategies, however, are the reduction of individual planning time and an uncomfortableness with supporting literacy in the science classroom. The current schedule at Hartman High School allows for most teachers to have a full planning period twice a week, with a shortened period, with half as much time, once a week. Implementing PLCs, which require teachers to meet on a regular basis, preferably weekly, would reduce this time considerably. This time is valuable to teachers, especially because it does not occur daily. During this time, teachers plan lessons, develop resources, and grade student work; while PLCs support most of this work, teachers are often protective of this time, and convincing them to give up this time will likely be, at least initially, a difficult task.

Administrators

Administrators are no doubt tasked with a number of responsibilities, one of which, of course, is ensuring that the instruction taking place in the school is effective. Often, this effectiveness, along with the overall quality of the school, is largely based on students' performance on standardized assessments such as the MEA for Science. Implementing PLCs to allow teachers to collaboratively plan instruction and analyze student performance data, which is intended to improve students' performance in general, but also on the MEA. Likewise, the focus on content-based literacy will also improve students' understandings of what they have read and ability to analyze information presented on the MEA for Science. Both of these factors will lead to improved scores, which reduces the amount of time administrators must spend on improving instructional practices. Because of this, administrators can shift their focus to other matters knowing that using the TLC in PLCs and incorporating literacy into daily instruction is improving results on the MEA for Science.

As scores improve from the use of PLCs and content-based literacy practices, the school may become more eligible for certain types of funding. Because scores are often used as an indicator of school quality, improved scores relay the message that Hartman High School is a high-quality school and is worth an investment. Various companies may grant monies to the school and special programs may be available because investors will be of the impression that the school is dedicated to high-quality instruction, especially in the sciences, which will allow for greater opportunities for students. This is most certainly a positive reflection on the administration of the school, as they are the instructional leaders in the building.

Several challenges face administrators when considering these two recommendations. First, creating buy-in for both PLCs and content-based literacy practices can be daunting. Often, teachers are reluctant to be on board early in the process because of its unfamiliarity and the feeling of being overwhelmed with initiatives. To combat this, administrators must be certain to eliminate other responsibilities in favor of these, preferably because these strategies will eliminate the need for the other tasks. Also, monitoring PLCs and the work within requires an active effort from administrators, who are already weighed down by other tasks. If administrators can participate in the PLC meetings, however, they can not only help with the buy-in because of their own investment in the process, but they can also monitor the progress in an organic manner, which will, in the long run, reduce the amount of time dedicated to monitoring, as they will not need to look back on minutes or notes; instead, they will have experienced the process as it happens.

School District

Just as administrators will be viewed as competent and strong instructional leaders because their school is performing well on assessments, the same is true for the school district. When the school performs well on a standardized assessment, which is part of a public profile of the school, the community as a whole, both locally and statewide, will see the district as high-performing. This translates to a greater confidence in the school district and its leadership, which can mean more support for initiatives in the future, especially those that require additional funding. When the community feels that the school is producing high-quality students, they are more likely to continue to invest both time and money into the school and its programs to continue the success. Because implementing PLCs and improving content-based literacy skills help to improve these scores, it is imperative that the district be involved in the process.

Despite the benefits that could be provided to the district, there are several challenges that the district would face. First, providing funds for the hiring of an additional full-time teacher in place of a part-time teacher to allow for common planning among teachers, would be necessary. Because of budget constraints, this may be a challenge. Similarly, the professional development that would be required to implement both PLCs and content-based literacy practices can be pricey. Other professional development opportunities may suffer or be eliminated in favor of PLCs and literacy if additional money is not allocated to this line item in the budget. Acquiring additional funds of moving money from one line item to another can be challenging to convince others of.

Community

As mentioned, the overall student performance on the MEA for Science, is publicly-reported statistic. The students' high performance can increase the school's ranking among other schools, making the school a desirable place to work and the community one in which parents would want to raise their families. Both of these mean that increased scores on the assessment can lead to economic development for the community as a whole.

When families move into the area desiring a high-quality education for their children, the housing market becomes increasingly competitive, meaning the house prices soar and new homes will be developed. This, of course, increases that income to the town in terms of tax revenue. It also provides more employable individuals in the community, which can lead to greater economic growth when companies choose to expand and hire local workers or new companies move into the area. Producing graduates who are well-qualified for the work force gives confidence to businesses that they will be able to sustain their work force, making the area even more enticing.

Some of the challenges, however, are related to this economic aspect. Because additional funding is needed to support the two recommended strategies, the community may be called on to support these efforts. This can mean an increase in taxes. An increase in population may also happen too rapidly for the community to keep up. There may not be enough affordable housing, which can lead to over-inflated home costs, which could become a deterrent. It's also possible that if families were to move to the area rapidly, the schools would not be able to support the increase in students, essentially creating an undesirable classroom setting or a need for additional facilities to be built, translating, of course, to additional funding, which falls back on the community if it cannot be found within the current school budget.

Evaluation Plan

To evaluate the effectiveness of PLCs and content-based literacy practices, both goal- and outcomes-based assessment should be used, targeting both formative and summative needs. The combination of these approaches allows for on-going reflection during the implementation process and as the strategies are being used throughout the school year, while also looking at the final results for the overall effectiveness. However, it is important that the results of both components are used in tandem rather than being used in isolation.

The goal-based evaluation should be used to address teachers' personal goals related to the literacy component, as well as each PLC's goal regarding meetings and progress. It is quite likely that most science teachers have limited experience with teaching literacy in the science classroom. As such, using strategies related to this will be new, and teachers should begin to implement instructional practices gradually. Goals should be written to reflect what strategies will be used and how often, as well as the intended, measurable outcome. Quantitative data may be used by teachers to measure their own effectiveness in terms of student performance using these strategies, but the overall goal should be related to implementing literacy-based strategies.

These goal-oriented evaluations should be conducted by both the teachers and administrators. Teachers should reflect upon their own use and look for areas of strength and weakness within their practice, identifying areas upon which they can improve moving forward. These goals should be revised on a regular basis throughout the academic year to ensure continued growth as teachers build efficacy with both the literacy and PLC components. Administrators should also monitor these goals through a review of documentation from PLCs and lesson plans. Observations of instruction should also be considered, as they provide evidence of the implementation of literacy strategies and decisions made by the PLC group as a whole.

Analyzing test scores after students have taken the MEA for Science in the spring aligns with an outcomes-based evaluation. Once the scores are sent to the school by the MDOE, administrators and department chairs are the first to receive this information. Overall student performance, as well as that of specific students, can be reviewed by these parties. Typically, this information is not shared with individual teachers or with the department as a whole, but it is recommended that this changes so that trends in student performance can be identified. Finding a way to do an item analysis is also recommended, as this sort of evaluative practice can lead to a more focused approach for future instruction.

Delimitations are decisions made by the researcher to set the boundaries and limitations of the study. In this study, delimitations include limiting participation to teachers within the science department at Hartman High School, along with administrators and special educators specifically involved in science instruction. This limitation was made so as to include only participants who are fully invested in science instruction at the school. Only considering students' performance on the MEA for Science as an indicator for student learning is another delimitation, as it is not the only evidence available of students' progress with the learning standards. This boundary was set because the universality of its application; that is, all 11th grade students at the school take the test, and the data provided allows for comparison between the students at Hartman High School and other students across the state. Additionally, interview participants were selected based upon the amount of time they have been employed at Hartman High School as a science teacher or administrator because of their familiarity with the test and its results.

Likewise, limitations, or weaknesses of the study that cannot be controlled, exist in this study, as well. The limitations of this study include the gender, age, ethnicity, socioeconomic status, and special services received of the students who take the assessment each year. These details can have an impact on students' performance on the assessment, but these factors are out of the control of the researcher and the school itself. The number of participants in this study is also a limitation of the study, as all but one science teacher in the department participated, and all administrators took part. This is a limitation due to the fact that no other perspectives can be provided to the study, and the history of the assessment data is limited because of the lack of longevity within the department. Also limiting the scope of this study is the lack of information available from the MDOE. An item analysis or score report by standard or reporting category could certainly help teachers and administrators improve instruction, but the reports provided do not include this information.

Further research is recommended to help solve the problem of low test scores on the MEA for Science. Recommendations include a consideration of the length of teachers' careers in relation to students' performance, as well as how unpacking standards in particular can influence students' scores. In particular, identifying if teachers who are new to the field, and likely have been subjected to more standardized tests as students, do better when preparing students for the MEA for Science than their older colleagues, who may not have had as much emphasis placed on testing during their formative years. In a similar manner, the standards that are required to be taught in each school have become more uniform as time has passed, and as such, identifying the ways in which schools are most effective at implementing these standards and the deeper understanding teachers have of the standards and the impact these practices have on students' performance on the assessment can provide more evidence for best practices in the classroom.

Summary

The goal of this study was to identify factors that are impacting students' performance on the MEA for Science at Hartman High School, a suburban high school located in southeastern Maine. Identifying these factors is important, as the scores on the MEA for Science are one of the publicly-available indicators of students' learning success, especially in science. By triangulating data from interviews, an online survey, and a review of documents, it is evident that there are areas of improvement upon which the faculty at Hartman High School can improve. This study has illustrated the importance of creating a uniform approach to planning instruction and assessments among teachers at Hartman High School, including the implementation of specific content-based literacy strategies in the science classroom.

By implementing PLCs at Hartman High School, teachers will be able to work collaboratively to develop content and materials that is aligned to both the standards and the rigor of the standards required at Hartman High School. Participants indicated a lack of planning time and professional development related to instructional practices, as well as a need to improve literacy skills among students. Offering professional development to best support collaborative planning efforts and using literacy-building activities into science curriculum is critical to moving the instructional practices forward at Hartman High School in a way that will improve scores on the MEA for Science. It should be noted, however, that the benefits of these efforts will not be fully realized until these strategies have been implemented for a total of three years, as students in the 9th grade will not take the test until their 11th grade year; the students who will experience instruction based on PLC practices and intentional literacy instruction in the science classroom for the duration of their high school experience will provide the best evidence of the effects of these strategies. Improving scores on the MEA for Science can lead not only to improved student and teacher efficacy, but also to various benefits for the school and the school district, as well as economic gains for the community as a whole.

REFERENCES

Achievement Network, LLC. (2020). Not just another initiative. *Achievement Network.* Retrieved from https://www.achievementnetwork.org/school-leaders

Akiba, M. & Liang, G. (2016). Effects of teacher professional learning activities on student achievement growth. *The Journal of Educational Research, 109*(1), 99-110, doi: 10.1090/00220671.2014.924470

American University (2019). What makes professional development effective for teachers? *School of Education, American University*. Retrieved from https://soeonline.american.edu/blog/what-makes-professional-development-for-teachers-effective

Aucejo, E. M., & Romano, T. F. (2016). Assessing the effect of school days and absences on test score performance. *Economics of Education Review, 55*(1) 70-87. doi:10.1016/j.econedurev.2016.08.007

Battelle for Kids (2011). Successfully implementing transformational change in education: Lessons learned about the importance of effective change leadership and strategic communications. Retrieved from https://www2.ed.gov/programs/racetothetop/communities/bfk-rttt-communications-lessons-learned.pdf

Berliner, D. (2001). *Our schools vs. theirs: Averages that hide the true extremes.* Retrieved from https://nepc.colorado.edu/publication/our-schools-vs-theirs-averages-that-hide-true-extremes

Bertolini, K., Stremmel, A., & Thorngren, J. (2012). Student achievement factors. *South Dakota State University College of Education and Human Science Department of Teaching,* *Learning, and Leadership.* Retrieved from https://files.eric.ed.gov/fulltext/ED568687.pdf

Betts, J. R., Reuben, K. S., & Danenberg, A. (2000). *Equal resources, equal outcomes? The distribution of school resources and student achievement in California*. Retrieved from https://www.ppic.org/content/pubs/report/R_200JBR.pdf

Bickman, L. & Rog, D. J. (2009). *The SAGE handbook of applied social research methods* (2nd Ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Binet, A. & Simon, T. (1916). *The development of intelligence in children: The Binet-Simon scale)*. Baltimore, MD: Williams & Wilkins, Co. doi: http://dx.doi.org.ezproxy.liberty.edu/10.1037/11069-000

Blatchford, P., Chan, K. W., Galton, M., Lai, K. C., & Lee, J. C. (Eds.). (2016). *Class size: Eastern and Western perspectives*. New York, NY: Routledge.

Bosworth, R. (2014). Class size, class composition, and the distribution of student achievement. *Education Economics*, *22*(1/2), 141-165. doi:10.1080/09645292.2011.568698

Boynton, M. & Boynton, C. (2005). *The educator's guide to preventing and solving discipline problems*. Alexandria, VA: ASCD.

Brenchley, C. (2015). What is ESEA? *Homeroom: The Official Blog of the U.S. Department of Education*. Retrieved from https://blog.ed.gov/2015/04/what-is-esea/.

Brown, G. and Desforges, C. (2006). *Piaget's theory* (1st ed.). London, England: Routledge. doi: https://doi-org.ezproxy.liberty.edu/10.4324/9780203715796

Caldas, S. J., & Bankston, C., III. (1997). Effect of school population socioeconomic status on individual academic achievement. *The Journal of Educational Research, 90*(5) 269–277. doi:10.1080/00220671.1997.10544583

California County Superintendents Educational Services Association (2016). Best practices in teacher and administrator induction programs. Retrieved from http://ccsesa.org/wp-content/uploads/2016/06/Best-Practices-in-Teacher-and-Administrator-Induction-Programs.pdf

Carnegie Mellon University (2019). Why should assessments, learning objectives, and instructional strategies be aligned? *Carnegie Mellon University Eberly Center.* Retrieved from https://www.cmu.edu/teaching/assessment/basics/alignment.html

Caro, D. H., McDonalds, J. T., & Willms, J. D. (2009). Socio-economic status and academic achievement trajectories from childhood to adolescence. *Canadian Journal of Education, 32*(3), 558-590.

Case, B. (2005). Horizontal and vertical alignment. *Pearson Policy Report.* Retrieved from https://images.pearsonassessments.com/images/tmrs/tmrs_rg/HorizontalVerticalAlignment.pdf?WT.mc_id=TMRS_Horizontal_and_Vertical_Alignment

Cerezo, L. (n.d.). Standardized testing. *Universidad de Murcia*. Retrieved from https://webs.um.es/lourdesc/miwiki/lib/exe/fetch.php?id=units&cache=cache&media=standardized_testing.pdf

Cherry, K. (2019a). Alfred Binet and the history of IQ testing. *VeryWell Mind*. Retrieved from https://www.verywellmind.com/history-of-intelligence-testing-2795581

Cherry, K. (2019b). The 4 stages of cognitive development: Background and key concepts of Piaget's theory. *VeryWell Mind.* Retrieved from https://www.verywellmind.com/piagets-stages-of-cognitive-development-2795457

Chingos, M. M. (2010). The impact of universal class-zie reduction policy: Evidence from Florida's statewide mandate. Cambridge, MA: Harvard University.

Christensen, V. (2018). What makes a test standardized? *The Classroom.* Retrieved from https://www.theclassroom.com/test-standardized-6680561.html

Clay, V. (2016). Intrinsic motivation vs. standardized tests. *Edutopia*. Retrieved from https://www.edutopia.org/blog/intrinsic-motivation-vs-standardized-tests-valencia-clay

College Board (2019a). Test dates and deadlines. *College Board: SAT Suite of Assessments.* Retrieved from https://collegereadiness.collegeboard.org/sat/register/dates-deadlines

College Board (2019b). *SAT school day Apr 09, 2019, 11th grade – question analysis – Form R: Evidenced-based reading and writing questions.*

Committee on Undergraduate Science Education (1997). *Science teaching reconsidered: A handbook.* Washington, D. C.: National Academy Press.

Cox, J. (2019). 5 teaching strategies to help boost test scores. *TeachHub.com: K-12 News, Lessons & Shared Resources by Teachers, For Teachers.* Retrieved https://www.teachhub.com/5-teaching-strategies-help-boost-test-scores

Creswell, J. W. (2015). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Boston, MA: Pearson Education, Inc.

Creswell, J. W. & Poth, C. N. (2018). *Qualitative inquiry & research design: Choosing among five approaches.* Los Angeles, CA: SAGE Publications, Inc.

Curry, D. P., Reeves, E., McIntyre, C. J., & Capps, M. (2018). Do teacher credentials matter? An examination of teacher quality. *Curriculum and Teaching Dialogue, 20*(1/2), 9-18, 176-177, 179-180.

Cuseo, J. (2007). The empirical case against large class size: Adverse effects on the teaching, learning, and retention of first-year students. *Journal of Faculty Development, 21*(5), 5-21.

Darling-Hammond, L. (2015). Can value added add value to teacher evaluation? *Educational Researcher*, *44*(2), 132-137. doi:10.3102/0013189X15575346

Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Heilig, J. V. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. Retrieved from https://www.nctq.org/nctq/research/1114011196655.pdf

Darling-Hammond, L., & Youngs, P. (2002). Defining "highly qualified teachers": What does "scientifically based research" actually tell us? *Educational Researcher, 31*(9), 13-25.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). Effective teacher professional development. *Learning Policy Institute.* Retrieved from https://learningpolicyinstitute.org/product/effective-teacher-professional-development-report

Davies, R. (2019). 5 reasons your students misbehave for a substitute & how to prevent it. *differentiated TEACHING.* Retrieved from https://www.thethirdwheelteacher.com/why-students-misbehave-for-the-substitute/

Day, C., Gu, Q., & Sammons, P. (2016). The impact of leadership on student outcomes: How successful school leaders use transformational and instructional strategies to make a difference. *Educational Administration Quarterly, 52*(2), 221-258.

DeMonte, J. (2015). A million new teachers are coming: Will they be ready to teach? *American Institutes for Research.* Retrieved from https://www.air.org/sites/default/files/downloads/report/Million-New-Teachers-Brief-deMonte-May-2015.pdf

Donaldson, Jr., 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.

Drake, S. M. & Burns, R. G. (2004). *Meeting standards through integrated curriculum.* Alexandria, VA: ASCD.

DuFour, R. (2004). What is a "professional learning community?" *Educational Leadership, 61*(8), 6-11.

Dweck, C. (2000). *Self-theories: Their role in motivation, personality and development*. Philadelphia, PA: Taylor & Francis.

Education Post (2019). The ABC's of ESEA, ESSA, and no child left behind. *Education Post*. Retrieved from https://educationpost.org/the-abcs-of-esea-essa-and-no-child-left-behind/

Education World (2019). Boosting test scores: "Principal" strategies that work. *Education World.* Retrieved from https://www.educationworld.com/a_admin/admin/admin366.shtml

Education World (2012). How to facilitate job-embedded professional development. *Education World*. Retrieved from https://www.educationworld.com/a_admin/job-embedded-professional-development.shtml

EF Academy (2019). 10 benefits of small class sizes. *EF Academy Blog.* Retrieved from https://www.ef.com/wwen/blog/efacademyblog/10-benefits-small-class-sizes/

Ellis, S. (2018). Why standardizes tests matter beyond college admissions. *U.S. News.* Retrieved from https://www.usnews.com/education/blogs/college-admissions-playbook/articles/2018-04-23/3-benefits-of-standardized-testing-beyond-college-admissions

Ehrenberg, R. G., & Brewer, D. J. (1995). Did teachers' verbal ability and race matter in the 1960s? Coleman revisited. *Economics of Education Review, 14*(1), 1-21. doi:10.1016/0272-7757(94)00031-Z

Finn, J. D., Pannozzo, G. M., & Achilles, C. M. (2003). The why's of class size: Student behavior in small classes. *Review of Educational Research, 73*(3), 321-368.

Fischer, C., Fishman, B., Dede, C., Eisenkraft, A., Frumin, K., Foster, B., McCoy, A. (2018). Investigating relationships between school context, teacher professional development, teaching practices, and student achievement in response to a nationwide reform. *Teaching and Teacher Education, 72*(1), 107-121. doi:10.1016/j.tate.2018.02.011

Fletcher, D. (2009). Standardized testing. *Time*. Retrieved from http://content.time.com/time/nation/article/0,8599,1947019,00.htm

Fordham Institute (2016). Why are 6.1 million students using Newsela? *Thomas B. Fordham Institute*. Retrieved from https://fordhaminstitute.org/national/commentary/why-are-61-million-students-using-newsela

Frey, B. B. (2018). Internal consistency. In: Frey, B. B. (Ed.), *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation.* Thousand Oaks, CA: SAGE Publications, Inc. doi: https://dx.doi.org/10.4135/9781506326139.n340

Gall, M. D., Gall, J. P., and Borg, W. R. (2007). *Educational research: An introduction* (8th edition). Boston, MA: Pearson Education, Inc.

Gallagher, C. J. (2003). Reconciling a tradition of testing with a new paradigm. *Educational Psychology Review, 15*(1) 83-99. doi:10.1023/A:1021323509290

Gallagher, N. K. (2015). New assessment tests for Maine students to be rolled out: Measured Progress, Inc. will replace the Smarter Balanced test; details will be unveiled Tuesday.

Portland Press Herald: Schools and Education. Retrieved from https://www.pressherald.com/2015/12/02/new-tests-for-maine-pupils-to-be-rolled-out/

Garcia, L. E. & Thornton, O. (2014). The enduring importance of parental involvement. *neaToday*. Retrieved from http://neatoday.org/2014/11/18/the-enduring-importance-of-parental-involvement-2/

Gehlbach, H., Brinkworth, M. E., King, A. M., Hsu, L. M., McIntyre, J., & Rogers, T. (2016). Creating birds of similar feathers: Leveraging similarity to improve teacher-student relationships an academic achievement. *Journal of Educational Psychology,* 108(3), 342-352. doi: http://dx.doi.org.ezproxy.liberty.edu/10.1037/edu0000042

Glossary of Education Reform, The (2015). Norm-referenced test. In *The Glossary of Education Reform*. Retrieved from https://www.edglossary.org/norm-referenced-test/

Greenberg, M. (2014). It's time for a new definition of accreditation. *The Chronicle of Higher Education, 60*(20).

Greenville Public School District (n.d.). New teacher induction program. Retrieved from https://www.gvillepublicschooldistrict.com/docs/district/depts/20/greenville%20public%20school%20district%20new%20teacher%20induction%20program.pdf?id=3990

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66*(3), 361-398.

Grissom, J. A., Kalogrides, D., & Loeb, S. (2015). Using student test scores to measure principal performance. *Educational Evaluation and Policy Analysis*, *37*(1), 3-28. doi:10.3102/0162373714523831

Guskey, T. R. (2003). How classroom assessments improve student learning, *Educational Leadership, 60*(5), 6-11.

Hartman High School (2019). Hartman High School program of studies 2019-2020. Retrieved from [URL hidden to protect identity of educational institution.]

Haver, N. (n.d.). Why teacher induction programs are a necessity for the start of the school year. *Whitby School*. Retrieved from https://www.whitbyschool.org/passionforlearning/why-teacher-induction-programs-are-a-necessity-for-the-start-of-the-school-year

Hawker Brownlow Education (n.d.). Quick reference guide: criterion-referenced vs. norm-referenced assessment. Retrieved from http://files.hbe.com.au/flyerlibrary/Brigance/Criterion-referenced%20vs.%20Norm-referenced%20Assessment.pdf

Heck, R. H., & Hallinger, P. (2014). Modeling the longitudinal effects of school leadership on teaching and learning. *Journal of Educational Administration*, *52*(5), 653-681.

Herman, J. L., Osmundson, E., & Dietel, R. (2010). Benchmark assessment for improved learning (AACC Report). Los Angeles, CA: University of California.

Hulleman, C. S. & Hulleman, T. (2018). An important piece of the student motivation puzzle. Retrieved from https://www.future-ed.org/reversing-the-decline-in-student-motivation/

Higgins, J. (2014, October 28). Does class size matter? Research reveals surprises. *The Seattle Times.* Retrieved from www.seattletimes.com

Hoxby, C. M. (1998). The effect of class size on student achievements: New evidence from population variation. Cambridge, MA: National Bureau of Economic Research.

Ingersoll, R. M. (2012). Beginning teacher induction: What the data tells us. *Education Week*. Retrieved from https://www.edweek.org/ew/articles/2012/05/16/kappan_ingersoll.h31.html

Ingersoll, R. M. & Strong, M. (2011). The impact of induction and mentoring programs for beginning teachers: A review of the research. *Review of Educational Research, 81*(2), 201-233.

Jacob, B. & Rothstein, J. (2016). The measurement of student ability in modern assessment systems. *Journal of Economic Perspectives, 30*(3), 85-108.

Jacobson, L. (2018). Standards for teacher induction programs released. *New Teacher Center*. Retrieved from https://newteachercenter.org/news-story/standards-for-teacher-induction-programs-released/

Jessie, L. G. (2012). Creating buy-in for PLCs. *All Things PLC.* Retrieved from https://www.allthingsplc.info/blog/view/161/creating-buy-in-for-plcs

Johnson, E., & Semmelroth, C. L. (2014). Special education teacher evaluation: Why it matters, what makes it challenging, and how to address these challenges. *Assessment for effective intervention*, *39*(2), 71-82.

Joyner, R. L., Rouse, W. A., & Glatthorn, A. A. (2013). *Writing the winning thesis or dissertation: A step-by-step guide* (3rd ed.). Thousand Oaks, CA: Corwin.

K12 Academics (2019). Common use for norm-referenced test. Retrieved from https://www.k12academics.com/standardized-testing/norm-referenced-test/common-use

Kastenbaum, S. (2012). The high stakes of standardized tests. *CNN: Schools of thought.* Retrieved from http://schoolsofthought.blogs.cnn.com/2012/05/17/the-high-stakes-of-standardized-tests/

Kalenze, E. (2016). How positive student-teacher relationships impact student motivation. *Search Instititute*. Retrieved from https://www.search-institute.org/positive-teacher-student-relationships-impact-student-motivation/

Kennedy, E. (2003). *Raising test scores for all students: An administrator's guide to improving standardized test performance.* Thousand Oaks, CA: Corwin Press.

Kielwitz, H. A. (2014). Comprehensive teacher induction: Linking teacher induction to theory. Retrieved from https://files.eric.ed.gov/fulltext/ED557137.pdf

Klein, Alyson (2015). The nation's main K-12 law at 50 years. *Education Week*. Retrieved from https://www.edweek.org/ew/section/multimedia/the-nations-main-k-12-law-a-timeline.html

Klein, Alyson (2018). How does science testing work under ESSA? *Politics K-12: Your Education Roadmap*. Retrieved from http://blogs.edweek.org/edweek/campaign-k-12/2018/10/science-essa-testing-states-plans.html

Koedel, C. & Polikoff, M. (2017). Big bang for just a few bucks: The impact of math textbooks in California. *Evidence Speaks Reports, 2*(5).

Kosanovich, M. & Rodriguez, E. (2019). Developing effective teachers through targeted professional development. *Middle School Matters*. Retrieved from https://greatmiddleschools.org/developing-effective-teachers-through-targeted-professional-development/

Ladd, H. F. (2017). No child left behind: A deeply flawed federal policy. *Journal of Policy Analysis and Management, 36*(2), 461-469. doi: https://doi.org/10.1002/pam.21978

Lane, J. F. (2018). Superintendent's memo: Testing requirements under the Every Student Succeeds Act and the 2017 standards of accreditation. Retrieved from http://www.doe.virginia.gov/administrators/superintendents_memos/2018/299-18.docx

Lavrakas, P. J. (2008). Test-retest reliability. In: Lavrakas, P. J. (Ed.), *Encylopedia of Survey Research Methods.* Thousand Oaks, CA: Sage Publications, Inc. doi: https://dx.doi.org/10.4135/9781412963947.n581

Lestrund, M. (2013). Educational interventions. In: Volkmar, F. R. (Ed.), *Encyclopedia of Autism Spectrum Disorders*. New York, NY: Springer.

Lewis, D., Madison-Harris, R., Muoneke, A., & Times, C. (2010). Using data to guide instruction and improve student learning. Retrieved from http://www.sedl.org/pubs/sedl-letter/v22n02/using-data.html

Liberty University (n.d.). *Institutional review board handbook.* Retrieved from https://www.liberty.edu/media/9997/IRB_Handbook.pdf

Liberty University (2019a). Applied dissertation template 2018-2019 [Blackboard]. Retrieved from https://learn.liberty.edu/bbcswebdav/pid-27185637-dt-content-rid-371327500_1/xid-371327500_1

Liberty University (2019b). Milestone 3: IRB review [Blackboard]. Retrieved from https://learn.liberty.edu/webapps/blackboard/content/listContent.jsp?course_id=_480077_1&content_id=_27185496_1

Loeb, S., & Figlio, D. (2011). School accountability. In Hanushek, E. A., Machin, S., & Woessmann, L. (Eds.), *Handbook of the Economics of Education, Vol. 3*. San Diego, CA: North Holland.

Lynch, M. (2015). The role of public schools in the advancement of communities. *The Edvocate*. Retrieved from https://www.theedadvocate.org/the-role-of-public-schools-in-the-advancement-of-the-communities/

Lynch, M. (2016a). The pro's and con's of alternate-route teacher preparation programs. Retrieved from https://www.theedadvocate.org/pros-cons-alternate-route-teacher-preparation-programs/

Lynch, M. (2016b). Understanding federal funding part I: 3 types of school funding. *The Edvocate.* Retrieved from https://www.theedadvocate.org/understanding-federal-funding-part-3-types-school-funding/

Maine Department of Education (n.d.a). Maine assessment and accountability reporting system (MAARS). Retrieved from https://lms.backpack.education/public/maine

Maine Department of Education (n.d.b). The Maine educational assessments: Why participation matters. *Maine Department of Education*. Retrieved from https://www.maine.gov/doe/sites/maine.gov.doe/files/inline-files/Why%20Participation%20Matters-1.pdf

Maine Department of Education (1990). *Maine's common core of learning*. August, ME: Maine Department of Educational and Cultural Services.

Maine Department of Education (2018a). 2017-18 MEA science test designs. Retrieved from https://www.maine.gov/doe/sites/maine.gov.doe/files/inline-files/2017-18TestDesignSheet.pdf

Maine Department of Education (2018b). Achievement level descriptors and score ranges. Retrieved from https://www.maine.gov/doe/sites/maine.gov.doe/files/inline-files/MEA_2018_ALDs%2BCut%20Scores_10-09-18.pdf

Maine Department of Education (2018c). Maine comprehensive assessment system (MECAS). Retrieved from https://www.maine.gov/doe/Testing_Accountability/MECAS

Maine Department of Education (2018d). MEA science. Retrieved from https://www.maine.gov/doe/Testing_Accountability/MECAS/materials/measscience

Maine Department of Education (2019). Maine student performance on state assessments. Retrieved from https://public.tableau.com/profile/maine.department.of.education#!/vizhome/MaineAssessments/MaineAssessments

Maine School Administrative District 75 (2019). *State assessment FAQs.* Retrieved from http://link75.org/district/teaching_and_learning/state_assessment_f_a_qs

Maine School Administrative District 75 (2020). *Benefit rate sheet*. Retrieved from https://ivisionsess.tylerhost.net/msad75ess/LinkClick.aspx?fileticket=z-BSVqU1pS4%3d&tabid=121&portalid=0&mid=474

Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). Making sense of data-driven decision making in education. *Rand Occasional Papers*. Retrieved from https://www.rand.org/content/dam/rand/pubs/occasional_papers/2006/RAND_OP170.pdf

Martinez, L. (2016). Motivating students before and after testing. Retrieved from https://engage2learn.org/motivating-students-before-and-after-testing/

Massachusetts Department of Education (2002). Induction of beginning educators. *Massachusetts Department of Elementary and Secondary Education*. Retrieved from http://www.doe.mass.edu/educators/mentor/teachers.html

Mathewson, T. G. (2019). How to unlock students' internal drive for learning: Intrinsic motivators can be key to student achievement – but extrinsic motivation dominates classrooms. *The Hechinger Report*. Retrieved from https://hechingerreport.org/intrinsic-motivation-is-key-to-student-achievement-but-schools-kill-it/

Matthews, J. (2012). Why textbooks don't work and hurt schools. *The Washington Post*. Retrieved from https://www.washingtonpost.com/blogs/class-struggle/post/why-our-textbooks-dont-work/2012/02/25/gIQAvI16ZR_blog.html?noredirect=on

McKay, S. (2015). Using new research to improve student motivation. *Carnegie Foundation for the Advancement of Teaching*. Retrieved from https://www.carnegiefoundation.org/blog/using-new-research-to-improve-student-motivation/

McLeod, S. (2018). Jean Piaget's theory of cognitive development. *Simply psychology*. Retrieved from https://www.simplypsychology.org/piaget.html

Mertler, C. A. (2007). *Interpreting standardized test scores: Strategies for data-driven decision making.* Thousand Oaks, CA: SAGE Publications, Inc.

Miller, R. T., Murnane R. J., & Willett, J. B. (2007, August). *Do teacher absences impact student test scores? Longitudinal evidence from one urban school district.* Retrieved from http://www.nber.org/papers/w13356

National Association of Elementary School Principles (n.d.). Using student achievement data to support instructional decision making. Retrieved from https://www.naesp.org/sites/default/files/Student Achievement_blue.pdf

National Association of Elementary School Principals (2006). Using data to steer alignment, instruction and improvement. Retrieved from https://www.naesp.org/sites/default/files/resources/1/Pdfs/EarlyChildCS/silverspring.pdf

National Center on Response to Intervention (n. d.). Using fidelity to enhance program implementation with an RTI framework. Retrieved from https://rti4success.org/sites/default/files/Using%20Fidelity%20to%20Enhance%20Program%20Implementation_PPTSlides.pdf

National Council of Teachers of English (2014). Why class size matters today. *Position Statements.* Retrieved from http://www2.ncte.org/statement/why-class-size-matters/

National Education Association (2008). Class size reduction: A proven reform strategy. Retrieved from http://www.nea.org/assets/docs/PB08_ClassSize08.pdf

National Education Association (2019a). History of standardized testing in the United States. *National Education Association: Great Public Schools for Every Student.* Retrieved from http://www.nea.org/home/66139.htm

National Education Association (2019b). Status of substitute teachers: A state-by-state summary. Retrieved from http://www.nea.org/home/14813.htm

National Research Council (2014). *Developing assessments for the Next Generation Science Standards*. Washington, DC: The National Academies Press. https://doi.org/10.17226/18409

Nest, P. (2019). Test prep & review strategies for grades 6-8. *National Education Association.* Retrieved from http://www.nea.org/tools/lessons/Test-Prep-Review-Strategies-Grades-6-8.htm

New Teacher Center (2016a). Analysis of student work. *New Teacher Center*. Santa Cruz, CA: New Teacher Center.

New Teacher Center (2016b). Coaching and observation strategies. *New Teacher Center*. Santa Cruz, CA: New Teacher Center.

New Teacher Center (2016c). Mentor formative assessment and support: Promoting growth. *New Teacher Center*. Retrieved from https://newteachercenter.org/wp-content/uploads/mentor-formative-assessment_practice-brief.pdf

New Teacher Center (2019). Our approach. *New Teacher Center*. Retrieved from https://newteachercenter.org/our-approach/

New York City Department of Education (2019). NY State High School Regents Exams. *New York City Department of Education.* Retrieved from https://www.schools.nyc.gov/school-life/learning/testing/ny-state-high-school-regents-exams

No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319 (2002).

O'Malley, K. (2012). Standardized testing: What is it and how does it work? *Pearson*. Retrieved from https://www.pearsoned.com/standardized-testing-work/

Okeke, C. I. O., Shumba, J., Rembe, S., & Sotuku, N. (2015). Demographic variables, work-stimulated stressors and coping strategies of pre-school educators: A concept paper. *Journal of Psychology*, *6*(1), 91-101.

Osborne, M. (2018). The class size debate: 5 important findings from research. North Carolina Center for Public Policy Research. Retrieved from https://nccppr.org/class-size-debate-5-important-findings-research/

Pennsylvania State University, The (2020). Newsela to help with assessing student reading level. *Penn State College of Agricultural Sciences*. Retrieved from https://aese.psu.edu/teachag/old/technology/newsela-to-help-with-assessing-student-reading-level

Phelan, C. & Wren, J. (2006). Exploring reliability in academic assessment. Retrieved from https://chfasoa.uni.edu/reliabilityandvalidity.htm

Piaget, J. (1951). *Play, dreams, and imitation in childhood.* London, England: Routledge. doi: https://doi-org.ezproxy.liberty.edu/10.4324/9781315009698

Piaget, J. (1952). *The origin of intelligence in children.* New York, NY: International Universities Press, Inc.

Piaget, J. & Inhelder, B. (1958). *The growth of logical thinking from childhood to adolescence*. London, England: Routledge. doi: https://doi-org.ezproxy.liberty.edu/10.4324/9781315009674

Piaget, J. (1964). *The early growth of logic in the child*. London, England: Routledge. doi: https://doi-org.ezproxy.liberty.edu/10.4324/9781315009698

Pianta, R. C., & Ansari, A. (2018). Does attendance in private schools predict student outcomes at age 15? evidence from a longitudinal study. *Educational Researcher, 47*(7), 419–434 doi:10.3102/0013189X18785632

Piro, J. S., Dunlap, K., & Shutt, T. (2014). A collaborative data chat: Teaching summative assessment data use in pre-service teacher education. *Cogent Education*, *1*(1), 968409.

Quinn, D. M. & Cooc, N. (2015). Science achievement gaps by gender and race/ethnicity in elementary and middle school: Trends and predictors. *Educational Researcher, 44*(6), 336-346.

Rahman, M. S. (2017). The advantages and disadvantages of using qualitative and quantitative approaches and methods in language “testing and assessment” research: A literature review. *Journal of Education and Learning, 6*(1), 102-112.

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, *50*(1), 4-36. doi:10.3102/0002831212463813

Salkind, N. J. (2010). Parallel forms reliability. In: Salkind, N. J. (Ed.), *Encyclopedia of Research Design*. Thousand Oaks, CA: Sage Publications, Inc. doi: https://dx.doi.org/10.4135/9781412961288.n301

Shandomo, H.M. (2010). The role of critical reflection in teacher education. *School-University Partnerships, 4*(1), 101-113.

Sierra, R. (2015). What is reflective teaching and why is it important? *RichmondShare*. Retrieved from https://www.richmondshare.com.br/what-is-reflective-teaching-and-why-is-it-important/

Sievertsen, H. H., Gino, F., & Piovesan, M. (2016). Cognitive fatigue influences students' performance on standardized tests. *Proceedings for the National Academy of Sciences of the United States of America, 113*(10), 2621-2624. doi: https://doi.org/10.1073/pnas.151947113

Solomon, S. (2016). What makes an effective teacher induction program? *GO Public Schools Oakland*. Retrieved from https://gopublicschoolsoakland.org/2016/12/what-makes-an-effective-teacher-induction-program/

Solution Tree (2020a). 2-day workshops. *Solution Tree*. Retrieved from https://www.solutiontree.com/events/workshops.html

Solution Tree (2020b). Our on-site professional learning for PLC at Work. *PLC at Work*. Retrieved from https://www.solutiontree.com/plc-at-work/pd-services

Springer, M. G. (2008). Accountability incentives: Do school practice educational triage? *EducationNext*. Retrieved from https://www.educationnext.org/accountability-incentives/

Stansbury, K. & Zimmerman, J. (2000). Lifelines to the classroom: Designing support for beginning teachers. *WestEd*. Retrieved from https://www.wested.org/online_pubs/tchrbrief.pdf

Steelman, S. (2018). How to mentor new teachers so they won't quit. *School Leaders Now*. Retrieved from https://schoolleadersnow.weareteachers.com/mentor-new-teachers/

Stenger, M. (2014). 5 research-based tips for providing students with meaningful feedback. Retrieved from https://www.edutopia.org/blog/tips-providing-students-meaningful-feedback-marianne-stenger

Stone, M. (2018). How Maine hurt education by trying to reform it. *Bangor Daily News: Maine Focus.* Retrieved from https://bangordailynews.com/2017/11/14/mainefocus/how-maine-hurt-education-by-trying-to-reform-it/

Stronge, J. H. (2018). *Qualities of effective teachers* (3rd ed.). Alexandria, VA: ASCD.

Sullivan, G. M. (2011). A primer on the validity of assessment instruments. *Journal of Graduate Medical Education, 3*(2), 119-120. doi: 10.4300/JGME-D-11-00075.1

Sweet, E. (2018). TLC for your PLC: 4 best practices. *Illuminate Education.* Retrieved from https://www.illuminateed.com/blog/2018/02/tlc-plc-4-best-practices/

Tarr, J. E., Reys, R. E., Reys, B. J., Chavez, O., Shih, J., & Osterlind, S. J. (2008). The impact of middle-grades mathematics curricula and the classroom learning environment on student achievement. *Journal for Research in Mathematics Education, 39*(3), 247-280.

Teaching Tolerance (2019). Teaching teachers: PD to improve student achievement. Retrieved from https://www.tolerance.org/professional-development/teaching-teachers-pd-to-improve-student-achievement

Texas Education Agency (2018a). Raw score conversion table: Biology. Retrieved from https://tea.texas.gov/WorkArea/DownloadAsset.aspx?id=51539624167

Texas Education Agency (2018b). STAAR biology blueprint. *Texas Education Agency*. Retrieved from https://tea.texas.gov/WorkArea/DownloadAsset.aspx?id=2147487706

Texas Education Agency (2019). STAAR resources. *Texas Education Agency*. Retrieved from https://tea.texas.gov/student.assessment/staar/

Thijs, J. & Fleischmann, F. (2015). Student-teacher relationships and achievement goal orientations: Examining student perceptions in an ethnically diverse sample. *Elsevier, 42*(August 2015), 52-63. doi: https://doi.org/10.1016/j.lindif.2015.08.014

Thompson, J. (1998). *Discipline survival kit for the secondary teacher*. West Nyack, NJ: The Centre for Applied Research in Education.

Tillman, L. C. (2006). A nation at risk. *Encyclopedia of Educational Leadership and Administration*. Thousand Oaks, CA: Sage Publications, Inc. doi: http://dx.doi.org/10.4135/9781412939584.n388

Tizard, B., Blatchford, P., Burke, J., Farquhar, C., & Plewis, I. (2017). *Young children at school in the inner city*. New York, NY: Routledge.

Tucker, P. D. & Strong, J. H. (2005). *Linking teacher evaluation and student learning*. Alexandria, VA: Association for Supervision and Curriculum Development.

Tyner, A. (2018). When students take tests, are they actually trying? *Thomas B. Fordham Institute*. Retrieved from https://fordhaminstitute.org/national/commentary/when-students-take-tests-are-they-actually-trying

Ultanir, E. (2012). An epistemological glance at the constructive approach: Constructivist learning in Dewey, Piaget, and Montessori. *International Journal of Instruction, 5*(2), 1308-1470.

United States Congress, Office of Technology Assessment (1992). *Testing in American schools: Asking the right questions*. Washington, DC: U.S. Government Printing Office.

United States Department of Education (2004). New no child left behind flexibility: Highly qualified teachers. Retrieved from https://www2.ed.gov/nclb/methods/teachers/hqtflexibility.html

United States Department of Education (2016). Overview of proposed regulations: Assessments. Retrieved from https://www2.ed.gov/policy/elsec/leg/essa/nprmassessmentppt07192016.pdf

United States Department of Education Office of Innovation and Improvement (2004). *Alternative routes to teacher licensure.* Jessup, MD: United States Department of Education.

United States Department of Labor Employment and Training Administration (1999). Understanding the test-quality concepts of reliability and validity. Retrieved from https://www.hr-guide.com/Testing_and_Assessment/Reliability_and_Validity.htm

United States National Commission on Excellence in Education (1983). *A nation at risk: The imperative for educational reform: A report to the Nation and the secretary of education,* *United States Department of Education.* Washington, DC: U.S. Government Printing Office.

University of Missouri-Columbia (2018). Students taught by highly qualified teachers more likely to obtain bachelor's degree: Schools with more teachers who majored in their teaching subject are more likely to have students succeed both short and long term. *Science Daily*. Retrieved from www.sciencedaily.com/releases/2018/05/180522114820.htm

University of the State of New York – New York State Education Department (2009). How are Regents exams scored? *Office of State Assessment*. Retrieved from http://www.p12.nysed.gov/assessment/concht/scoring-regents.html

University of the State of New York – New York State Education Department (2018). New York state education department test development process. *Office of State Assessment.* Retrieved from http://www.p12.nysed.gov/assessment/teacher/testdevprocess.html

Urbandale Community School District (2019). 627—Instructional materials: Selection, inspection, and reconsideration. Retrieved from http://www.urbandaleschools.com/policy/article-600-educational-program/627-instructional-materials-selection-inspection-and-reconsideration/

US News (2018). Best high school rankings. *US News.* Retrieved from [URL hidden to protect identity of educational institution.]

Viadero, D. (2008). Student engagement found to rise as class size falls. *Education Week*. Retrieved from https://www.edweek.org/ew/articles/2008/03/25/29aerasize_web.h27.html

Virginia Department of Education (2016). Virginia standards of learning (SOL) tests cut scores as adopted by the Virginia Board of Education. Retrieved from http://www.cityschools.com/qclick/files/2015/05/sol-cut-scores.pdf

Virginia Department of Education (2019). Standards of learning (SOL) and testing. Retrieved from http://www.doe.virginia.gov/testing/index.shtml

Virginia Department of Education (n.d.). Test blueprint: Chemistry. Retrieved from http://www.doe.virginia.gov/testing/sol/blueprints/science_blueprints/2010/2010_blueprint_chemistry.pdf

Ware, F. (2002). Black teachers' perceptions of their professional roles. In J. J. Irvine (Ed.), *In search of wholeness: African-American teachers and their culturally specific classroom practices.* New York, NY: Palgrave.

Westrick, P. A., Le, H., Robbins, S. B., Radunzel, J. M., & Schmidt, F. L. (2015). College performance and retention: A meta-analysis of the predictive validities of ACT® scores, high school grades, and SES. *Educational Assessment*, *20*(1), 23-45.

Whelan, D. L. (2008). Teacher absenteeism affects student achievement. *School Library Journal,1*(10), 1-10.

Wiggins, G. P. & McTighe, J. (2005). *Understanding by design* (2nd ed.) Alexandria, VA: ASCD.

Wilcox, D. R. & Samaras, A. P. (2009). Examining our career switching teachers' first year of teaching: Implications for alternative teacher education program design. *Teacher Education Quarterly 36*(4), 173-191.

Wiliam, D. (2018). *Embedded formative assessment* (2nd ed.). Bloomington, IN: Solution Tree Press.

Williams, A. (2008). What is the difference between criterion-referenced and norm-referenced testing? *LetsGoLearn: Frequently Asked Quetions*. Retrieved from http://www.letsgolearn.com/faqs/comments/what_is_the_difference_between_criterion_referenced_and_norm_referenced_tes/

WiseWire (2016). What does it mean to “unpack” a standard? Retrieved from https://www.wisewire.com/wp-content/uploads/items/54051/pdf_cta_LMD_0022.pdf

Wolf, L. F. & Smith, J. K. (1995). The consequence of consequence: Motivation, anxiety, and test performance. *Applied Measurement in Education, 8*(3), 227-242.

Yin, R. K. (2014). *Case study research: Design and methods (5th ed.).* Thousand Oaks, CA: SAGE Publications, Inc.

Zhang, L. (2016). Is inquiry-based science teaching worth the effort? *Science & Education, 25*(7-8), p. 897-915. doi:10.1007/s11191-016-9856-0

APPENDIX A

LIBERTY UNIVERSITY
INSTITUTIONAL REVIEW BOARD

October 15, 2019

Suzanna Brawn
IRB Exemption 3946.101519: Improving Students' Scores on the Maine Educational Assessment for Science

Dear Suzanna Brawn,

The Liberty University Institutional Review Board has reviewed your application in accordance with the Office for Human Research Protections (OHRP) and Food and Drug Administration (FDA) regulations and finds your study to be exempt from further IRB review. This means you may begin your research with the data safeguarding methods mentioned in your approved application, and no further IRB oversight is required.

Your study falls under exemption category 46.101(b)(2), which identifies specific situations in which human participants research is exempt from the policy set forth in 45 CFR 46:101(b):

> (2) Research that only includes interactions involving educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior (including visual or auditory recording) if . . . the following criteria is met:
>
> > (iii) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and an IRB conducts a limited IRB review to make the determination required by §46.111(a)(7).

Please note that this exemption only applies to your current research application, and any changes to your protocol must be reported to the Liberty IRB for verification of continued exemption status. You may report these changes by submitting a change in protocol form or a new application to the IRB and referencing the above IRB Exemption number.

If you have any questions about this exemption or need assistance in determining whether possible changes to your protocol would change your exemption status, please email us at irb@liberty.edu.

Sincerely,

[redacted]

Administrative Chair of Institutional Research
Research Ethics Office

LIBERTY
UNIVERSITY.
Liberty University | Training Champions for Christ since 1971

APPENDIX B

[redacted]

[redacted]
Superintendent

[redacted]
Assistant Superintendent

[redacted]
Business Manager

[redacted]

[redacted]
Director of Special Services

[redacted]
Director of Adult and Community Education

[redacted]

16 August 2019

Suzanna Brawn
Researcher/Doctoral Candidate
Liberty University
1971 University Way
Lynchburg, VA 24502

Dear Suzanna Brawn

After careful review of your research proposal entitled *Improving Students' Scores on the Maine Educational Assessment for Science*, I have decided to grant you permission to conduct your study at [redacted] in [redacted]

Check the following boxes, as applicable:

[X] The requested data WILL BE STRIPPED of all identifying information before it is provided to the researcher.

Sincerely,

[redacted]
Maine School Administrative District 75

[redacted]

August 16, 2019

Suzanna Brawn
Researcher/Doctoral Candidate
Liberty University
1971 University Way
Lynchburg, VA 24502

Dear Suzanna Brawn,

After careful review of your research proposal entitled *Improving Students' Scores on the Maine Educational Assessment for Science*, I have decided to grant you permission to conduct your study at [redacted]

Check the following boxes, as applicable:

X The requested data WILL BE STRIPPED of all identifying information before it is provided to the researcher.

X I am requesting a copy of the results upon study completion and/or publication.

Sincerely,

[redacted]

APPENDIX C

The Liberty University Institutional
Review Board has approved
this document for use from
10/15/2019 to --
Protocol # 3946.101519

CONSENT FORM

Improving Students' Scores on the Maine Educational Assessment for Science
Suzanna Brawn
Liberty University
School of Education

You are invited to be in a research study regarding the improvement of students' scores on the MEA for Science. You were selected as a possible participant because you have been employed at Mt. Ararat High School for three years or more, and you are a science teacher, special educator, or administrator who is familiar with the MEA for Science and science instruction at the school. Please read this form and ask any questions you may have before agreeing to be in the study.

Suzanna Brawn, a doctoral candidate in the School of Education at Liberty University, is conducting this study.

Background Information: The purpose of this study is to solve the problem of low test scores on the MEA for Science and to formulate a solution to address the problem, answering the research question "How can the problem of low test scores on the Maine Educational Assessment for Science be solved?"

Procedures: If you agree to be in this study, I would ask you to do the following things:

1. Participate in an interview. The interview will ask questions appropriate to each participant's responses on the survey. Follow-up questions during the interview will be developed on the spot. Interviews will last approximately 60 minutes and will be recorded for transcription purposes.
2. Complete a survey. This survey includes 15 Likert-scale questions regarding various details surrounding the MEA for Science. This task should take approximately 30 minutes to complete.

Risks: The risks involved in this study are minimal, which means they are equal to the risks you would encounter in everyday life.

Benefits: Participants should not expect to receive a direct benefit from taking part in this study.

Compensation: Participants will not be compensated for participating in this study.

Confidentiality: The records of this study will be kept private. In any sort of report I might publish, I will not include any information that will make it possible to identify a subject. Research records will be stored securely, and only the researcher will have access to the records. I may share the data I collect from you for use in future research studies or with other researchers; if I share the data that I collect about you, I will remove any information that could identify you, if applicable, before I share the data.

- Participants will be assigned a pseudonym. I will conduct the interviews in a location where others will not easily overhear the conversation.
- Data will be stored on a password-protected computer in password-protected files and may be used in future presentations. After three years, all electronic files will be deleted.
- Interviews will be recorded and transcribed. Recordings will be stored on a password-protected computer as password-protected files for three years and then erased. Only the researcher will have access to these recordings.

Voluntary Nature of the Study: Participation in this study is voluntary. Your decision whether or not to participate will not affect your current or future relations with Liberty University. If you decide to participate, you are free to not answer any question or withdraw at any time without affecting those relationships.

How to Withdraw from the Study: If you choose to withdraw from the study, please contact the researcher at the email address/phone number included in the next paragraph. Should you choose to withdraw, data collected from you, will be destroyed immediately and will not be included in this study.

Contacts and Questions: The researcher conducting this study is Suzanna Brawn. You may ask any questions you have now. If you have questions later, **you are encouraged** to contact her at sbrawn@liberty.edu. You may also contact the researcher's faculty chair, Dr. Russell Claxton, at rlclaxton@liberty.edu.

If you have any questions or concerns regarding this study and would like to talk to someone other than the researcher, **you are encouraged** to contact the Institutional Review Board, 1971 University Blvd., Green Hall Ste. 2845, Lynchburg, VA 24515 or email at irb@liberty.edu.

Please notify the researcher if you would like a copy of this information for your records.

Statement of Consent: I have read and understood the above information. I have asked questions and have received answers. I consent to participate in the study.

☐ The researcher has my permission to audio record me as part of my participation in this study.

____

Signature of Participant Date

____

Signature of Investigator Date

The Liberty University Institutional
Review Board has approved
this document for use from
10/15/2019 to --
Protocol # 3946.101519

CONSENT FORM

Improving Students' Scores on the Maine Educational Assessment for Science
Suzanna Brawn
Liberty University
School of Education

You are invited to be in a research study regarding the improvement of students' scores on the MEA for Science. You were selected as a possible participant because you are a science teacher, special educator, or administrator at Mt. Ararat High School who is familiar with the MEA for Science and science instruction at the school. Please read this form and ask any questions you may have before agreeing to be in the study.

Suzanna Brawn, a doctoral candidate in the School of Education at Liberty University, is conducting this study.

Background Information: The purpose of this study is to solve the problem of low test scores on the MEA for Science and to formulate a solution to address the problem, answering the research question "How can the problem of low test scores on the Maine Educational Assessment for Science be solved?"

Procedures: If you agree to be in this study, I would ask you to do the following things:

1. Complete a survey. This survey includes 15 Likert-scale questions regarding various details surrounding the MEA for Science. This task should take approximately 30 minutes to complete.

Risks: The risks involved in this study are minimal, which means they are equal to the risks you would encounter in everyday life.

Benefits: Participants should not expect to receive a direct benefit from taking part in this study.

Compensation: Participants will not be compensated for participating in this study.

Confidentiality: The records of this study will be kept private. In any sort of report I might publish, I will not include any information that will make it possible to identify a subject. Research records will be stored securely, and only the researcher will have access to the records. I may share the data I collect from you for use in future research studies or with other researchers; if I share the data that I collect about you, I will remove any information that could identify you, if applicable, before I share the data.

- Data will be stored on a password-protected computer in password-protected files and may be used in future presentations. After three years, all electronic files will be deleted.

Voluntary Nature of the Study: Participation in this study is voluntary. Your decision whether or not to participate will not affect your current or future relations with Liberty University. If you decide to participate, you are free to not answer any question or withdraw at any time without affecting those relationships.

How to Withdraw from the Study: If you choose to withdraw from the study, please contact the researcher at the email address/phone number included in the next paragraph. Should you choose to withdraw, data collected from you, will be destroyed immediately and will not be included in this study.

Contacts and Questions: The researcher conducting this study is Suzanna Brawn. You may ask any questions you have now. If you have questions later, **you are encouraged** to contact her at sbrawn@liberty.edu. You may also contact the researcher's faculty chair, Dr. Russell Claxton, at rlclaxton@liberty.edu.

If you have any questions or concerns regarding this study and would like to talk to someone other than the researcher, **you are encouraged** to contact the Institutional Review Board, 1971 University Blvd., Green Hall Ste. 2845, Lynchburg, VA 24515 or email at irb@liberty.edu.

Please notify the researcher if you would like a copy of this information for your records.

Statement of Consent: I have read and understood the above information. I have asked questions and have received answers. I consent to participate in the study.

☐ The researcher has my permission to audio record me as part of my participation in this study.

____

Signature of Participant Date

____

Signature of Investigator Date

APPENDIX D

Improving Students' Scores on the Maine Educational Assessment for Science
An Applied Research Qualitative Interview

1. What skills are assessed with the Maine Educational Assessment (MEA) for Science?
2. How are these skills incorporated through students' 9-12 curriculum map?
3. How do the day-to-day instructional strategies in your classroom prepare students for the assessment?
4. How do the instructional resources available, including textbooks and other in-class materials, as well as electronic tools, support daily instructional practices?
5. How do students demonstrate mastery of these skills and concepts in your class?
6. The MEA is administered during a student's 11th grade year of high school and is a cumulative exam. How are skills needed to be successful on the MEA learned during the 9th and 10th grade years reflected in the curriculum map?
7. What specific preparation is offered to students prior to taking the MEA?
8. For the past three years, our students have scored in the bottom half (and sometimes the bottom quarter) of all schools in the state. What factors do you believe contribute to such low test scores?
9. When the assessment results are shared with the department, what do you think should be the next steps for teachers?
10. What strategies have been explored to increase students' performance on the test?
11. How were these strategies decided upon?
12. How were they implemented?

13. How do you think students feel about taking this assessment and what evidence do you have to support your opinion?

the assessment, they will likely be indifferent towards their scores (Tyner, 2018).

14. In what ways do you feel that teacher effectiveness impacts students' performance on the MEA for Science?
15. What factors positively impact students' performance on this test?
16. What other thoughts or feedback about how students' scores on the MEA could be increased at our school could you add?

APPENDIX E

Improving Students' Scores on the Maine Educational Assessment for Science
An Applied Research Quantitative Survey

Demographic Research Questions
Instructions: Mark one answer for each demographic question.

1. Which category best describes your age in years?

   21-29 | 30-39 | 40-49 | 50-59 | 60 or older

2. What is your race?

   White | Black or African-American | Asian | Native Hawaiian/Other Pacific Islander | Two or More Races | Other

3. What is your gender?

   Male | Female

4. What is the highest educational degree you have received?

   Less than High School Diploma or Equivalent (GED) | High School Diploma or Equivalent (GED) | Associate's Degree | Bachelor's Degree | Graduate Degree | Doctorate

5. What grades do you teach? (You may select more than one).

   9th Grade | 10th Grade | 11th Grade | 12th Grade

Content Research Questions
Instructions: Choose the answer that best describes your opinion.

1. Before teaching a new concept, the associated Next Generation Science Standards (NGSS) are intentionally built into instruction.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

2. Classroom instruction is delivered with the same rigor specified in the NGSS standards.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

3. All students are assessed in the same manner across grade levels.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

4. Time is provided during contract hours for collaboration between teachers of the *same* courses.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

5. Data collected from both assessments are used to plan future instruction.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

6. Content taught in each course is clearly communicated throughout the department.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

7. Professional development focuses on new instructional strategies has been offered to science teachers.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

8. Teachers regularly participate in meaningful data analysis conversations.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

9. Students are excited and ready to learn in my class each day.

1	2	3	4	5
Never	Rarely	Sometimes	Often	Always

10. The science department at Hartman High School is made of teachers that are highly-qualified educators.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

11. On-going support is provided to new teachers.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

12. Professional development that serves the instructional needs of teachers is provided on a regular basis.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

13. Teachers in the science department at Hartman High School have high attendance rates.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

14. When teachers are absent from school, high-quality substitute teachers are hired.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

15. Class sizes in the science department at Hartman High School are conducive to student learning.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

APPENDIX F

Protocol for Implementing Professional Learning Communities & Content-based Literacy

- Establish common planning for content-related teachers (Spring)
- Create buy-in among key teachers in the department (Summer)
- Professional Development (August)
- Implement PLC/TLC (September-March)
    - Planning standards-based
    - Rigor
    - Scaffolding
    - Differentiation
    - Common lessons/assessments
    - Data analysis
    - Re-teaching/new planning
    - Spiral review
- Targeted remediation based on data analysis from PLCs (April)
- Continued and on-going professional development (year-long)