

Student Perceptions Towards Mandated Assessments

**By
© 2019
Jennifer L. Woolever**

**Submitted to the graduate degree program in Educational Leadership & Policy Studies
and the Graduate Faculty of the University of Kansas in partial fulfillment of the
requirements for the degree of Doctor of Education (Ed.D).**

Chair: Rick Ginsberg, Ph.D.

Thomas A. DeLuca, Ph.D.

Argun Saatcioglu, Ph.D.

Kelli R. Thomas, Ph.D.

Jennifer Ng, Ph.D.

Date Defended: 15 May 2019

The dissertation committee for Jennifer L. Woolever certifies that this is the approved version of the following dissertation:

Student Perceptions Towards Mandated Assessments

Chair: Rick Ginsberg, Ph.D.

Date Approved: 15 May 2019

Abstract

This study examined student perceptions towards testing mandates. The specific purpose was to gain an understanding into the perceptions students have towards testing mandates and determine their perceived value in the areas of (1) Improvement, (2) External attribution, (3) Affective benefits, and (4) Irrelevance. In analyzing the perceptions of students, this study also focused on particular demographics that were identified in the literature as significant predictors of student performance (i.e. gender, race/ethnicity, primary language, parent education, GPA and plans after high school) to determine if these variables not only affect student performance, but also perceptions towards mandated tests. This study investigated 360 ninth and tenth grade students from five high schools within one suburban school district. A demographic questionnaire, Brown's student conceptions of assessment survey (SCoA-VI), and three open-ended response questions were utilized in the study. The findings showed 9th and 10th grade students disagreed about the importance these testing mandates had towards their "Improvement", "External Attribution" and "Affective Benefits". High school students agreed mandated assessments were irrelevant. In addition to the SCoA-VI and open-ended question, certain demographic characteristics were found to influence student perceptions towards mandated assessments. Females were more likely to view the assessments as unfair, ignored, not enjoyable or helpful and not a good measure for the quality of the school and/or their future employment when compared to male students. English speaking students were more likely to view mandated assessments as irrelevant when compared to other languages examined. Also, white students, those planning to attend college or technical training, and those with higher levels of parent/guardian education tended to disagree that mandated assessments were enjoyable

and/or helpful to classmates. In conclusion, students had a dislike for testing mandates. Despite the negative feelings, students offered insights and suggestions for creating a more conducive measure that is relevant, reliable and offers opportunities for improvement.

Acknowledgements

I would like to thank my doctoral advisor, Dr. Rick Ginsberg, for his ongoing encouragement and leadership. I would also like to recognize the dissertation committee for their feedback and guidance throughout the dissertation process. Also, I would like to thank Dr. Dan Gruman, for offering his advice and math expertise to help me analyze my quantitative data. Most importantly, I would like to thank my mom, Dr. Kathy Fletcher. She would not let me quit and continued to advise my family members to stop asking, “Is it done yet?” My family’s sense of humor at times can be annoying, but also encouraging. They always mean well and I could not do it without them. Thank you.

Table of Contents

Chapter 1 Introduction	3
Background	3
Problem Statement.....	6
Research Questions	7
Significance of the Study.....	7
Organization of the Study	10
 Chapter 2 Literature Review	 11
Standardized Test in America	12
An Overview of the Mandated Assessment Movement.....	13
Perceptions of Mandated Assessments.....	18
Correlates in Research that Affect Student Performance	24
Student Perceptions Towards Mandated Assessments	27
Summary	33
 Chapter 3 Methodology	 35
Introduction	35
Research Design.....	35
Setting and Sample	36
Instruments.....	39
Procedure.....	42
Data Analysis	43
Research Question 1	44
Research Question 2	45
Limitations.....	46
 Chapter 4 Results	 48
Qualitative Data: Open-Ended Questions.....	55
Students' Perceived Benefits from Testing Mandates	55
Students' Perceived Drawbacks from Testing Mandates	60
Student Opinions Towards Mandated Assessments	65
Research Question 1 Summary.....	69
Research Question 2 Summary.....	76
 Chapter 5 Conclusions	 78
Summary of the Findings	78
Conclusions	86
Implications for Future Research.....	93
 Appendix	 97
A: Permission to Utilize SCoA-VI	97
B: IRB Approval.....	99
C: Parent/Guardian Email	100
D: Demographic Information	101
E: SCoA-VI.....	102
F: Open-Ended Questions	107

Bibliography 108

Chapter 1

Introduction

Background

Over the past several decades educational policies have increasingly mandated student achievement tests at the national, state and district level. Mandated assessments have become an annual ritual for students throughout their elementary and secondary education. Many states and school districts mandate testing programs to provide a multitude of ways for their schools to accurately assess student achievement and to have the ability to compare results with other students, schools, teachers and administrators. Several states have taken further steps towards increasing accountability by endorsing value-added models as a means to evaluate an educator's contribution to student learning within a given school year and/or time frame. Having accountability measures tied to assessment results establish achievement goals or standards for students, teachers, and administrators. Assessment scores can bring public praise, accolades, financial rewards, embarrassment and/or heavy sanctions. In the United States, assessment mandates assigned at the state and district level vary greatly on the magnitude these stakes and accountability measures hold for those involved in the testing process. Uncertainty and value of the assessments has caused testing mandates to be a contentious topic within educational research (Adamson, Astrand & Darling-Hammond, 2016; Ravitch, 2010; Nichols & Berliner, 2007)

Proponents for testing mandates highlight several benefits students receive from these assessments based on the perceptions of their teachers, school personnel, district leaders and policymakers. Reeves (2009) found that not only do formative assessments provide information

to students and teachers about learning, they can also be very rewarding by providing benchmarks for students to sustain their learning. Several other studies emphasized the positive shift in instruction due to the implementation of these mandates linking school incentives to results. Much of the changes identified from these various studies were aimed at improving student achievement by increasing total instruction time, curriculum alignment with learning standards and adding more professional development for teachers (Ferguson, Kober, & Stark-Rentner, 2017; White and Rosenbaum, 2008; Rouse et. al, 2007; Hamilton, Stecher & Klein 2002). These studies suggest the positive impact testing mandates have had on teachers in their instruction, professional growth and development, and on the overall learning objectives achieved in the classroom. Much of the literature supporting common assessments focused on the idea that having these mandates provides a consistent foundation for all stakeholders to measure and assess student achievement (Brown 2010; Wolf, 2007).

On the other hand, opposing perspectives found in the research suggest a more negative outlook towards these assessment mandates in the areas of instructional practices, content and the increase in time devoted for preparing and assessing students (Davis, 2011; Wagner, 2008, Amrein & Berliner, 2002a). Amrein & Berliner (2002a) noted that these practices were more common near testing dates and among teachers of poor, minority students. In addition, Cunningham and Sanzo (2002), for example, found that high-stakes testing negatively impacted creative and effective teachers by leading them to cram for tests rather than focusing on instruction. Several other studies supported similar effects on how instruction was adjusted based on the assessment criteria which caused other non-tested subjects to become limited or obsolete (Davis, 2011; Rothstein et al., 2008; Wagner, 2008; Au, 2007; Koretz and Hamilton, 2003; Amrein & Berliner, 2002a). Many of these studies that acknowledged the downside of testing

mandates were conducted with on-site faculty at the building level who oversee instruction and curriculum.

With a large amount of literature surrounding the perceived benefits and drawbacks of testing mandates, little research has actually been examined on the perceptions of those most directly impacted by mandated tests- the students who take the assessments. Student viewpoint concerning assessment mandates has been understudied and/or overlooked. The information gained from student perceptions could not only broaden the current research, but also provide additional clarity on the rationale for why these assessments are mandated at the district, state, and federal level. Understanding how students perceive these experiences with mandated assessments could be useful when making decisions about refining educational policies, strengthening teaching practices and providing better ways for staff and parents to support student growth and performance. Understanding how students feel towards these mandates could offer legislators, administrators, teachers and parents with more purposeful ways to improve testing experiences so that it is more significant to students regarding their learning and future or it could point to the need for different kinds of assessment/measurement practices. The purpose of this study is to gain an awareness into the perceptions of students to determine their perceived value of the assessment mandates in the areas of (1) Improvement, (2) External attribution, (3) Affective benefits, and (4) Irrelevance.

Triplett and Barksdale (2005) found that student perceptions of high-stakes testing were largely disregarded due to the prevailing attitude among researchers that students would naturally report a negative bias toward all forms of accountability assessments. Several studies have compared the beliefs and attitudes of administrators, teachers, pre-service teachers, counselors, and parents (Dworkin & Tobe, 2014; Eslinger, 2014; Husband, 2014; Dawson, 2013). However,

very few studies have considered students' perceptions of state mandated accountability assessments (Dutro & Selland, 2012; Nelson, McMahan, & Torres, 2012; Heilig, 2011; Triplett & Barksdale, 2005). Given the use of these tests to measure the quality and effectiveness of individual schools and districts, it is critical to consider all possible factors that might contribute to a student's success or lack of success on these tests (Horn, 2003). It is important to understand student perceptions as they are the objects for these testing mandates.

Student performance on standardized tests is a key component of federal, state and district mandates and is ultimately a means to determine the quality and success of schools. Raising test scores is an imperative goal for most schools and districts. A wealth of research has shown how demographic variables (i.e. gender, race/ethnicity, income, and English language ability) are a significant predictor that affects student performance (Olszewski-Kubilius and Lee, 2011; Darling-Hammond, et. al (2009); Davis-Kean (2005); Abedi, 2002; Codero-Guzman, 2001; Hartman, 2001; Alexander, Entwisle & Bedinger (1994)). In analyzing the perceptions of students, this study focused on particular demographics that were identified as significant predictors of student performance (i.e. gender, race/ethnicity, primary language, parent education, GPA and plans after high school) to determine if these variables not only affect student performance, but also their perceptions regarding mandated tests.

Problem Statement

The study is designed to investigate student perceptions (beliefs) regarding mandated assessments and whether key demographic variables (gender, race/ethnicity, primary language, parent education, GPA and plans after high school) influenced student perceptions towards mandated assessments. Investigating these demographic variables allowed for connections to

existing research on the academic achievement differences associated with these factors.

Furthermore, the study investigated the relationship these demographics have on achievement and how it corresponds to the students' perception towards testing mandates.

Research Questions

R1: What are high school students' perceptions toward mandated testing on improvement, external attribution, affective benefits and irrelevance as measured by Brown's SCoA-VI survey?

A) Do high school students feel testing mandates improve teaching and student learning (Improvement)?

B) Do high school students feel mandated assessments measure the school's quality and predict the student's future in education and employment (External attribution)

C) Do high school students feel these assessments are beneficial to themselves and classmates (Affective benefits)?

E) Do high school students feel mandated assessments are unfair and ignored (Irrelevance)?

R2: Do high school student perceptions toward mandated assessments differ based on selected demographic characteristics including gender, race, primary language, parent education, GPA, and plans after high school?

Significance of the Study

This study is significant because it expands the research on mandated assessments, which has been a theme for educational policies throughout the past several decades (PDK International, 2017). Standardized assessments continue to be at the forefront for policymakers

as an efficient way to measure schools, teachers and student performance. As testing mandates remain a heavily emphasized topic, this study is important because it examines an area that has not been given much credit and/or recognition, which is student perceptions towards testing mandates. Understanding how students perceive testing mandates makes this study significant for three major reasons: it could (1) improve educational policy, (2) strengthen practice for educators, and (3) support student growth and performance. The information gained from this study could be impactful to several stakeholders that are responsible for implementing and supervising these testing mandates.

First, the information gained could be invaluable for federal/state legislators and district administrators as they continue to develop or invest in educational policies supporting standardized assessment mandates. Elwood and Lundy (2010), for example, noted that if given the opportunity students could provide information about the implementation of educational policies, the benefits assessments have towards their learning, and further examination of the systems fairness. Heritage (2013) suggested that in order for this process to be essential to twenty-first century learning, students should be involved in the inquiry and action. Therefore, it would be imperative to involve students in the development of the policies so that they have a role and an opportunity to contribute their input. The information obtained could create greater student buy-in that increases motivation and better understanding of the exams purpose.

Second, understanding student perceptions could better equip school personnel and parents on ways to prepare, handle and advise during testing circumstances for students. Understanding the perceptions of students towards these tests could provide guidance on different types of professional development for teachers needed at the building and classroom levels. Fletcher (2005) and Mitra (2008), for example, emphasized encouraging “student voice”

and the need to understand their perspective in order to effectively change/influence the practice of educators. Understanding how students perceive mandated assessments could provide more direction on how schools and teachers support student learning, the testing process, and further professional development.

Finally, knowing how students perceive these testing mandates could help provide guidelines for faculty and parents to offer improved support and communication to increase student motivation when completing the assessments. Eklof (2007), for example, discussed how student motivation can have significant weight on student performance and assessment results. It is important for strengthening our understanding of student perceptions towards mandated assessments, as further insight on what is significant and/or irrelevant to students can potentially provide staff and parents with better direction on how to offer more meaningful incentives and/or motivational support for students.

As policies continue to emphasize mandated assessments, it is imperative all stakeholders' perspectives be explored. The significance of the study is that it offers insight into students perceptions towards testing mandates, which is a voice that has not been often heard within the current literature. Furthermore, the study provides recommendations from students for testing changes so that the process could be more meaningful to them and serve as a better measure of the student's learning and academic performance. The study also provided an opportunity for students to explain and elaborate on their concerns, experiences, and perspectives towards testing mandates that have been implemented throughout their elementary and secondary education. Due to the amount of resources utilized with mandated testing, it is important that these exams measure accurately and are valuable to teachers and students.

Under the current testing circumstances, students in the study described less effort and motivation put forth because of the set-up, time and pressure surrounding these testing mandates. This study's findings can be very beneficial for school administrators, teachers, parents, policymakers, and state and federal government officials because it gives a glimpse of the students opinions about the tests. Students suggested several areas of improvement to make the testing process more applicable to learners, such as accuracy, test content, and better feedback to assist with future learning. In doing so, this study may lead to a positive change within the practices surrounding mandated assessments.

Organization of the Study

This dissertation is organized into five chapters. Chapter 1 introduces the study and includes the following sections: background, problem statement, research questions and significance of the study. Chapter 2 provides a critical review of the research and literature related to the study. Chapter 3 explains the methods and procedures used to conduct the study. Chapter 4 describes the findings of the study. Chapter 5 concludes with a review of the study and its findings, a discussion of those findings, conclusions, and implications of the study for educators, administrators, parents, and policymakers, as well as recommendations for future research.

Chapter 2

Literature Review

The review of literature will be guided by the research conducted on mandated assessments. According to the Glossary for Education Reform (2015), the term “assessment” refers to a variety of methods and/or tools that educators use to evaluate, measure, and document academic readiness, learning progress, skill acquisition, and educational needs of a student. While the term assessment can take a wide variety of forms in education, this study focuses on “mandated assessments,” those that are required and initiated by the federal, state government and/or school district.

The literature review will be divided into 5 sections. The first section will give an overview of the mandated assessment movement. In this section, the historical background will outline the process that has occurred to make mandated assessments so influential in our school systems today. The second section will discuss research regarding the perceptions of mandated assessments and the perceived effects that have occurred as a result of these testing mandates. The information presented in this section will provide a better picture for how different aspects of the educational setting have changed based on the assessment mandates as perceived by stakeholders. The third section will present information about correlates in research that affect student performance on testing mandates. This section will provide evidence about how certain contributing factors such as demographics, past performance, and student perception can affect student performance. The fourth section will provide research on student perceptions towards mandated assessments. This section will present information regarding students’ perceptions towards mandated assessments and how it affects their effort and motivation to perform. The

fifth section of the literature review will provide information on various instruments used in research to measure perceptions of assessment. The last section will offer more insight into past studies on the tools used to measure perceptions towards assessments involving various stakeholders and to promote the rationale for this study. The chapter ends with a summary of key findings and rationale for this study.

Standardized Tests in America

Standardized assessments have played a role in American education since the 1800s, however it was not until the 20th century before support for standardized tests became more widespread. As for the earliest record of standardized test, China was the first to use assessments as a means to access knowledge and assign government jobs (Fletcher, 2009). Similarly, the U.S. Army adopted achievement tests to identify potential officers from large groups of recruits during World War I. National leaders used these exams to identify academic and managerial talents (Edwards, 2006). The use of standardized tests quickly became more mainstream during World War II and the Cold War because of its efficient and affordable means to access troops (Edwards, 2006).

As for American education, standardized testing began to emerge in the 1800s, however it was not until the 20th century before it became more widespread (Stiggins, 2002). For example, a nonprofit group of universities and educational organizations known as the College Entrance Examination Board was established in the 1900s and began to explore standardized testing for higher education. The board sought to create an exam for northeastern elite universities so they could standardize student admissions (Fletcher, 2009).

Initially the standardized test consisted of essays and was not meant for widespread use. However further development and calibration was made by French psychologist Alfred Binet, who created a standardized test of intelligence, later named the Stanford-Binet Intelligence Test (Fletcher 2009). Then in 1926, the College Board founded the Scholastic Aptitude Test (SAT). The SAT became highly regarded for its accessibility and by the end of World War II, the test was being implemented by numerous universities. The SAT quickly became a standard rite of passage for college-bound high school seniors (Fletcher, 2009).

Furthermore, in 1959 Everett Lindquist presented the American College Test (ACT). As mandated assessments began to evolve, the accessibility for applicants and little cost to administer were key factors contributors to its popularity among the government, universities and other organizations. In the 1970s state testing became mainstream and by the 1980s America was assessing nationally (Stiggins, 2002).

An Overview of the Mandated Assessment Movement

It is well documented that mandated assessments have caused a fundamental change within schools (Claibron, 2009; McNeil, 2000). The mandated assessment movement for public schools started when Lyndon Johnson declared a “War on Poverty”. President Johnson, who believed that “full educational opportunity” should be “our first national goal”, signed the Elementary and Secondary Education Act (ESEA) into law (U.S. Department of Education, 1965). To help combat the effects of poverty, schools were provided additional federal funds. The ESEA offered a source of income, referred to as title funds, for schools to use on instructional resources, professional development and various educational programs supporting low-income students. As Title I began to emerge across the nation, little evidence existed as to

whether or not the achievement gap was narrowing between at-risk students and their peers. In order to statistically show improvement, Congress and the U.S. Department of Education knew further measures needed to be taken to ensure that Title I allocations and interventions were more effectively evaluated (Borman and D'Agostino, 2001). In 1969, three years after the implementation of Title I, the ESEA developed the National Assessment of Educational Progress (NAEP) to measure the impact funding had on our schools and students. The NAEP assessment was federally mandated and used to measure student achievement nationally. The assessment was given to a random sample of 4th, 8th, and 12th grade students. As states began collecting student data and allocating title funding, a culture quickly established surrounding standardized assessments. For many stakeholders, the federal mandate, NAEP, created the notion of what we refer to as now "testing stakes". A connection between funding resources and student achievement began to emerge. School funding was quickly being tied to results.

A landmark report known as "A Nation at Risk" alerted constituents and urged policymakers to support tests that included higher standards of academic achievement in lieu of the minimal competency tests that were currently utilized in public schools (National Commission on Excellence in Education, 1983). In the report, the commission stated, "minimum competency examinations fall short of what is needed, as the minimum tends to become the maximum, thus lowering educational standards for all" (p.20). Furthermore, the commission warned that, "the educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a Nation and as people" (p.5).

This report raised concern in the public because school systems were not being accountable, and students were failing. A Nation at Risk argued that Americans test scores were lagging behind their international competitors (Guthrie & Springer, 2004). These results included

NAEP, SAT scores, and data from the International Assessment of Education Achievement to reinforce the claim that the American school system was failing our children. As a result, state legislatures quickly began to revise their states' educational policies to promote more rigorous academic standards. *A Nation at Risk* created a sense of awareness about the lack of accountability occurring in schools and ultimately pressed those in leadership roles to set forth plans to prevent further decline.

By 1985, almost all states had initiated educational reforms that called for more rigorous graduation requirements, longer school days extending the year and requiring more homework and testing (Shepard, 2002). These changes were only the beginning for educational reforms, as an emphasis on outcomes quickly followed. In 1994, congress renewed ESEA to encourage states to meet high standards by establishing Goals 2000. Students in 4th, 8th, and 12th grades had to complete a competency test in English, math, science, history, and geography. These policies were to offer a sense of assurance to the public because school systems were now being held accountable for maintaining these high standards (Hurst et al., 2003). For many states, the high stakes testing era was established and policymakers were the driving force behind these accountability measures. As each policy began to unfold and intertwine with others, testing continued to play a strong role in on how federal funding would be distributed across states and eventually to schools.

In 2002, Congress reauthorized ESEA when President George W. Bush Jr. signed into law No Child Left Behind (NCLB). Schools were mandated to test ninety-five percent of their students. This policy also placed a cap on modified alternative assessments for at-risk and/or special education students. The defining characteristic of NCLB is the concept of Adequate Yearly Progress (AYP). The state established yearly benchmarks to measure student growth

through the observation of standardized tests scores in math and reading. In many states, high performing schools were awarded extra money and low performing schools were given assistance and a series of consequences over a four-year period if no improvement was evident. Schools not making AYP for two consecutive years would be classified as “failing” and would be required to offer tutoring and other supplemental services (Hurst et al., 2003). According to NCLB, if a Title I school failed to attain AYP, the school district was placed on an improvement plan and at that time parents had the ability to move their children to higher performing schools and/or the ability to receive supplemental educational services for their children. Additionally, schools that continued to display no improvement in meeting AYP would be at risk of losing title funding altogether. Like past policies, NCLB tied educational resources to test results and performance. The quantitative measures established by NCLB greatly increased pressure for many school administrators and educators because school districts relied on the additional federal funding that ESEA provided. NCLB put in place measures that exposed achievement gaps among underserved students and their peers, as well as prompted dialogue nationwide on the importance for improvement in education (Brenchley, 2015). The focus on accountability provided a means to monitor, measure and potentially provide a quality education for all children. Yet, challenges in the implementation process for NCLB became a challenge for many schools and districts.

In 2012, President Barrack Obama granted flexibility to states regarding specific requirements of NCLB in exchange for rigorous and comprehensive state-developed plans designed to close achievement gaps, increase equity, improve the quality of instruction, and increase outcomes for all students (Brenchley, 2015). The waivers permitted by the federal government allowed for flexibility as states and schools created a new design to increase student

performance for their own region. In 2015, President Obama signed the Every Student Succeeds Act (ESSA). The ESSA included provision that would ensure success for students and schools by upholding critical protections for disadvantage and high-needs students. A new accountability system encouraged states to work together on common academic standards that were built on college and career readiness (Kansas State Department of Education, 2016). The ESSA law offered flexibility to find the best local solutions, while still providing accountability measures to monitor student progress on their academic performance. At the state level that meant local leaders and school districts were now responsible for identifying and intervening when schools exhibited a large disparity in their numbers in regard to graduation rates, attendance, assessment results, etc... similar to past policies, testing was still a prominent part of the ESSA law.

According to the ESSA, statewide assessments are to be conducted annually in an effort to measure students' progress towards mastering common core standards. Like many state level plans, Kansas utilized a variety of assessments. At the state level, the Kansas Assessment Program (KAP) requires several mandated assessments yearly in various fields of study throughout a student's public schooling. In English Language Arts (ELA) and mathematics, students are administered exams during their 3rd through 8th and 10th grade school year. KAP alternates yearly science assessments for 5th, 8th and 11th grades with a History/Government assessment in 6th, 8th, and 11th grade. The Kansas English Language Proficiency Assessment (K-ELPA) assesses K through 10th grade students who are identified as English Language Learners (ELL) in the areas of listening, reading, writing, and speaking. Lastly, the Career Pathways Assessment (c-Pass) measures 9th through 12th grade students on their readiness in postsecondary education and/or entry into the workforce. The State Board of Education designed the assessment program to (1) measure specific claims related to the Kansas College and Career

Ready Standards (KCCRS), (2) provide information for calculating Annual Measurable Objectives (AMOs) and for state accreditation, (3) report individual student scores along with the student's performance level and (4) provide subscale and total scores that can be used with local assessment scores to assist in improving a building's or district's programs (The Kansas State Department of Education, 2016).

In conclusion, how the mandated assessment movement developed and progressed shows the magnitude at which policymakers and stakeholders sought for improvement in public schools by endorsing standardized accountability systems. Over the years, the level of control and decision making was pushed from heavy federal mandates to more local control within each state government and school district. However, the theme for accountability and routine assessments remained a constant factor throughout as a means for improving schools. As policies changed general accountability measures remained the same- to assist those schools and areas identified as low-performing. As testing mandates continue to be endorsed, it is important to reflect on how these routine assessments have transformed schools for better or worse.

Perceptions of Mandated Assessments

The idea of schools, teachers, and students being held accountable is a controversial topic in American education. Creating common rigorous academic standards and developing ways to measure how students perform on those standards created opportunities for conversation, comparison, critique and multiple consequences. Much of the debate focused on who (individuals and schools), how, and to what extent assessments should be used to hold accountable. In this section, perceptions of mandated assessments and the intended and unintended effects will be examined from the organization level (district and school) and the

classroom level.

Perceptions at the Organizational Level

Research has shown that districts and schools have responded to accountability in contrasting ways which has led to some intended and unintended results (Ravitch, 2010; Sauder and Espeland, 2008; Nichols and Berliner, 2007). First, from the district and/or school perspective, the intended or positive effects accountability has had on achievement included (1) a more consistent calibration across classrooms (Spillane et al., 2011; Hallett, 2010), (2) better alignment between curriculum frameworks and the content taught within classrooms (Spillane et al., 2011; Brown and Clift, 2010;), and (3) a perceived increase in the amount of effort and productivity displayed by teachers (Reback et al., 2011). Overall, the positive effects perceived from the district and school about standardized accountability systems was that it created consistency in the content, increased collaboration and professional dialog, and served as a motivator for staff. As a result, many school districts across the country began implementing routine mandated achievement test at the building level. These assessments assisted district and building staff in ways to monitor student growth, instruction, teacher performance and school improvement in a smaller more personal setting. Several studies showed that by creating achievement targets, it created a more positive effect on increasing student test performance (Dee & Jacob, 2011; Hanushek & Raymond, 2005; Carnoy & Loeb, 2002). This method has been described by Amrein-Beardsley (2008), as Value-Added Models (VAMs). These assessments are typically used to measure a student's learning over time, which can show the impact a teacher has as a source of variance in a student's outcome. This value-added philosophy holds schools and teachers accountable for the learning gains of the students they serve.

As for the unintended effects perceived at the district and school level, studies showed unnecessary reclassification or exemption of certain students from taking the tests (Jennings & Beveridge, 2009; Cullen and Reback, 2006; Figlio and Getzer, 2006; Jacob, 2005). In addition, Jacob (2005) found an increase in special education placement and student retention in grades prior to the tested grade level in Chicago Public Schools.

Regarding intended effects, Berliner & Nicholos (2007) examined the impact high-stakes testing programs had on education since the No Child Left Behind act. These researchers investigated the accounts of teachers, administrators, and journalists. Their findings showed various forms of cheating by students, teachers, and administrators, data manipulation from the state and district, and excluding less-performing students from the test-taking pool. Berliner & Nicholos (2007), further described the impact assessments had on teacher and student morale. Their analysis drew from Campbell's law, the notion that when social consequences are connected to a quantitative value (test scores), the more likely the tests will be corrupted, along with the social processes it was intended to monitor (Campbell, 1979). Their analyses concluded that testing has these effects.

Furthermore, other unintended outcomes perceived by administrators and teachers were in regard to the resources and support provided to students. Au (2007), for example, analyzed 49 qualitative studies to determine how high-stakes testing affected curriculum. The primary effect found was curriculum being narrowed to tested subjects and content fragmented into test related criteria. Additional studies (Rothstein et al., 2008; Black & William, 2004; Koretz and Hamilton, 2003; Stecher et al., 2000) found similar findings with less emphasis on non-tested subjects, as well as content areas becoming limited due to the demands set forth by testing mandates.

Another drawback described by researchers (Hong & Youngs, 2008) is that due to the pressure caused by mandated assessments, educators had an increased focus on students described as "bubble kids", those who are approaching or near passing the test. Similar studies (Deming et al., 2013; Neal & Schanzenbach, 2010) conducted on administrators and teachers also found that an increase in focus and time was spent on marginal students thus taking away focus and time from students who scored at both ends of the achievement distribution.

After reviewing the information, it is clear that several issues need to be considered when examining mandated assessments. Proponents maintain that testing mandates are needed for all students to have a fair and equal opportunity to receive a quality education (Reback et al., 2011; U.S. Department of Education, 2010; Supovitz, 2009). The opposition, however, maintains that the implementation of testing mandates has created a culture of control and anxiety, which limits the access of students and teachers. Given the proponents on either side of the debate, it is important to note that no proven results have led to the conclusion that testing mandates are an effective process of increased learning (Jones & Egley, 2006). It is evident that the outcomes surrounding testing mandates can vary and how districts and schools respond to these measures is what ultimately leads to both intended and unintended results (Ravitch, 2010; Sauder and Espeland, 2008; Nichols and Berliner, 2007).

Perceptions at the Classroom Level

A wealth of research has focused on teacher perception of mandated assessments and accountability. Knowing teachers are a major contributing factor to student achievement, it is important to understand how teachers perceive the purpose for these testing mandates and how it has changed their curriculum and instruction. Harris, Harnett, & Brown (2009) conducted teacher interviews and found many negative thoughts about assessments when teachers did not

believe that the assessments were effective in aiding improvement in teacher and student learning. It is important to note that Darling- Hammond & Rustique-Forrester (2005) advised that just by adding more assessments to the teaching and learning process does not guarantee or provide educators with the tools they need to gain a better understanding of student learning.

Similar results were presented in a survey conducted by the National Board on Educational Testing and Public Policy (Pedulla et al., 2003). Educators in the survey reported that their state-testing program contradicted their own notions of good educational practices. Both groups of teachers, 76% of high stakes and 63% of moderate to low-stakes settings, felt the implementation of state testing programs had an adverse effect on their teaching style causing a negative impact on the overall quality of instruction. In addition, pressure to raise test scores was experienced in both low and high-stakes settings from superintendents and principals. Lastly, a substantial number of teachers in both high-stakes (85%) and moderate to low-stakes (67%) settings reported teaching test-taking skills to prepare students for the state test. These findings among several, show that by placing accountability measures on schools and/or teachers it can create a great deal of pressure for educators to improve student performance on tests, as well as pushing them to use less effective teaching habits that are geared specifically toward improving test scores (Pedulla et al., 2003; Hoffman, Assaf, and Paris, 2001; Smith et al., 1991).

The Glossary of Education Reform describes high stakes as test scores that are used to determine accolades, advancements, punishments, or compensation for schools, students and employees (Abbott, 2014). These stakes can range from funding reductions and negative publicity to awards, public celebration, grade promotion and salary increases. As for low stakes tests, these exams are used to measure academic achievement, identify learning problems, and/or

inform instructional practices. In general, what distinguishes high stakes from low stakes tests are not how the tests are designed but rather their function in how these results are used.

As states and districts adopt new mandates, it is imperative to understand how the transition from one assessment to another affects the classroom. Korte et al. (1991) conducted a 5-year study in a large school district that was transitioning to a new standardized test. Throughout this process, the school mean score dropped more than half a grade level when compared to the data in the previous years. As the district progressed with the new assessment, the average score for each school noticeably increased. To determine if the increase in scores was equated to teaching to the test or actual growth in student learning, the district re-administered the original standardized test and compared it to the new assessment. The conclusion drawn from the study showed that the growth in student performance on the new test was due in part to teaching to the test. This drawback is a real concern for several stakeholders, as teaching practices are influenced in undesirable ways due to fear of not meeting the stakes set forth by these assessment mandates.

Additionally, perceptions portrayed by teachers also showed a shift in subject content and teaching practices. A survey conducted on reading teachers in Texas revealed, that on average, teachers spent 8 to 10 hours per week preparing students for the Texas Assessment of Academic Skills (TAAS) (Hoffman, Assaf, and Paris, 2001). These training activities included teaching test taking skills, marking answer sheets, reviewing topics on the test, and using commercially prepared resources created from the state department. Therefore, when test scores become the goal of the teaching process, it caused those heavily emphasized tested indicators to lose educational value and distort the overall educational process in undesirable ways (Aviv, 2014).

Much of the research surrounding mandated assessments were completed with samples of administrators, teachers, and support staff members. Very little knowledge exists about how these assessments affect students. As for the research conducted on testing mandates, it is evident that this type of testing environment has created an increased amount of pressure and stress for staff and students being held accountable (Hoffman, Assaf, and Paris, 2001; Koretz et al., 1991). With such a variance in how each individual and/or school organization has responded to the pressure, it would be equally valuable to know how students have responded to these testing measures. Lloyd-Smith & Tarr (2000), indicated that very seldom do we utilize pupils as analysts for schooling and/or measures of appropriateness. Conversely, Heritage (2013) stressed the importance for students to be seen as equal stakeholders, especially in assessments, if it is to be used for improving student learning and understanding. By including student perceptions, another voice will be added to the research on the perceived effects of testing mandates that currently lacks accurate representation.

Correlates in Research that Affect Student Performance

A wealth of literature has examined the demographic makeup of schools and its relationship to the academic performance of students (Berends and Penaloza, 2010; Ewijk & Slegers, 2010; Hanushek, Kain, and Rivkin, 2009; Rumberger and Palardy, 2005). Much of this literature highlights achievement gaps between students with demographics such as gender, race/ethnicity, socio- economic states and English language learners. To further understand the achievement gap, knowing how students perceive these measures could help identify specific strategies to assist those students who have certain demographic factors that have been identified as having greater challenges.

The National Center for Education Statistics (NCES) defines that an achievement gap occurs when a group of students categorized by a common demographic make-up significantly outperforms another group of students that share a different demographic make-up (Snyder, Brey & Dillow 2016). NCES identified that if the average scores for these groups were larger than the margin of error, it can be described as statistically significant meaning an achievement gap has been established between these groups of students. Different student subgroups are faced with different sets of challenges. To understand and control factors leading to the performance gap between subgroups, it is important to learn about the specific issues and perceptions for each subgroup.

It has been shown that an achievement gap has been identified between different races and ethnicities. Vanneman et al. (2009) noted that the academic achievement for students of varying races overall have improved their scores over time, yet the difference in the achievement gap still exists. When looking at the data from the first year of the NAEP assessment (1990), White-Black students had a 32 point difference and White-Hispanic students had a 20 point difference (NAEP Mathematics Report Card, 2018). In 2015 and 2017, a 32 point White-Black score difference existed which was not very different from the 33-point score gap in 1990. Similarly, there was no change in the 24 point White-Hispanic score gap in 2017 compared to the 22 point score difference in 2015 and the 24 point difference in 1990.

In addition to race/ethnicity, Hirsch & Moats (2001), also referenced to the "language gap" and identified language as a predictor that affects student performance. Students who are English language learners (ELL) lack the language skills that help them assess and understand math, science, language arts, and social studies. Many ELL students struggle with these content areas. Children who come to school speaking a different language tend to struggle to meet

literacy requirements, and often feel overwhelmed in the classroom. Furthermore, Francis et al. (2006) argued that academic language can be the single most important determinant of academic success.

Another variable examined in research was gender. Many national studies in the U.S. found that on average males outperform females on math tests and females outperform males on English Language Arts (ELA) tests (Lee, Moon, & Hegar, 2011; Fryer & Levitt, 2010). Other studies highlighted gender stereotypes and the availability of socioeconomic resources as possible contributors to gender disparities in academic interests and achievement (Baker & Milligan, 2013; Raley & Bianchi, 2006). Although there is no strong evidence that parents spend more money on male or female children (Hao & Yeung, 2015), there is evidence that parents invest their time and money in gender specific activities (Raley & Bianchi, 2006). For example, parents engage in more reading, storytelling, and verbal activities with their female children as early as 9 months of age (Baker & Milligan, 2013), but spend more time on science and math for male children (Raley & Bianchi, 2006). In particular, stereotypes may contribute to shaping students' beliefs about their academic capability (Jacobs et al., 2002), their interest in different subjects (Cech, 2013; Charles & Bradley, 2009), and their academic performance (Tomasetto, Alparone, & Cadinu, 2011).

Another variable identified as a strong predictor for students' academic success was parental/guardian education. For example, a child exposed to parents who model achievement-oriented behavior (obtaining advanced degrees; reading frequently; encouraging a strong work ethic) and provided achievement-oriented opportunities (visits to the library and museum; after-school programs; educational books and videos) were more likely to develop a perception that achievement is valued, pursued, and anticipated (Dubow, Boxer and Huesmann, 2010).

In conclusion, understanding student perceptions towards these measures could offer more information on how to support these students with these demographic factors who have been identified as having greater challenges. Having knowledge of students conceptions towards their academic performance on mandated assessments could explain why certain students devote more time and effort into preparing and completing these assessments. Researchers, educators, and policymakers continue to seek to understand the reason for these gaps and to develop policies to reduce the magnitude of these gaps in student academic performance (Hanushek, Kain, and Rivkin 2009; Kelly 2009; Reardon 2008; Stiefel, Schwartz, and Ellen 2007). Given that demographic characteristics can be meaningful predictors of student performance, this study, will examine gender, race/ethnicity, language, highest level parent/guardian education, student GPA and plans after high school.

Student Perceptions Towards Mandated Assessments

As testing mandates increase at the federal, state and district level it is imperative to acknowledge how student perceptions can influence performance. Wise and DeMars (2005) described test-taking effort as the extent to which students give their “best effort to the test, with the goal being to accurately represent what one knows and can do in the content area covered by the test” (p. 2). Student perceptions towards an assessment and its purpose can have direct impact on the effort and overall performance. Understanding what students find purposeful and/or irrelevant from these tests could support how stakeholders (educators, administrators and parents) communicate to students about the purpose and significance for putting forth the effort to improve performance.

Research conducted on elementary and secondary students shows that older students tend to have more skeptical views about assessments in general and of accountability tests than younger students. For example, Paris, Roth, and Turner (2000) surveyed 2nd through 11th grade students in Michigan, California, Arizona, and Florida. Their investigation found that students generally held positive views about mandated assessments, but with increasing age and familiarity students developed more negative attitudes towards these tests. Older students reported distrust in the accuracy and validity of the test, disappointment with the feedback after taking the test, and concern for the social comparisons following the results of the test. Furthermore, older students reported less effort and strategies to assist with taking the tests than compared to younger students. Similar findings by Paris, Roth, and Turner (2000), who surveyed 4th, 7th, and 10th grade students in Michigan, found 10th grade students tended to harbor more negative attitudes towards the state-mandated test than their younger peers. For example, 10% of the 10th grade students agreed with the statement “It is important for me to do well on this test.”, compared to 7th grade students with 62%. Additionally, tenth graders thought that it was more acceptable to cheat, fill in answers, and try half-heartedly. The results of these studies indicated that as students got older, their attitudes towards standardized tests became more negative which adversely influenced their test-taking behavior.

Another factor affecting student motivation was the level of stakes placed on students for their performance regarding the mandated assessment. Wise and DeMars (2005) completed a comprehensive review that examined the effects of test-taking motivation on low-stakes test performance and they concluded that on average unmotivated students performed one half of a standard deviation lower than motivated students. Research showed that unmotivated students scored lower on low-stakes assessments than motivated students, after controlling for ability

(Eklof, 2010; Steedle, 2010; Wise & DeMars, 2005; Sundre & Kitsantas, 2004;). Eklof (2007) warned that ignoring the component of test-taking motivation in a low-stakes testing environment could lead to confounding knowledge and motivation thereby threatening the validity of the test results altogether.

Research on test-taking motivation showed that students who had low-stakes accountability may not have been properly motivated to demonstrate what they know and/or are able to do, compromising the validity of the test scores (Liu, Bridgeman, & Adler, 2012; Haladyna & Downing, 2004). Similar studies (Thelk, 2006; Wise & Bhola, 2006; Wise and Kong, 2005;) found a moderately positive correlation ($\sim .30$) between student effort and performance. For example, Wise & DeMars (2005), hypothesized that how students perceive the importance of a test can have an indirect effect on their test performance, mediated through their effort taken on the test. Another study by Wheelock, Bebell, and Haney (2000) examined 4th, 8th, and 10th grade students' drawings of themselves taking the Massachusetts high-stakes test. The findings portrayed students as anxious, angry, bored, pessimistic and withdrawn from testing. It was evident that many students felt stressed while taking mandated tests, and as a result, students perceived to have a decrease in social functioning, self-worth, and achievement. Adjusting the way stakeholders approach these mandates could provide better meaning and purpose for students, increase their effort and motivation, as well as provide more accurate data representing their true understanding and knowledge.

As testing mandates increase it is important to understand how student perceive these initiatives. Knowing the effect motivation can have on student performance, further studies have explored ways to help improve student motivation such as material incentives, feedback, and interventions. (Brown, 2010, Wise and DeMars, 2005). In addition to these strategies, Brown

(2010) stressed the importance of highlighting the value of the test and appealing to the students' willingness to improve their school as ways to increase motivation.

In summary, student perceptions towards assessments and the rationale for having them can directly impact their effort, motivation and overall performance on the tests. Over time students become more negative about these assessments. As students advance to the next grade or have certain stakes tied to their performance it can cause their perception towards testing mandates to sway. Therefore, knowing what students find purposeful and/or irrelevant from these assessments could provide stakeholders with more direction on how to effectively approach students and communicate a more meaningful message from their viewpoint.

Instrument Measuring Perceptions of Assessments

Studies examining student perceptions have explored areas on how students perceive classroom instruction (Carnine, Silbert, Kame'enui & Tarver, 2010; Wiliam & Leahy, 2007;) feedback (Rowe, 2011; Price et al., 2007; Holmes & Papageorgiou, 2009; Poulos & Mahony, 2008), teacher effectiveness (Rudduck, 2007) and types of assessment (William & Thompson, 2007; Black et. al, 2004). Another large area of research conducted on student perceptions is the area of student motivation. For example, the Student Opinion Survey (SOS) by Sundre (1999), is a ten question Likert-scale that measures participants' self-reported effort and task completion. As mentioned in the research, both of these constructs are considered to be important for test-taking motivation. Examples of an item related to effort is "I engaged in good effort throughout this test," and an example of an item related to importance is "Doing well on this test was important to me." Another survey is the Sense of Community Index-2 (SCI-2) by Chavis, Lee, & Acosta (2008). It is a 25 item Likert-scale that measures participants' level of feelings and connectedness to their institution. The survey asks participants to indicate how important it is for

them to feel a sense of community with other community members. Examples of these items include “I can recognize most members of this community,” “Fitting into this community is important to me,” and “I feel hopeful about the future of this community.” For instance, a study by Brown and Walberg (1993) manipulated the sense of shared responsibility amongst elementary school students in Chicago by telling them prior to taking a test that their scores would be used to evaluate the quality of their schools and their teachers. The results of the study demonstrated that students who received such instructions performed better than controls. Similarly, Wise and Smith (2012) hypothesized that student participants may be motivated internally by appealing to their sense of ‘academic citizenship.’ Academic citizenship has been framed in the context of students helping for the overall benefit of their school (e.g., Schmitt et al., 2007), and taking responsibility for a task whether or not it is asked of them (e.g., Gore, Kiefner, & Combs, 2012).

As for student perceptions towards mandated assessments, large-scale survey studies in New Zealand have observed important aspects of students’ attitudes and perceptions about assessment (e.g. Brown & Hirschfield, 2007, 2008). For example, one instrument measuring student perceptions towards assessments was created by Dr. Gavin Brown (2004), the Students Conceptions of Assessment Survey (SCoA). A variety of assessments have been conducted using the SCoA instrument.

Brown (2004) developed the first inventory of the SCoA by analyzing empirical studies on students’ perceptions about testing in New Zealand. The first inventory had a total of 33 items aggregated into a hierarchical model of 7 factors distributed across inter-correlated factors: Improvement (teacher improves students, self- improvement), Beneficial (personal enjoyment, class enjoyment), External (school accountability, student future), and Irrelevance (first-order

factor). The Cronbach alpha is a measure of the scale reliability for how closely related a set of items are as a group. For the four factors in the SCoA, it has relatively good internal consistency: Improvement, $\alpha = 0.88$; Beneficial, $\alpha = 0.85$; External factors, $\alpha = 0.78$; and Irrelevance, $\alpha = 0.80$. The first study found that several conceptions of assessment significantly predicted mathematics achievement (Brown & Hirschfeld, 2008). The second study conducted in New Zealand had 11 questionnaire items which generated four conceptions of assessment that were found to have a strong relationship to reading achievement (Brown & Hirschfeld, 2008). The four conceptions path regression weights to achievement were: assessment makes students accountable ($\beta = .42$), assessment makes schools accountable ($\beta = -.27$), assessment is fun ($\beta = -.24$), and I ignore assessment ($\beta = -.14$).

Further studies that have utilized versions of SCoA have been conducted in Hong Kong (Wang & Brown, 2014), Brazil (Matos et al., 2009), Germany (Hirschfeld & von Brachel, 2008), Netherlands (Segers & Tillema, 2011), Iran (Brown, Pishghadam & Sadafian, 2014), United States (Wise & Cotton, 2009) and Cyprus (Michaelides & Sidiropoulou, 2014). These studies identified that student conceptions toward assessments usually differed in levels of endorsement, revealing the significance for local educational policies and contexts in shaping how students perceive assessments and its purposes. For instance, if students agree that assessments evaluate them, their performance and effort on the assessment are likely to increase. In contrary, if students perceive assessments are used to evaluate schools or can be ignored, their achievement and effort tend to decrease (Brown and Hirschfeld, 2008; Brown et al., 2009; Wise and Cotton, 2009). Brown (2011) reiterated the importance of acknowledging the connection for how students perceive the purpose and nature of the assessment because it can greatly affect their behavioral responses to assessment performance.

To conclude, studies focusing on student perceptions towards assessments have explored several areas at the classroom level in the United States and internationally with other assessment systems. Studies utilizing versions of the SCoA have highlighted the significance and value for understanding how students perceive assessments and its purposes. In the United States, with local control established at the state and district level, understanding how students perceive mandated assessments is important in order to get a better gauge about how students perceive these tests, the value, and their performance. This study specifically focuses on mandated assessments using the SCoA-VI and provides students an opportunity to share their thoughts in ways other students have not.

Summary

This study will examine student perceptions towards mandated assessments. This study will expand the current research and extend the information on the perceived effects that have occurred as a result of testing mandates. Reviewing the historical development of mandated assessments, it is evident that testing mandates have become very influential. These educational policies have included additional federal funding to help combat poverty and at-risk students, as well as requiring yearly reporting/publication of results to increase school and teacher accountability. Many of these studies conducted about mandated assessments focused on the perceptions of administrators, staff, and parents.

Research has shown that stakeholders have responded to testing mandates in differing ways. The intended outcomes reported were creating more consistency across classrooms, better alignment of the curriculum, increasing the amount of effort by faculty, and improvement in student test performance. As for the unintended effects, these included non-tested content

becoming limited, unnecessary student placement/identification for testing accommodations, less effective teaching strategies, and teaching to the test. In addition to the research on perceived effects, other factors identified as pertinent to the performance of students are demographic characteristics and personal effort and motivation exerted by students. It is evident from the research that factors contributing to a student's make-up and how they perceive the testing experience and/or expectation can affect their performance. Studies have also explained correlates affecting student performance on mandated tests. Understanding the perceptions of students and how they perceive the testing experience could help substantiate or contradict why testing mandates are a legitimate process to measure schools, teachers, and student performance.

During the past two decades of educational policy implementation, testing mandates continue to be a priority set forth by policymakers as a means to measure schools, teachers and students. The information gained from this study could be impactful to several stakeholders responsible for implementing and supervising testing mandates. Also, demographics can equally contribute to student performance, examining students in the study by these factors will provide more information on how these particular subgroups perceive and respond to these testing mandates. Student perceptions towards mandated assessments has been under-studied. The information gained from this study could not only contribute to the current research regarding testing mandates but expand other facets of research affecting student performance.

Chapter 3

Methodology

Introduction

The purpose of this section is to describe the methodology utilized to investigate the perceptions of ninth and tenth grade students towards mandated assessments in regard to improvement, external attribution, affective benefits, and irrelevance. The study further examined student perceptions toward mandated assessments based on demographic characteristics including gender, race/ethnicity, primary language, parent education, GPA and plans after high school. The methodological elements discussed include 6 areas: research design, sample and setting, instruments, procedure, data analysis and ethical considerations/limitations.

Research Design

This is a descriptive study consisting of both quantitative and qualitative data to investigate ninth and tenth grade students' perceptions toward mandated assessments. The demographic questionnaire and the SCoA-VI survey yielded quantitative data and open-ended questions provided qualitative data. This approach allowed students an opportunity to elaborate in greater detail the benefits and drawbacks of having participated in mandated tests. Creswell and Plano (2007) described that mixing both methods was more than simply collecting and analyzing both kinds of data, but rather involving the use of both approaches in tandem so that the overall strength of a study is greater than either qualitative or quantitative research alone. By utilizing this approach, the study sought to create a solid foundation for drawing conclusions about the perception's students have regarding mandated assessments.

Setting and Sample

The setting for the study took place in a large school district located in the Kansas City metropolitan area. Approximately 27,648 students ranging in grades kindergarten through twelve grade were enrolled for the 2017-2018 school year. The school district has a total of thirty-three elementary schools, five middle schools, five high schools and six instructional centers. Of the five traditional public high schools, each school has approximately 1,000 students in grades nine through twelve.

A purposive sample of ninth and tenth grade students were selected from all five high schools. The participants had to meet the following characteristics: (1) currently enrolled as a full-time student and (2) having participated at some point in federal, state and/or district mandated testing throughout their elementary and secondary education. The school district currently utilizes large-scale low-stakes accountability assessments to meet external mandates, measure student growth, evaluate teacher/student performance, and guide instructional practice (See Table 1).

Table 1: State & District Mandated Tests

Mandated Assessment	Subject	Grades	Times per Year
Individual Growth and Development Indicators Test (MyIgdIs)	Reading, Math	PreK	3
Measure of Academic Process (MAP):	Reading, Math	Non-Title: 3-9 Title: K-9	3
Dynamic Indicators of Basic Early Literacy Skills (DIBELS)	Reading	K-6th	3
KS Assessment	ELA, Math, Science, History/Government	ELA/Math: 3-8, 10 Science: 5, 8, 11 H/G: 6th	1
National Assessment of Educational Progress (NAEP)	Reading, Math, Science, Writing, Technology Engineering Literacy, Arts, Civics, Geography, Economics, US History	4, 8, 12	1
International Mathematics and Science Study (TIMSS)	Math, Science	4, 8	1

Table 2: Sample & District Demographic Comparison

Demographic	Sample	District
Gender		
• Female	57.8%	49.1%
• Male	42.2%	50.9%
Race/Ethnicity		
• Non-Hispanic White	81.1%	63.9%
• Black or African American	5.0%	8.9%
• Hispanic or Latino	5.8%	18.7%
• Other	8.1%	8.6%
Primary Language		
• English	95.8%	87.9%
• Other	4.2%	12.1%
Parent/Guardian Level of Education		
• High	64.2%	65.3%
• Low	35.8%	34.7%

Next the demographic information from the district and sample are reported in table 2.

The demographic comparison revealed that the sample's demographic subgroups were not equal to the proportion of the subgroups within the district. Out of 360 study participants, 208 were female and 152 were male. The district's population is about 50/50 male and female. As for the next demographic, race/ethnicity, showed a relatively large Non-Hispanic White population for the sample. This may have been due to the process required to get students to participate in the study. Similarly, the district roughly has a 64% Non-Hispanic White population. However, the study produced an even larger percent with 81% Non-Hispanic White. This was substantially more than any other race/ethnicity represented in the study. This created a large discrepancy within the sample and from the district's population.

The third demographic examined in the study was students' primary language. The district has a relatively large student population, whose primary language is English. The study also had a considerable number of participants with 345 out of the 360 selecting English as their primary language. The difference between the district and sample size was 7.9%.

The fourth demographic comparison was parents' level of education. Two demographic categories were created; "High" if participants identified a parent or guardian having attained an associate degree or higher and "Low" if a parent or guardian did not achieve an associate degree. The study had 287 participants labeled "high" and 73 "low" regarding the level of parent/guardian education. When comparing the parent/guardian level of education the district's population to the sample had a 2.1% difference.

The last two demographic questions did not have district totals to compare; Question E, Grade Point Average (GPA) and Question F, plans after high school (See Table 3).

Table 3: Sample Grade Point Average and Plans after High School

Demographic	Sample
GPA	
3.5 and Above	66%
2.5-3.4	26%
1.5-2.4 or Other	8%
Plans After HS	
College or Technical Training	79%
Full-time Job or Gap Year	7%
Not Sure	14%

Of the 360 study participants, 238 designated their GPA as "3.5 and Above" which was considered to be mostly A's. In addition, 283 students selected college or technical training as their future plans after high school. It is important to note that all of these students in the study were either starting or partially into their high school education. Therefore, responses from these 9th and 10th grade students could change due to the time remaining in their high school career.

In summary, the sample had more females, white non-Hispanics, and English speaking than the total district population. These demographics tend to be higher performing, which was

reflected in high GPAs and plans for after high school. The level of parent education identified by the participants did not differ from the district's levels.

Instruments

For this study, a demographic questionnaire, Brown's student conceptions of assessment survey, and three open-ended response questions were utilized. In the following section, each of these instruments will be discussed.

Demographic Questionnaire

A demographic questionnaire was constructed by the investigator to gather relevant data in order to profile the subjects (See Appendix D). Nebeker et al. (2015) described descriptive studies as information-gathering activities in which the researcher attempts to accurately describe a naturally occurring health state, behavior, attitude or a particular group. The demographic variables included gender, race/ethnicity, primary language, highest level of education for parent or guardian, the students' high school GPA (They were provided a range), and the student's plan after high school. These demographic variables were selected because they are either (1) descriptors of the study sample so it could be compared to the district and (2) variables that could potentially affect the perceptions of the students towards mandated assessments. The subjects in the study had the ability to skip, select "other", or "not sure" on the demographic questionnaire.

Students' Conceptions of Assessment (SCoA-VI)

The Students' Conceptions of Assessment (SCoA-VI) inventory (Brown, Irving, Peterson, & Hirschfeld, 2009; Brown, Peterson, & Irving, 2009; Brown, 2008) was designed to elicit self-reported responses from participants about testing (See Appendix E). Brown (2008),

identified four major conceptions of assessment: (1) Improvement, (2) External Attribution, (3) Affective Benefits and (4) Irrelevance (See Table 4).

Table 4: SCoA-VI Conceptions & Subfactors

Conceptions & Subfactors		# of Questions
Conception 1: Improvement		11
Subfactor 1	Student Learning (Self)	5
Subfactor 2	Teaching (Teacher)	6
Conception 2: External Attribution		6
Subfactor 1	School Quality	2
Subfactor 2	Future Education/Employment	4
Conception 3: Affective Benefits		8
Subfactor 1	Enjoyable	2
Subfactor 2	Helpful	6
Conception 4: Irrelevance		8
Subfactor 1	Unfair/Bad	5
Subfactor 2	Ignored	3

The “Improvement” conception measures perceptions that assessments improve the teacher’s teaching and the student’s learning. The second conception, “External Attribution” examines the perception that assessments serve an external purpose such as judging a school’s quality or predicting a student’s future educational and employment success. Next, “Affective Benefits” conception assesses a student’s perception that assessments are enjoyable and helpful to classmates. The fourth conception, “Irrelevance”, measures students’ perceptions that testing is “unfair/bad” and “ignored”.

Students were asked to respond to each statement using a 7-point Likert scale ranging from 1 (completely disagree) to 7 (completely agree). In the original survey, Brown (2004) noted that students tended to respond positively to all items, therefore the responses on the SCoA-VI used a Likert-scale that was fully labeled with all responses positively-packed. The Cronbach alpha is a measure of the scale reliability for how closely related a set of items are as a group. The internal consistency of the four factors of the SCoA-VI were: Improvement, $a = 0.88$; Beneficial, $a = 0.85$; External factors, $a = 0.78$; and Irrelevance, $a = 0.80$.

For this study, the 7-point Likert scale was modified by displaying only the endpoints. This concept of “satisficing”, coined by Simon (1955), suggests that when responses become taxing, subjects tend to put forth limited effort by selecting satisfactory rather than adjusting their responses as needed. The reduced positive labeling on the scale was considered less demanding for the subjects who looked for cues and therefore less vulnerable to response bias. While students completed the online survey, up to seven questions were shown at a time. The subjects were required to answer all questions before clicking “next” to be able to move to the subsequent set of questions. Students also had the option to select “prev” to go back to the previous set of answered questions to edit as needed. Student participants were required to answer each question by selecting a point on the Likert scale from strongly disagree to strongly agree.

Open-Ended Response Questions

In addition to the demographic questionnaire and the modified survey provided by Brown et al. (2009), three open-ended questions were added at the end of the survey (1) Overall, what do you see as the benefits for you having completed mandated tests in school? (2) What do you see as the drawbacks for you from having completed mandated tests in school? and (3) Do you have any other opinions about mandated tests you would like to share? Providing these open-ended questions created an opportunity for students to elaborate further on their experiences with standardized assessments and to provide descriptive details to examine the study’s research questions (See Appendix F). All three qualitative questions were displayed together at the end of the survey. Students were provided a box to submit their answers for each question but had the option to leave the response empty if desired to complete the survey.

Procedure

The procedure section will describe the steps that occurred during the data collection process. First, permission and approval by Dr. Gavin Brown, Director of Quantitative Data Analysis and Research Unit at the University of Auckland, to utilize the Student Conceptions of Assessment Inventory Version VI (SCoA-VI) was sought and granted for the study (See Appendix A). Second, the demographic questionnaire, SCoA-VI survey, and the three open-ended questions were submitted into an online survey platform to present to the IRB of the school district and the university. Once approval was obtained (See Appendix B), the Assessment Coordinator compiled the subjects who met the study criteria which were (1) currently enrolled as full-time student and (2) having participated at some point in federal, state and /or district mandated testing.

Next, an email was sent requesting parent/guardian consent and the link to the student survey to the Assessment Coordinator (See Appendix C). The coordinator then sent out the email using the district's distribution communication system (Skyward) to 4,187 ninth and tenth grade parents/guardians seeking permission for their student to complete the survey. The informational email was sent on March 26, 2018 at 5:00pm. The email provided a brief overview to ninth and tenth grade parents and guardians about the study and requested them to forward the survey link on to their child thus granting their permission for the student to participate in the study. It is not known how many parents/guardians opened and read the email.

Two questions were sent back from parents to the coordinator (1) requesting a copy of the questions asked on the survey and (2) if the email could be translated in Spanish. Both of these requests were granted and provided to these individuals through email by the Assessment Coordinator. A reminder email was sent to all parents and guardians again on Monday, April 9,

2018. The survey was open for responses over a four-week period starting on March 26, 2018 and ending on April 20, 2018. A bulk of the responses were completed on the days that the informational email was sent to parents (3/26/18) which had 256 online surveys submitted and again on the reminder email (4/9/18), which had 122 surveys completed. Thus 378 responses were received from the total mailing population of 4,187. Again, it is not known how many of the 4,187 emails sent to parents were actually read.

The survey was administered within the district and state's testing window for the State assessment and MAP assessment. Administering the survey at this time frame allowed for students to be able to quickly reference and reflect on their experiences from these Spring assessments. At no time during this process were the names of the students and schools collected therefore securing the anonymity for all parties. The total responses back from the survey were 378, but only 360 were fully completed by the participants. The response percentage from the online survey was 9%, slightly below the average percentage rate 10-15% for online surveys (Fryrear, 2015). However, given that it is not known how many parents or guardians saw or even opened the email, the 9% rate can be a little misleading. Parents/Guardians were sent an email that a request had been posted in their student's Skyward portal. Next, the parent/guardian had to login into Skyward to be able to access the email message. Thus, the 9% response rate is likely lower than the rate of those who actually were aware of the study.

Data Analysis

To examine the research questions, the responses from the SCoA-VI were used to determine a mean and standard deviation for every question on the survey, as well as the four conceptions (Improvement, External Attribution, Affective Benefits, and Irrelevance) and the

eight sub factors (student learning, teaching, school quality, future education/employment, enjoyable, helpful, unfair/bad, and ignored). Further investigation of the demographic characteristics using T-tests and analysis of variance (ANOVAs) were utilized to examine the relationship of the demographic variables to the mean scores for each of the four conceptions. Further discussion about the analysis is presented below by each research question.

Research Question 1

Research question 1 asked: What are high school students' conceptions of the effects mandated testing has had on their academic experiences? First, the mean and standard deviation was calculated for each question on the SCoA-VI and then for each of the four conceptions (Improvement, External Attribution, Affective Benefits, and Irrelevance) and their sub factors (student learning, teaching, school quality, future education/employment, enjoyable, helpful, unfair/bad, ignored). Therefore, a total of 45 mean scores (33 question items, 4 conceptions, and 8 sub factors) were configured and analyzed to determine the sample mean and standard deviation.

Next, information obtained from the three open-ended questions were analyzed to learn more specifics about attitudes towards these testing mandates and to synthesize an overall meaning. Rossman & Rallis (1998) described this coding process as "organizing the material into chunks or segments of text before bringing meaning to information" (p.171). Segmented text from each qualitative question were categorized into groups and labeled with a common term or phrase used by the participants in the study. Creswell (2009) described the interpretation of the results, "as a means for the researcher to draw conclusion from the result for the research

question” (p. 152). This method created a basis for the language and words used to express how the participants perceived mandated assessments.

Research Question 2

Research question two asked: Do student conceptions of mandated assessments differ based on their demographic characteristics? A nominal scale was assigned to each participant’s demographic characteristic (gender, race/ethnicity, primary language, parent education, high school GPA, and plans for after high school) with a number to serve as a label with no quantitative value. T-tests were used to help explain the relationship between the independent variables: gender (female or male) and primary language (English or Spanish) to the dependent variables (overall sample mean score on the SCoA-VI and the mean scores for the four conceptions). These items are identified as A and C on the demographic information questionnaire (See Appendix D). To determine if the demographic variables: race/ethnicity (non-Hispanic, White, Black/African American, Hispanic/Latino or other), highest level of parent education (did not graduate high school, graduated high school, some college/technical school, associate’s degree, bachelor’s degree or graduate degree), the students high school GPA (3.5 and above Mostly A’s, 2.5-3.5 Mostly B’s, 1.5-2.5 Mostly C’s, and Other) and the student’s plan after high school (4-year college, 2-year college or technical training, Full time job or Take time off “Gap year”) impacts the overall sample mean score and the mean scores for the four conceptions in the study, ANOVA tests were conducted. These question items are B, D, E and F on the demographic questionnaire (Appendix D). The results of the tests were analyzed to determine if the demographic characteristics accounted for the variance in the mean of each conceptions at a .05 statistical level of significance ($\alpha = .05$).

Limitations

The results of this study must be considered in view of several methodological limitations. First, a convenience sample was utilized because of the accessibility of the sample, hence all conclusions are not generalizable to any specific population. Second, having students under the age of 18 as participants in the study created a challenge and added several steps in order to garner consent for participation in the study. Therefore, a smaller return rate was produced in the study due to parental permission having to be granted before student participation would be requested. For that reason, permission could be denied or limited in some way, by either the parent or student. A third limitation involved the SCoA-VI survey being conducted online. Certain populations might be less likely to have internet access preventing them from participating and/or responding to the survey. Also, frequent internet users are constantly bombarded by messages/junk mail and could easily delete the consent and invitations to participate in the study.

A fourth limitation was with Skyward, the district's communication portal. Skyward sends an alert to the parents/guardians' email to let them know a message has been posted in Skyward. After receiving the email alert, parents/guardians logged into Skyward in order to view the message. Once the parent/guardian viewed the post, they were requested to forward the message on to their student signifying their approval/consent. Unfortunately, parents/guardians do receive numerous alerts from Skyward which can cause it to be overlooked. Utilizing this communication method and having subjects under the age of 18, created multiple steps thus hindering the study's turnout/response rate.

A fifth limitation is that students self-reported responses were without validity checks. Using an online platform makes it difficult to know if the sample providing the responses is the

correct and intended participant (i.e. could be a family member, friend, etc.). Also, the length of the survey may have caused subjects to be less likely to stay fully engaged during the entire survey since it required more than 8-10 minutes of their personal time.

Chapter 4 Results

The purpose of this chapter is to present, analyze and interpret the data gathered from 9th and 10th grade students to identify how these subjects perceive mandated assessments and if it varies based on their demographic characteristics.

Research Question 1: What are high school students' perceptions toward mandated testing on improvement, external attribution, affective benefits and irrelevance as measured by Brown's SCoA-VI survey?

A) Do high school students feel testing mandates improve teaching and student learning (Improvement)?

B) Do high school students feel mandated assessments measure the school's quality and predict the student's future in education and employment (External attribution)

C) Do high school students feel these assessments are beneficial to themselves and classmates (Affective benefits)?

E) Do high school students feel mandated assessments are unfair and ignored (Irrelevance)?

First a mean and standard deviation for each question on the SCoA-VI was calculated (See Table 5). The Likert-scale used to score the responses from the SCoA-VI ranged from 1 (completely disagree) to 7 (completely agree).

Table 5: Mean and Standard Deviation on the SCoA-VI

Questions on the SCoA-VI *	\bar{x}	σ
I pay attention to my assessment results in order to focus on what I could do better next time	4.00	1.78
Assessment encourages my class to work together and help each other	2.24	1.40
Assessment is unfair to students	4.11	1.70
Assessment results show how intelligent I am	2.75	1.67
Assessments help teachers track my progress	3.91	1.78
Assessment is an engaging and enjoyable experience for me	1.75	1.29
I ignore assessment information	3.28	1.73
Assessment is a way to determine how much I have learned from teaching	3.20	1.75
Assessment is checking off my progress against achievement objectives or standards	3.61	1.66
I make use of the feedback I get to improve my learning	3.28	1.85
Assessment provides information on how well schools are doing	3.42	1.76
Assessment motivates me and my classmates to help each other	1.98	1.31
Assessment interferes with my learning	4.18	1.77
I look at what I got wrong or did poorly on to guide what I should learn next	3.17	1.67
I use assessments to take responsibility for my next learning steps	2.99	1.61
Assessment results predict my future performance	2.39	1.58
Our class becomes more supportive when we are assessed	2.04	1.34
Teachers are over-assessing	4.60	1.67
I use assessments to identify what I need to study next	2.66	1.58
Assessment is important for my future career or job	2.57	1.84
When we do assessments, there is a good atmosphere in our class	2.61	1.67
Assessment results are not very accurate	4.57	1.64
My teacher uses assessment results to help me improve	2.86	1.54
Assessment measures the worth or quality of schools	2.73	1.67
Assessment makes our class cooperate more with each other	1.90	1.22
Assessment is value-less	4.37	1.76
Teachers use my assessment results to see what they need to teach me next	3.03	1.67
When we are assessed, our class becomes more motivated to learn	1.86	1.22
I ignore or throw away my assessment results	3.72	1.97
The state assessment shows whether I can analyze and think critically about a topic	3.31	1.70
I find myself really enjoying learning when I am assessed	1.91	1.39
Assessment has little impact on my learning	4.79	1.87
Mandated test tells my parents how I am doing	2.89	1.83
Survey Sample (n= 360)	3.11	1.85

* *Likert: 1=Completely Disagree 7=Completely Agree*

As noted in table 5, the sample mean was 3.11 with a standard deviation of 1.85. The highest mean scores from the SCoA-VI were classified under the “Irrelevance” conception.

Specifically:

- “Assessment has little impact on my learning” ($\bar{x} = 4.79$)
- “Teachers are over assessing” ($\bar{x} = 4.60$)
- “Assessment results are not very accurate” ($\bar{x} = 4.57$)
- “Assessment is value-less” ($\bar{x} = 4.37$)
- “Assessment interferes with my learning” ($\bar{x} = 4.18$)
- “Assessment is unfair to students” ($\bar{x} = 4.11$)

The participants in the study were likely to agree with the “Irrelevance” of mandated assessments. Within the “Irrelevance” conception, 6 out of the 8 statements had the highest mean scores, placing it on the agree side of the scale ($\bar{x} > 4.0$) For students, “Irrelevance” meant mandated assessments were perceived to have little impact and were not as important to what was currently going on within the students’ academic environment.

Conversely, questions with the lowest means that participants were likely to disagree with were in the “Affective Benefits” conception. Specifically:

- “Assessment is an engaging and enjoyable experience for me” ($\bar{x} = 1.76$)
- “When we are assessed, our class becomes more motivated to learn” ($\bar{x} = 1.86$)
- “Assessment makes our class cooperate more with each other” ($\bar{x} = 1.90$)
- “I find myself really enjoying learning when I am assessed” ($\bar{x} = 1.91$)
- “Assessment motivates me and my classmates to help each other” ($\bar{x} = 1.98$)
- “Our class becomes more supportive when we are assessed” ($\bar{x} = 2.04$)
- “Assessment encourages my class to work together and help each other” ($\bar{x} = 2.24$)

Students were more inclined to disagree that mandated assessments were “enjoyable to self” and “helpful to classmates” because 7 out of the 8 statements concerning “Affective Benefits” had the lowest mean scores on the SCoA-VI. This is consistent with the results of the

“Irrelevance” factor. Therefore “Affective Benefits” was not viewed favorably among the participants and had the closest mean score to the endpoint on the disagree side of the scale causing it to be the most disliked conception out for the four from the SCoA-VI. This meant students perceived mandated assessments did little to improve or promote their learning and classmates.

In conclusion, based on the sample mean and standard deviation from each statement on the SCoA-VI, students were more likely to view negatively testing mandates by agreeing with the “Irrelevance” for these exams and further describing them as “unfair/bad” and “ignored”. As for the “Affective Benefits” conception, participants disagreed that testing mandates were “enjoyable for themselves” and “helpful to their classmates”.

Next, as noted in Table 6, the mean score and standard deviation were calculated for each conception (Improvement, External Attribution, Affective Benefits, and Irrelevance) and the subfactors (Student learning, teaching, school quality, future education/employment, enjoyable to self, helpful to classmates, unfair/bad, and ignored).

Table 6: Mean and Standard Deviation for the SCoA-VI Conceptions and Subfactors

SCoA-VI Conceptions and Subfactors		\bar{x}	σ
Conception 1	Improvement	3.27	1.77
	Student Learning (Self)	3.22	1.76
	Teaching (Teacher)	3.32	1.72
Conception 2	External Attribution	2.79	1.76
	School Quality	3.08	1.75
	Future Education/Employment	2.65	1.75
Conception 3	Affective Benefits	2.04	1.39
	Enjoyable	1.83	1.35
	Helpful	2.10	1.39
Conception 4	Irrelevance	4.20	1.82
	Unfair/Bad	4.36	1.72
	Ignored	3.93	1.96

* *Likert: 1=Completely Disagree 7=Completely Agree*

Table 7: The Conceptions on the SCoA-VI

Conceptions	\bar{x}	f- value	f- critical	P-value
Improvement	3.27	204.2	2.61	2.69
External Attribution	2.79			
Affective Benefits	2.04			
Irrelevance	4.20			

The conception “Irrelevance” had subfactors with the highest mean scores that fell on the agree side of the Likert scale. Participants tended to agree that mandated assessments were “unfair/bad” ($\bar{x} = 4.36$) “ignored” ($\bar{x} = 3.93$). The subfactor “unfair/bad” referred to mandated tests as “interfering with learning”, students are being “over assessed”, results are not “accurate” and “value-less”. As for the subfactor “ignored”, mandated assessments were perceived to have “little impact” causing subjects to ignore the information and results. Out of the four conceptions, “Irrelevance” was the only conception to have a mean score on the “agree” side of the scale. As a result, students were likely to perceive mandated assessments as an irrelevant measurement that was often seen as “unfair/bad” and “ignored”.

As for the disagreement side of scale, the other three conceptions on the SCoA-VI: “Affective Benefits”, “External Attribution” and “Improvement”, all had mean scores below the midpoint ($\bar{x} < 4.0$) the Likert Scale. The conception that had subfactors with the lowest mean scores was “Affective Benefits”. These subfactors were “enjoyable” ($\bar{x} = 1.83$) and “helpful” ($\bar{x} = 2.10$). On the SCoA-VI, subjects were likely to disagree that mandated tests were an “engaging and enjoyable experience” that helped “encourage” and/or “motivate” the class. Consequently, making “Affective Benefits” the closest conception to the endpoint “completely disagree” than any of the other conceptions on the SCoA-VI.

The next conception participants were likely to disagree with was “External Attribution”. The subfactors for “External Attribution” were “school quality” ($\bar{x} = 3.08$) and “future

education/employment” ($\bar{x} = 2.65$). Subjects tended to disagree mandated test provided accurate and valuable information on the overall school’s performance. Particularly:

- Mandated tests provide information on how well schools are doing ($\bar{x} = 3.42$)
- Mandated tests measure the worth or quality of schools ($\bar{x} = 2.73$)

In addition to the subfactor, “school quality”, “future education/employment” within the “External Attribution” conception showed to have an even greater dislike among participants.

Specifically:

- Mandated test results show how intelligent I am ($\bar{x} = 2.75$)
- Mandated tests results predict my future performance ($\bar{x} = 2.39$)
- Mandated tests are important for my future career or job ($\bar{x} = 2.57$)
- Mandated tests tell my parents/guardians how much I have learned ($\bar{x} = 2.89$)

Participants tended to disagree more within these two subfactors that mandated tests “predict future performance” and are important for the student’s “future career or job”. The last conception on the disagreement side of the scale was “Improvement”. The subfactors were “student learning” ($\bar{x} = 3.22$) and “teaching” ($\bar{x} = 3.32$). “Student learning” focused on the usage students had for the exam’s results and if the feedback provided guidance on what students needed to do next.

For example:

- I pay attention to my results on mandated tests in order to focus on what I could do better next time ($\bar{x} = 4.00$)
- I make use of the feedback I get to improve my learning ($\bar{x} = 3.28$)
- I look at what I got wrong or did poorly on to guide what I should learn next ($\bar{x} = 3.27$)
- I use mandated tests to take responsibility for my next learning steps ($\bar{x} = 2.99$)
- I use mandated tests to identify what I need to study next ($\bar{x} = 2.66$)

Examining the subfactor student learning, showed students were more likely to pay attention to their results, but disagreed that mandated assessments were a means to help them

identify what they needed to study next and/or assist in taking responsibility for their next steps in learning. As for the next subfactor, “Teaching”, participants disagreed that mandated test helped with tracking and examining students’ academic progress. For instance:

- Mandated tests help teachers track my progress ($\bar{x} = 3.91$)
- Mandated tests are a way to determine how much I have learned from teaching ($\bar{x} = 3.20$)
- A Mandated test is checking off my progress against achievement objectives or standards ($\bar{x} = 3.61$)
- My teacher uses mandated test results to help me improve ($\bar{x} = 2.86$)
- Teachers use my mandated test results to see what they need to teach me next ($\bar{x} = 3.03$)
- Mandated tests show whether I can analyze and think critically about a topic ($\bar{x} = 3.31$)

Further examination of the subfactors on the disagreement side of the scale showed students disagreed that mandated assessments contributed to the teacher’s “teaching”. The disagreement side also revealed participants showed even stronger disagreement towards mandated assessments and the ability to support “student learning”, “school quality” and “future education/employment”.

In conclusion, student responses showed an undesirable outlook towards the significance of testing mandates. Advocates supporting mandated tests would not want students in agreement with the “Irrelevance”, rather these proponents would like to see “Affective Benefits”, “External Attribution” and “Improvement” on the agreement side of the scale. However, that was not the result. “Affective Benefits”, “External Attribution” and “Improvement” had mean scores below the midpoint of the scale ($\bar{x} < 4.0$) indicating it was on the disagreement side. Likewise, the “Irrelevance” conception, where a high score suggests a negative sense towards testing, was above the midpoint ($\bar{x} > 4.0$) and to the side of the scale representing agreement.

Qualitative Data: Open-Ended Questions

Next, in order to analyze the open-ended questions, similar topics were clustered together into themes emerging from the students' responses. Rossman & Rallis (1998) described this coding process as "organizing the material into chunks or segments of text before bringing meaning to information" (p.171). After categorizing the data, each theme was then labeled with a common term that was frequently expressed through the participants' responses. Creswell (2009) describes the interpretation of the results, "as a means for the researcher to draw conclusion from the result for the research question" (p. 152). Therefore, the qualitative analysis for this study utilized the language and words used by students to express how they perceived mandated assessments.

Students' Perceived Benefits from Testing Mandates

Of the 360 high school students surveyed, 294 students responded to the open-ended question, "Overall, what do you see as the benefits to you as a high school student for having taken mandated tests in school?" From these responses, three themes emerged from the student responses: (1) No benefit, (2) how test may help and (3) non-student benefits.

What was clear from the data when subjects were asked about the benefits was that the strongest response was little to no benefits. Out of the 294 responses, 132 students reported "little to no benefit" for having taken mandated tests, which had the most participants (n =132) out of all the themes categorized from the benefits open-ended question. This first and most dominant theme meant students perceived no value and found very little significance for having completed these exams. For instance:

- “I honestly don’t see any benefits for taking mandated test. They take up time and I don’t even get to see my results and what I got wrong or right.”
- “I really don’t see any because a student can be having an off day and a test doesn’t really show how good they really are at something.”
- “There are no benefits and I click random ones sometimes.”
- “I don’t really see many benefits. Ever since they have started they have brought a boring, gloomy atmosphere and never seem to help me with anything.”
- “I honestly don’t. I’m one of the more lucky people who tend to test well, however many collapse under the stress of testing. Many whom may be quite intelligent people, just unable to deal with the stress.”
- “I don’t see much use to the mandated test it is just a test we take every once in a while. We take, and it is boring and drastically reduces are readiness to learn. I do not know what they go to and have little impact on my education from my viewpoint.”
- “Absolutely nothing, they take time out of the day where I could actually be learning something. If I’m being honest half these scores aren’t accurate because a lot of people stop trying after the first 30 questions, I know I do. After question 30 it becomes a guessing game until it’s over.
- “The benefits for mandated test are limited. I don’t feel as though I get anything out of them and quite frankly, they feel like a waste of time and a step back in productivity.”
- “For me personally, mandated test are just overall extremely stressful and can make me look poorly on myself. The 50+ questions are tedious and consume time that could be used for students. I have never learned something new from a mandated test; they do not show how I got answers wrong and show me how to improve. I doubt teachers use them to better the learning experience because teachers aren’t the ones who make the tests. It’s the system.”

It was evident by these statements that students perceived no benefit for having participated in mandated assessments. This theme connected to the “Irrelevance” conception on the SCoA-VI. Under the “Irrelevance” conception, participants viewed mandated tests as “unfair/bad” or “ignored”. For example, a student stated that “I don’t really see many benefits. Ever since they have started they have brought a boring, gloomy atmosphere and never seem to help me with anything”. Another student further described the benefits as, “Absolutely nothing”. These statements echoed the “Irrelevance” students perceived to have towards these testing mandates concerning their learning and classmates. Participants commented on the amount of time the exams took, which caused them to want to “click random ones” or “guess”. As a result,

subjects expressed the lack of effort and neglect they had for their testing performance. These particular strategies described by students can be very detrimental to the accuracy of the results produced from these testing mandates.

However, not all responses from the survey question concerning benefits were negative. The second theme identified how mandated assessments may help students. Out of 294 responses, 130 participants described experiences that could potentially enhance academics and/or assist with their personal development. These components included; (1) identifying personal academic growth and comparing with other students (2) developing test-taking skills and understanding the testing environment, and (3) assisting with class placement. As for identifying personal academic growth and student comparison responses indicated:

- “Being able to measure my learning progress and see how I improved over time.”
- “I see it as a general benchmark of where students are compared to each other.”
- “A way to see general knowledge but at the expense of others being judged by others.”
- “I like to see my progress and compare it to my classmates. It makes me more competitive.”
- “To see how I stack up on this test to other people.”
- “You get a better idea of where you are in terms of intelligence to other kids in your grade. Obviously, this isn’t important, it’s just interesting.”
- “Bragging rights with friends about higher scores.”

The next component focused on the development of test-taking skills and understanding the testing environment. The participants shared:

- “It helps me practice for ACT/SAT and college prep.”
- “Mandated test can help prepare students for taking test in the future. They can teach students how to properly use test-taking skills.”
- “The only use for mandated tests are the ones that in the end will help us get better ACT scores so that we might have actual hopes of getting into a good college.”
- “Mandated test prepare me for more rigorous test taking in college and assess my test-taking skills.”
- “I have learned test-taking skills but not any real usable skills as we never get our results back telling us which problems we specifically got right or wrong.”

The third component described mandated assessments as a means to assist with class placement. For instance, students reported:

- “It shows where I am compared to other students and usually where I should be placed (honors, AP, etc.).”
- “The score informs the tester how they are progressing in a class and which classes they should take in the future.”
- “The overall benefits of taking mandated tests, are that learning professionals have a better understanding of what information their students are retaining. What their future classes should focus on more. In addition to class placement and where our academic goals are aimed towards.”
- “ I see minor benefits for high school students as measuring academic progress but more importantly they help with class placement which helps with attaining more college credit in the long run.”

In conclusion, how these exams may help students did not receive as many responses from those participants citing no benefit. Despite the overall negative rating toward “Improvement”, students did report some benefits such as; identifying academic strengths and weaknesses, test-taking skills, test environment, and class placement. For example, one student reported that, “It shows where I am compared to other students and usually where I should be placed (honors, AP, etc.).” Another stated, “It helps me practice for ACT/SAT and college prep.” Several of these comments, highlighted the subfactor “future education/employment” within the “External Attribution” conception. Subjects reflected on mandated assessments as a means to prepare students for more rigorous academic testing in their future, acceptance into higher academic courses at their school, and championing for college recruitment. These perceptions focused on the extent that mandated assessments can lead to greater prospects/opportunities for students later on within their academic career.

The last theme that surfaced from the question, “Overall, what do you see as the benefits to you as a high school student for having taken mandated tests in school?” was about the

benefits these assessments had for others which was identified as non-student benefits. This theme had the least amount of responses compared to participants reporting “no benefit” and “may have benefits for students”. 32 responses out of 294 described benefits that included stakeholders such as: (1) administration and government officials, (2) teachers, and (3) parents. The first component provides the perceptions students believe administrators and government officials receive from their participation in testing mandates. For example:

- “The benefits are for the state and government to see what all you have learned at your school over the year.”
- “The teachers/state can see how different schools are ranked and if students typically miss in a certain category.”
- “Which schools need better funding and staff.”
- “A benefit from mandating testing would be that it is an easy way to see the overall improvement of the school as a whole.”
- “Standardized testing while it shouldn’t be used to judge the intelligence of an individual, it does provide extremely helpful insight to understand achievement gaps and even wealth gaps. Without them, these gaps could go unnoticed and unopposed.”

Although rated negatively, a small percentage of the comments shared pointed to the subfactor “school quality” and being able to judge the value or worth of the school. For example, one student stated, “None, just to see how the school is doing.”. Another subject commented, “I don’t see a single benefit for students, although the schools use it to show the academic achievement from one school to another.” While participants regarded their own personal benefits negatively, some of the reflections shared recognized non-student benefits such as “school quality”, within the “External Attribution” conception.

The second component describing non-student benefits, included the benefits teachers gain. Specifically:

- “It helps teachers know the average knowledge of the students so they can teach what they need to know and move one.”

- “The overall benefits of taking mandated tests, are that learning professionals have a better understanding of what information their students are retaining. What their future classes should focus on more.”
- “The benefits of mandated tests is to give feedback to teachers as to how well students are progressing through the years.”
- “Teachers being able to see how well they are teaching.”

The third component describing benefits for others included parents/guardians. Such as:

- “It shows my parents how well I’m doing in school compared to my peers.”
- “It lets my parents know how well I can test under certain circumstances.”

In conclusion, three themes emerged from the qualitative question, “Overall, what do you see as the benefits to you as a high school student for having taken mandated tests in school?”

These themes were (1) no benefit, (2) test may benefit students and (3) non-student benefits. The largest response rate identified no benefit. It was apparent that students had a difficult time naming and/or describing any benefit for having participated in mandated assessments. The next largest response rate offered insight into what students perceive as their benefits for having taken mandated assessments. These benefits reflected on academic progress, test-taking and class placement. The third theme offered insight into non-student benefits such as; parents, schools and government personnel. Overall, it was evident from the reflections, students had difficulty seeing any value for taking mandated assessments, though they did provide some areas of benefit.

Students’ Perceived Drawbacks from Testing Mandates

For the second qualitative question, “What do you see as the drawbacks for you as a high school student having taken mandated tests in school?” 285 out of the 360 participants provided

responses. The perceived drawbacks described by students were categorized into three themes (1) Time, (2) Test content and feedback, and (3) physical symptoms.

The first theme time refers to the amount of time taken by the process, activity, or the person doing it. 130 out of 285 responses described time as the biggest drawback. The components for the theme, time was (1) length of the exam and (2) the loss of class time and learning. The responses shared by participants concerning the length of the exam stated:

- “Mandated tests are extremely time consuming. Even after the testing period is over, students are tired and unmotivated to learn.”
- “Having a huge 2-hour test, students will start doing their best then get bored and just guess the rest.”
- “Well, some drawbacks are your eyes being strained after staring at a screen for a really long time, which is not healthy and sometime causes a headache. We don’t even get to see our scores on some of them and when we do it is numbers that I don’t even know what they mean, so I feel like I just wasted the past hour of my life.”
- “Each test takes about 2 hours of class time and can cut into other classes as well. It doesn’t count as a grade and many students guess or give up, so there’s really no point.”

Along with the length of testing mandates, the loss of class time and learning was another conception within the theme of time. Participants described these drawbacks as:

- “It ends up just being a day taken away from learning. The teachers don’t care about it. The kids don’t care about it. No one in the room cares about these stupid long tests that are over 50 questions that will take the entirety of at least one if not two class periods.”
- “It is a waste of time that I could have spent actually learning.”
- “We would take a week to two weeks to just focus on our assessment coming up. The teachers seemed so motivated to get good test scores that they missed out on teaching us the other stuff. They were only teaching to the assessment.”
- “It is a waste of time. Instead of going more in-depth on a topic, we have to set aside time for a test that doesn’t do much.”
- “Takes valuable time away from learning.”

In summary, the drawback time showed to have a major bearing on how students perceived mandated assessments. The responses provided by the students reinforced the high

ratings on the “Irrelevance” conception from the SCoA-VI. Due to the amount of time, many subjects felt testing mandates were “unfair/bad” and often “ignored”. For example, one student stated, “Having a huge 2-hour test, students will start doing their best then get bored and just guess the rest.” Another affirmed that, “Each test takes about 2 hours of class time and can cut into other classes as well. It doesn’t count as a grade and many students guess or give up, so there’s really no point.” The two components within time emphasized concerns for the length of the test and the disappoint for lost class time/learning. Thus, creating perceptions among the participants that mandated assessments are “unfair/bad” and “ignored”.

The next theme that emerged from the subjects’ responses concerning drawbacks was the exam’s content and feedback. The exam’s content refers to how the test is set-up and the material/learning standards measured. The test feedback is information participants receive after completing the exam. Of the 285 responses, 127 subjects reported that the content and structure of the test caused it to be a significant obstacle when completing these exams. The components within this theme are (1) test content and (2) feedback. Descriptions provided by the subjects’ concerning test content stated:

- “Most mandated tests are the same thing but worded differently.”
- “They have questions in them we have never seen before. Sometimes other classes are not as far ahead and when everyone takes them at the same time it is unfair.”
- “Having taken mandated test starting at a very young age, you start to lose interest and motivation for the same test year after year.”
- “Drawbacks of mandated tests include biased scoring on irrelevant topics that most students are not currently studying. I prefer hands on learning and making real world applications.”
- “They are not very efficient because they are about memorization which is a skill that is not really needed today.”
- “They often include question formats that are confusing, even if you generally know the material.
- “It’s a predetermined set of questions and they don’t give us the progress we need.”

Along with test content, the other component emphasized was the lack of feedback students gained after completing testing mandates. Specifically:

- “They don’t really help us learn more or improve our knowledge.”
- “They really do not show what is learned because the test questions are non-related to curriculum.”
- “The main drawback in taking mandated test is that we don’t learn anything. For these test to actually yield some benefits the testing process needs to be reformed.”
- “They don’t give us the progress that we need.”
- “I don’t think they show what we have learned. I think your grades in classes should determine that.”

Similar to the first drawback, time, test’s content and feedback also connected to the high rating \bar{x} from the “Irrelevance” conception on the SCoA-VI. For example, one participant stated, “Having taken mandated test starting at a very young age, you start to lose interest and motivation for the same test year after year.” Another added, “It feels we are being taught to test well, not to retain information.” Many of the responses described boredom and uninterest in the testing process/experience because of the exam’s content.

Likewise, participants felt the feedback from the test offered little advice on what or how to do better. For instance, one student stated, “They don’t really help us learn more or improve our knowledge.” The test content and feedback component reiterated the “Irrelevance” conception that assessments are perceived as ‘unfair/bad’ and “ignored”.

The third theme that emerged from the question “What do you see as the drawbacks for you as a high school student having taken mandated tests in school?” was the physical symptoms reported by students. 28 out of 285 responses reported stress, anxiety, and low self-esteem, and resulted in students feeling physical symptoms from having participated in these testing mandates. These two related components are (1) psychological symptoms and (2) physical manifestations.

The first component, the psychological symptoms of stress, anxiety and low self-esteem was described by participants as :

- “They make me feel bad about myself and make me feel stupid.”
- “It makes me feel like I don’t know a lot of stuff. Also, it makes me self-conscious when people want to compare scores. I also feel pressured to improve every time.”
- “Higher stress levels and being taught that a test score determines my whole future.”
- “Taking tests causes me anxiety even though I am usually well prepared.”
- “They stress people out and make it hard to focus.”

These statements emphasized the amount of stress and anxiety students perceived to encounter when completing mandated assessments and the negative effect it had overall on their self-esteem. As for the next component, physical manifestation of stress, anxiety and low self-esteem, participants stated:

- “I don’t get enough sleep because I’m either nervous or studying. They aren’t fun at all.”
- “I get stressed out, I fear my grade, I can’t stay focused and feel drained after testing.”
- “Stress and worries causes students to have problems with other assignments.”
- “Lots of added stress, can tank grades, generally creates a stressful and sometimes slightly hostile learning environment.”

The components portraying students’ feelings and signs of stress, anxiety and low self-esteem had clear adverse effects on the students. For example, one student stated, “Higher stress levels and being taught that a test score determines my whole future.” Another participant added, “It makes me feel like I don’t know a lot of stuff. Also, it makes me self-conscious when people want to compare scores. I also feel pressured to improve every time.” These statements made by subjects portrayed a “stressful” and “sometimes hostile learning environment” which resulted in students feeling “bad” or “stupid” and sometimes further physical symptoms.

In summary, the perceived drawbacks from students for having taken mandated tests was the (1) time, (2) test content and feedback, and (3) physical symptoms. The drawbacks described by participants overwhelmingly highlighted the “Irrelevance” conception on the SCoA-VI, which had the highest scores of all on the sub-scales. The “Irrelevance” conception continued to be the main conception focused on by students within the qualitative data. The other conception from the drawbacks noted by participants underscored the negative ratings from the “Affective Benefits” conception. Participants in the study voiced less favorable comments towards their own personal liking or enjoyment towards mandated assessments.

Student Opinions Towards Mandated Assessments

For the last open-ended question, “Do you have any other opinions about mandated test you would like to share?” Out of 177 responses, 65 students reported no other opinions, which left 112 student opinions remaining. From those responses, three themes emerged: (1) Test Process (2) Measurement Accuracy and (3) Test Content and Feedback. Time, test content and feedback were themes described earlier in the study when subjects were asked about the benefits and drawbacks for having participated in mandated assessments. It was clear participants felt compelled to reiterate these themes again.

Of the 112 opinions shared, 42 students expressed opinions concerning the test process and the amount of time it took to complete these testing mandates. Time was a reoccurring theme as it was addressed in the perceived drawbacks towards having completed mandated assessments. The component time refers to the amount of time taken by the process, activity, or the person completing the task. Students elaborated on the testing process and their concern about the amount of time by stating:

- “I think it is necessary to have the test to see if students are learning the material. But the test could be much shorter and possibly less complex so that it is more enjoyable for the students.”
- “Sometimes you just need to lessen the amount of questions, it’s a tad overwhelming especially when 5 or 6 ask the same exact question just phrased differently. It feels a little tedious.”
- “As a very energetic and easily distracted person, mandated tests are a pain. It is nearly impossible for me to sit down and take a 50+ minute test without losing interest.”
- “High school students take so many mandated tests that it gets tiring. Since students take them a lot, students dread the days they have to take them and may do badly on them because they don't see the point of doing so many.”
- “I think mandated tests could be spaced out through a few days more and not have as much weight put on them.”

In addition to these statements, participants shared thoughts for improving the test process such as, “lessen the amount of questions”, “split into smaller chunks”, or “space out through a few days”. The objections subjects had towards the current testing process again related to the high ratings of the “Irrelevance” conception on the SCoA-VI. Further inquiry regarding students’ opinions revealed participants were not necessarily opposed to taking mandated tests, rather they offered recommendations for making the exam more relevant by adjusting the time or limiting the amount of questions to increase student engagement, focus, and attentiveness during the testing process.

The second theme from the opinions communicated by the participants was about the measurement accuracy. Subjects shared their distrust for how effectively mandated tests evaluated their learning and understanding. Of the 112 responses, 38 students reported concerns towards the exams accuracy and reliability. Measurement accuracy is described as how close a measured value is to the actual value. Statements concerning the accuracy of mandated assessments in detail were:

- “Not sure if they are a good indication of my learning .”
- “It’s not accurate and causes students to stress out.”

- “I feel that they are pointless because they do not measure intelligence. They measure ones skill to remember.”
- “I have friends who are really smart but are terrible test takers and I feel like they don't get accurate results because of that. Also, I think friends can spend too much time preparing to take the test than understanding the material.”
- “I do not like them. In my mind, the only thing they test is how well a student can take a test. They're also only based off of a student's performance in one or two days, instead of over an entire semester like it should be.”
- “Mandated tests are not an accurate assessment of intelligence and aren't necessary in middle and high school.”
- “Personally, my score on standardized tests fluctuates a lot and sometimes, it's not reflective of how much I've learned. I dropped 9 points on the MAP test at the end of my freshman year on the reading section, even though I had an excellent teacher and learned a lot. My score did not reflect that accurately. Not only is it not accurate, it breeds competition among students, comparing scores after each test as an unhealthy way to "measure up" each other.”

These statements showed unease among the students for how well they perceived mandated tests measured their academic understanding and overall educational performance. One student reported concerns that students “spend too much time preparing to take the test than understanding the material.” Another described the measurement as, “how well a student can take a test”. The doubts portrayed by students concerning the assessments ability to accurately measure a students' learning and understanding related to the negative ratings towards the “Affective Benefits” from the SCoA-VI. How students perceive these mandates to be “helpful” or lack thereof aligns to the “Affective Benefits” subfactor. In conclusion, the participants reflected their apprehension towards the measurement accuracy and reliability testing mandates have for finding a student's true level of learning or mastery.

The third theme shared from the opinions of students towards mandated assessments was testing content and feedback. 32 out of 112 responses addressed the this theme making it the weakest of the three sub-themes. Test content and feedback means the exam's material and the explanation of the findings from the results of the assessment. The components for this theme

was (1) Test content and (2) Test feedback. The descriptions provided by students are concerning test content were:

- “I don’t like mandated test. I personally do better with hands on work, it helps teach me and I learn better.”
- “I wish we integrated them more into school. It feels like they come and go and nothing is done about them. We don’t prepare and are usually told about them the day before it happens.

For the next conception, test feedback, students reported:

- “They don’t even tell you what you got wrong or how to fix it so how would that help me study?”
- “They should at least tell us what our score actually means or show us the answers to what we got wrong or commonly missed questions, so we can learn from them.”
- “Students don’t use the information after the test. I don’t think many teachers do either. We just take them and then they’re done. It’s very old fashioned in my opinion and it’s not how students learn.
- “I really wish we had feedback on how we do on the mandated tests. They don’t seem to affect my future with college, jobs, or grades so I don’t pay attention to the score that much. I am proud of myself to see my score improve but personally I don’t think mandated tests encourage me or my peers.”
- “I need an in-depth description of what the scores mean. I need to be able to explain it to my parents.”

Students provided suggestions for improvement concerning the exam’s content and feedback. For example, one student stated, “I wish we integrated them more into school. It feels like they come and go and nothing is done about them. We don’t prepare and are usually told about them the day before it happens.” As for the feedback, a participant suggested that the test could, “provide answers to what students get wrong and/or give the commonly missed questions, so they can learn from the test”. Another stated, “I need an in-depth description of what the scores mean. I need to be able to explain it to my parents.” Overall, the subjects responses revealed the desire students generally have to want to perform well. The recommendations for improvement given was the need for additional information after the exam so students are able to learn from the test and/or substantiate their score/results.

Research Question 1 Summary

The students' perspectives offered insight into the fundamental purpose of the study: to investigate and understand high school students' perceptions towards mandated assessments. Research question one framed and guided the overall purpose for understanding how students perceive testing mandates:

A) Do high school students feel testing mandates improve teaching and student learning (Improvement)?

Participants were more likely to “disagree” that mandated tests helped with improving the teacher ($\bar{x} = 3.32$) or self ($\bar{x} = 3.22$). According to the 7-point Likert scale, the “Improvement” conception was on the disagree side of the scale. In addition to the SCoA-VI, open-ended questions also reflected uncertainty for how the subjects perceived testing mandates improved teaching and learning. It was evident by the qualitative responses, subjects wanted the opportunity for “Improvement”, but disagreed with the current testing process. Students offered insight and suggestions on ways to improve the testing experience such as less questions, dividing the test into smaller sections, etc.. However, under the current testing circumstances participants disagreed that these mandates offered any kind of “Improvement” for both teacher and students.

B) Do high school students feel mandated assessments measure the school's quality and predict the student's future in education and employment (External attribution)?

On the SCoA-VI subjects tended to “disagree” with how mandated assessments measured “External attribution” ($\bar{x} = 2.79$). Students were likely to “disagree” that mandated assessments measured the school's quality ($\bar{x} = 3.08$) and were even more likely to “disagree” that it predicts a student's future in education and employment ($\bar{x} = 2.65$). Students supported this belief further

by expressing their opinions within the qualitative responses, questioning whether the accuracy of the results produced from these testing mandates were a true and reliable measure. In general, high school students do not perceive mandated assessments measure the quality of the school and predict a student's future in education and employment.

C) Do high school students feel these assessments are beneficial to themselves and classmates (Affective benefits)?

High school students perceived mandated assessments provided no benefits to themselves and classmates. On the 7-point Likert scale, "Affective Benefits" was on the disagree side of the scale, near the endpoint "completely disagree". "Affective Benefits" had the lowest mean score out of all four conceptions on the SCoA-VI. It was also apparent in the qualitative data when participants were asked about the benefits and drawbacks for having taken mandated assessments. For benefits, subjects cited, "little to none" and for drawbacks concerns about the exam's time, content, and feedback were provided. Students perceived mandated test offered no "Affective Benefits" for them and classmates.

E) Do high school students feel mandated assessments are unfair and ignored (Irrelevance)?

High school students agreed when asked if mandated assessments were "unfair/bad" ($\bar{x} = 4.36$) and "ignored" ($\bar{x} = 3.93$). The average response for "Irrelevance" ($\bar{x} = 4.20$) was above the midpoint and on the agree side of the scale. Further examination of the interval levels from all four conceptions on the SCoA-VI showed "Irrelevance" had the highest mean score resulting in this being the only conception on the agree side. Furthermore, the students' qualitative responses showed drawbacks to be the amount of time, stress, and anxiety felt as well as the uncertainty for how accurate these testing mandates are for their learning/understanding.

Research Question 2: Do high school student perceptions toward mandated assessments differ based on selected demographic characteristics including gender, race, primary language, parent education, GPA, and plans after high school?

To answer the second research question, a nominal scale was assigned to each participant's demographic characteristic (gender, race/ethnicity, primary language, parent education, high school GPA, and plans for after high school) with a number to serve as a label with no quantitative value. T-tests were used to help examine the difference between the independent variables with two values: gender (female or male), primary language (English or Spanish), and (Education high or low) to the dependent variables (overall sample mean score and the mean scores for the four conceptions). These items are identified as A, C, and D on the demographic information questionnaire (Appendix D).

Table 8 below provides a further description of the results concerning demographics and the statistical significance to their perspective regarding mandated tests. The p-value was set at α 0.05 to determine if that data were statistically significant.

Table 8: Student Perceptions Towards Mandated Assessments Based on Gender

Conception	Mean: Gender		t- stat	t- critical	P-value
	Female n = 208	Male n = 152			
1) Improvement	3.18	3.40	-1.61	1.97	0.11
2) External Attribution	2.64	3.00	-2.63	1.97	*0.01
3) Affective Benefits	1.92	2.19	-2.37	1.97	*0.02
4) Irrelevance	4.32	4.04	2.15	1.97	*0.03

* *Significant at $p < .05$*

Table 8 presents the results concerning student perceptions based on gender, the p value was less than the α 0.05 for “External Attribution”, “Affective Benefits” and “Irrelevance”. This indicated that students’ gender was statistically significant for these three conceptions. Females tended to have a greater level of disagreement than males with the perception that testing

mandates were for “External Attribution” and “Affective Benefits”. Females also were likely to agree more about the “Irrelevance” for these testing mandates than males. However, both male and female students felt little impact on how mandated assessments provided “Improvement” to the student’s learning and teacher.

Table 9: Student Perceptions Towards Mandated Assessments by Language

Conception	Mean: Language		t- stat	t- critical	P-value
	English n = 345	Other n = 15			
1) Improvement	3.27	3.49	-0.66	1.97	0.51
2) External Attribution	2.77	3.38	-1.77	1.97	0.08
3) Affective Benefits	2.03	2.35	-1.16	1.97	0.25
4) Irrelevance	4.23	3.53	2.16	1.97	*0.03

* *Significant at $p < .05$*

Table 9 displays results by language. For student perceptions based on language, only one conception, “Irrelevance” showed to be statically significant. Participants who identified their primary language as English were more likely to agree with the “Irrelevance” for having completed these testing mandates compared to other participants whose primary language was “other”. It is important to note that only 15 students from the study reported “other” as their primary language which does not provide a good representation for that particular demographic. As for “Improvement”, “External Attribution”, and “Affective Benefits” all three conceptions showed no statistical significance.

Table 10: Student Perceptions Towards Mandated Tests Based on Parent/Guardian Education

Conception	Mean: Level of Education		t- stat	t- critical	P-value
	High n = 287	Low n = 73			
1) Improvement	3.04	3.58	-3.46	1.97	*0.00
2) External Attribution	2.75	3.06	-1.74	1.97	0.08
3) Affective Benefits	1.96	2.35	-2.79	1.97	*0.01
4) Irrelevance	4.26	3.96	1.85	1.97	0.06

* *Significant at $p < .05$*

Table 10 shows the results of the analyses based on education level. For parent or guardians' highest level of education, two factors were compared high education level (meaning a parent or guardian has attained an associate degree or higher) and a low education level (meaning they have attained some college/technical experience, high school graduation or less). From the t-test, two conceptions were shown to be statistically significant. "Improvement" and "Affective Benefits". Students in the study who had a parent or guardian with low levels of education tended to agree more with the perception that mandated assessments provided "Improvement" and "Affective Benefits" for themselves and others than students from more higher educated parents/guardians. This was the only demographic in the study to show a statistical difference with the conception "Improvement". As for "External Attribution" and "Irrelevance" these conceptions showed no statistical significance when comparing the level of education attained by a parent/guardian.

Next, ANOVAs were performed to determine if the demographic variables: race/ethnicity (non- Hispanic, White, Black/African American, Hispanic/Latino or other), the students high school GPA (3.5 and above Mostly A's, 2.5-3.5 Mostly B's, 1.5-2.5 Mostly C's, and Other) and the student's plan after high school (4-year college, 2-year college or technical training, Full time job or Take time off "Gap year") impacts the overall sample mean score and the mean scores for

the four conceptions in the study. These question items are B, E and F on the demographic questionnaire (Appendix D).

Table 11: Student Perceptions Towards Mandated Assessments Based on Race/Ethnicity

Conception	Mean: Race/Ethnicity				F-value	F-critical	P-value
	White n = 292	Black n = 18	Hispanic Latino n = 21	Other n = 29			
1) Improvement	3.20	3.64	3.74	3.60	2.42	2.63	.07
2) External Attribution	2.63	3.16	3.27	1.96	3.28	2.63	*.02
3) Affective Benefits	2.08	2.66	2.65	2.63	4.92	2.63	*.00
4) Irrelevance	4.24	4.06	3.82	3.84	1.68	2.63	.17

* *Significant at $p < .05$*

As noted in table 11, based on the race/ethnicity of the subjects surveyed, “External Attribution” and “Affective Benefits” were statically significant. As for “External Attribution”, students who selected “other” and white as their race/ethnicity were more likely to disagree that mandated assessments have “External Attribution” compared to black and Hispanic or Latino students. On the other hand, white students were more likely to disagree than black, Hispanic or Latino, and other students about the “Affective Benefits”. Students of color appeared to see more benefits for having participated in these testing mandates than others. As for “Improvement” and “Irrelevance” these conceptions showed no statistical difference based on the students race/ethnicity.

Table 12: Student Perceptions Towards Mandated Assessments Based on Student GPA

Conception	Mean: GPA			F- value	F-critical	P-value
	3.5 and Above n = 238	2.5-3.4 n = 93	Other n = 29			
1) Improvement	3.17	3.04	3.46	1.35	3.02	0.26
2) External Attribution	2.77	2.85	2.02	0.12	3.02	0.89
3) Affective Benefits	1.99	2.12	2.08	0.52	3.02	0.59
4) Irrelevance	4.18	4.04	4.16	0.42	3.02	0.66

* *Significant at $p < .05$*

Table 12 displays the students' GPA, which showed to have no statistical significance with all four conceptions: "Improvement", "External Attribution", "Affective Benefits" and "Irrelevance". According to the data, all GPA categories showed high mean scores for "Irrelevance" and low mean scores for "Affective Benefits". No matter the students GPA, all students were more likely to agree with the "Irrelevance" and disagree with the "Improvement" "Affective Benefits", and "External Attribution" for having participated in testing mandates.

Table 13: Student Perceptions Towards Mandated Test based on Plans After High School

Conception	Mean: Plan After HS			F- value	F-critical	P-value
	College or Technical Training n = 283	Job or Gap Year n = 26	Not Sure n = 51			
1) Improvement	3.24	3.65	3.52	1.99	3.02	0.14
2) External Attribution	2.82	3.25	3.20	2.57	3.02	0.08
3) Affective Benefits	2.07	2.50	2.48	4.38	3.02	*0.01
4) Irrelevance	4.14	3.70	4.25	1.65	3.02	0.19

* *Significant at $p < .05$*

According to the data in table 13, "Affective Benefits" was statistically significant. 9th and 10th grade students who planned to attend college or receive technical training after high school were more likely to disagree with the perception that testing mandates had "Affective Benefits" from other students who were either unsure, taking a gap year or planning to work after high school. As for "Improvement", "External Attribution", and "Irrelevance" these

conceptions had no statistically significance based on students plans for after high school.

Similar to the last demographic, the conception “Irrelevance” showed the highest mean scores for plans after high school in all three categories.

Research Question 2 Summary

In summary, student perceptions towards mandated assessments differed based on certain demographic characteristics when compared to the four conceptions from the SCoA-VI. First, the gender of students had the greatest statistical significance compared to other demographic information studied within the SCoA-VI conceptions. Gender showed statistical significance with “External Attribution”, “Affective Benefits” and “Irrelevance”. Female students were more likely to disagree with “External Attribution” and “Affective Benefits”, but agree more than males about the “Irrelevance” for these testing mandates. The next demographic examined was language, which showed “Irrelevance” to be statically significant. Students who selected English as their primary language were more likely to agree with the “Irrelevance” compared to students who selected “other” as their primary language.

The third demographic observed was the level of education a parent/guardian had attained. Two conceptions appeared to show significance “Improvement” and “Affective Benefits”. Subjects who had a parent or guardian with low levels of education tended to agree mandated tests offered “Improvement” and “Affective Benefits” than did those with parents/guardians having higher education levels. Parent/guardian education was the only demographic to show statistical significance with the “Improvement” conception.

The next demographic examined was race/ethnicity. Students who selected “other” were more likely to disagree that mandated assessments have “External Attribution” compared to

students who identified themselves as white, black, Hispanic or Latino. Furthermore, white students were more likely to disagree with “Affective Benefits” than students who were identified as black, Hispanic/Latino, or other.

The fifth demographic considered was students’ GPA which had no statistical significance among any of the four conceptions. Students’ GPA had the least amount of statistical significance out of all of the demographic information studied.

The last demographic examined was the plan students had for after high school. Students who planned to attend college or receive technical training after high school were more likely to disagree with the “Affective Benefits” than students who were unsure, taking a gap year or planning to work after high school. In the next chapter the results are further examined and include interpretations of the findings, limitations of the study, and implications and recommendations for future research involving testing mandates.

Chapter 5

Conclusions

The purpose of this study was to identify how 9th and 10th grade students perceive mandated assessments using the SCoA-VI survey and open-ended questions. Chapter 5 includes a summary of the findings, conclusions, and implications for future research in the areas of testing mandates.

Summary of the Findings

The first research question investigated what high school students' perceptions were towards mandated testing on improvement, external attribution, affective benefits and irrelevance as measured by Brown's SCoA-VI survey. According to the SCoA-VI, 9th and 10th grade students "disagreed" that mandated assessments provided improvement for "teaching" ($\bar{x} = 3.32$) and "student learning" ($\bar{x} = 3.22$). Students tended to "disagree" with statements that stated, "mandated assessments help me identify what I should study next" ($\bar{x} = 2.66$) and "teachers use test results to help me improve" ($\bar{x} = 2.86$). Further examination of the "Improvement" conception showed students were neutral about the statement concerning "I pay attention to my results on mandated tests in order to focus on what I could do better next time" ($\bar{x} = 4.0$). Likewise, on the subfactor "teaching", students expressed neutral perceptions regarding teachers being able to "track progress" ($\bar{x} = 3.91$).

High school students "disagreed" that mandated assessments provided "External Attribution" ($\bar{x} = 2.79$). From the SCoA-VI, students perceived mandated tests did not effectively measure the "school's quality" ($\bar{x} = 3.08$) and predict "future education/employment"

($\bar{x} = 2.65$). On the SCoA-VI, students disagreed more strongly about the statements referencing “future performance” ($\bar{x} = 2.39$) and “job/career” ($\bar{x} = 2.57$). Students also disagreed with mandated assessments being able to determine “how well the school is doing” ($\bar{x} = 3.42$) or showing “parents/guardians what they learned” ($\bar{x} = 2.89$).

Ninth and tenth grade students perceived that mandated assessments did not provide “Affective Benefits” ($\bar{x} = 2.04$). From the SCoA-VI, 7 of the 8 statements regarding “Affective Benefits” had the lowest mean scores out of all the statements assessed on the survey. Students showed the greatest disagreement towards “Affective Benefits” than they did with any of the other conceptions on the SCoA-VI. As a result, students disagreed that mandated assessments were “enjoyable to self” ($\bar{x} = 1.83$) and “helpful to classmates” ($\bar{x} = 2.10$).

According to the SCoA-VI, 9th and 10th grade students “agreed” about the “Irrelevance” ($\bar{x} = 4.20$) of mandated assessments. Six of the eight statements concerning “Irrelevance” had the highest mean scores out of all the statements examined on the survey indicating it was on the agree side of the scale. Further analysis showed 9th and 10th grade students “agreed” mandated assessments were “unfair/bad” ($\bar{x} = 4.36$). For example, student tended to agree with statements indicating these exams had “little impact on my learning” ($\bar{x} = 4.79$) and “not very accurate” ($\bar{x} = 4.57$). Students were also more neutral regarding whether or not students “ignored” ($\bar{x} = 3.93$) mandated assessments, such as “I ignore assessment information” ($\bar{x} = 3.28$). For students, mandated assessments were not perceived as an accurate measure of their learning. Although many students reported “little impact”, students in the study did not “ignore” the tests and information provided to them.

In summary, it was evident that students disliked the current process of testing mandates. The survey results from students were not surprising, as research conducted on other

stakeholders have shown over the last two decades mandated assessments continue to be a controversial topic in the field of education (Hagopian, 2014).

In addition to the SCoA-VI, three open-ended questions were asked to 9th and 10th students regarding their perceptions on the (1) benefits, (2) drawbacks and (3) opinions towards having participated in mandated assessments. Three themes emerged from the qualitative question, “Overall, what do you see as the benefits to you as a high school student for having taken mandated tests in school?” The these were: (1) no benefits, (2) student benefits and (3) non-student benefits.

The largest response from students describing the benefits of the test was the first theme, no benefit. It was apparent that 9th and 10th grade students had a difficult time naming and/or describing any benefit for having participated in mandated assessments. Students described no benefits because of the “gloomy atmosphere” and “stress of testing”. Furthermore, they considered this testing experience a “waste of time” and a “step back in productivity”.

Unlike the first theme, the second theme did highlight some benefits for students such as “identifying growth”, “class placement”, “student comparisons”, and assisting with “test-taking skills/environment”. The third theme described benefits, not for students, but rather for other stakeholders such as; government officials, teachers, and parents. The non-student benefits that emphasized government benefits included “ranking schools”, “funding” and monitoring “achievement gaps”. For teacher benefits, students commented on educators being able to monitor “student progression” and assisting with “future courses”. Students also stated that “comparison to peers” and showing what level of “test-taker” the student was could also be potential benefits for parents/guardians.

Although a majority of 9th and 10th grade students found no benefit for having taken mandated assessments, a small percentage did emphasize some limited benefits for students and non-students. These responses concerning perceived benefits for students and non-students was consistent with findings in the SCoA-VI.

Students were also asked “What do you see as the drawbacks for you as a high school student having taken mandated tests in school?” Three themes developed from the responses by these 9th and 10th grade students. The drawbacks were: (1) time, (2) test content and feedback, and (3) physical symptoms. Time was described as a major drawback because of the “length of the test” and “loss of class time/instruction”. For example, students reported, “a day taken away from learning”, “each test takes about 2 hours of class time which cuts into other classes”, and “the amount of questions make it difficult to focus”.

Similar to the first theme, the exam’s content and feedback was also described as a major drawback. Ninth and tenth grade students responded negatively towards test content, describing it as “non-related to curriculum”, “irrelevant topics”, “same questions just worded differently” and “about memorization”. Students elaborated on the lack of real-world application and the content being unrelated to their current studies as major shortcomings for testing mandates.

Students also described their uneasiness with feedback provided after completing the exam. Students stated that some of these mandated exams, did not show a “score”, “problems missed”, and/or “reasons/explanations”. Students further commented on the lack of feedback by stating, “In my mind, the only thing they test is how well a student can take a test” and “Not sure if they are a good indication of my learning”. These perceptions showed students were skeptical about the overall value for these testing mandates due in part to the content not being applicable and limited feedback.

The third theme shared by students, describing the drawbacks, emphasized the psychological and physical manifestations such as; stress, anxiety and low self-esteem. For example, students stated, “I get stressed out, I fear my grade, I can’t stay focused and feel drained after testing” and “I don’t get enough sleep because I’m either nervous or studying”. The last drawback described by students, portrays a very “stressful” and “sometimes hostile learning environment” which has resulted in many students feeling “bad” or “stupid” and sometimes even further physical symptoms, like a “headache” or “upset stomach”.

The last open-ended question, “Do you have any other opinions about mandated tests you would like to share?” generated three themes (1) test process, (2) measurement accuracy, and (3) test content and feedback. The first theme, test process, referred to the exam’s layout and length. For example, students discussed how the exam had “5 or 6 of the same questions, just phrased differently”. Furthermore, a student expressed that having taken mandated tests at a very young age caused the person to “start to lose interest and motivation for the same test year after year.” Several of the opinions shared provided recommendations, such as; “spacing out over a few days” and “lessening the amount of questions”. The desire to decrease the number of questions could be due to the redundancy many students expressed as to why they “lost interest”.

The second theme, measurement accuracy, introduced a new theme. Measurement accuracy referred to the students’ distrust for how effectively mandated tests evaluated their learning and understanding. Students stated that mandated assessments were, “not a good indication of my learning.” and “not an accurate assessment”. High school students described the length of the exam caused them to “click random ones” or “guess”. The explanations provided by students, does raise uncertainty on whether or not these current conditions surrounding testing mandates could produce an accurate analyses of students learning and understanding.

The third theme, test content and feedback, was a reoccurring theme. Test content and feedback continued to be a major barrier identified by students as an area needing improvement. As for opinions, students described other measures such as “hands-on activities” or integrating smaller assessments into course work, so testing mandates don’t just “come and go” which can create a “one and done” mentality. As for the exam’s feedback, students reiterated their recommendations for more advice/explanation with regards to “how to fix missed questions” or “more detailed information concerning my score”. Students want feedback so the testing experience can be meaningful to them and their learning.

Although negative ratings were indicated by the students from the SCoA-VI, thoughtful insight and recommendations were also conveyed in the students’ responses on the open-ended questions. Heritage (2013) emphasized the importance for students to be viewed as equal stakeholders especially when assessments are used to improve/gauge student learning and understanding. These responses demonstrated how reflective and sincere students can be towards the testing process. Despite the negative outlook, students want to improve and they desire to have a more helpful and useful measure to assess their learning.

The second research question examined high school student perceptions toward mandated assessments based on selected demographic characteristics. The first demographic, gender, showed high school students perceptions towards mandated assessments differed on “External Attribution”, “Affective Benefits”, and “Irrelevance”. According to the SCoA-VI, female students rated “External Attribution” and “Affective Benefits” ($\bar{x} = 2.63$, $\bar{x} = 1.92$) lower than male students ($\bar{x} = 3.00$, $\bar{x} = 2.19$) in the study. Female students tended to disagree more strongly than male students when asked about statements concerning the “school’s quality”, “future education/employment”, “enjoyment for self” and “helpful to classmates”. As for the agreement

side of the scale, female students agreed more ($\bar{x} = 4.32$) than male students ($\bar{x} = 4.04$) that mandated assessments are “unfair/bad” and often “ignored”. Male students tended to have a more neutral standpoint ($\bar{x} = 4.0$) concerning the “Irrelevance” conception.

The next demographic characteristic evaluated high school students perceptions towards mandated assessments based on language. Ninth and tenth grade students who identified “other” as their primary language ($\bar{x} = 3.53$) “disagreed” with the “Irrelevance” of mandated assessments. Whereas, students that identified “English” as their primary language ($\bar{x} = 4.23$) “agreed” that mandated assessments were irrelevant, meaning they were “unfair/bad” and often “ignored”. It is evident that the difference between English and other languages students influenced their perceptions towards the relevance of these exams.

Another demographic variable, parent/guardian education was examined in this study. High school students perceptions towards mandated assessments based on parent/guardian education differed on “Improvement” and “Affective Benefits”. High school students who identified parents as having a “high” level of parent/guardian education ($\bar{x} = 3.04$, $\bar{x} = 1.96$) disagreed more strongly towards the “Improvement” and “Affective Benefits” than students who identified “low” levels of parent/guardian education ($\bar{x} = 3.58$, $\bar{x} = 2.35$). Both educational levels had negative perceptions towards “Improvement” and “Affective Benefits”. However, students with higher levels of parent/guardian education had a more negative impression that these mandates were beneficial to the student and their education.

The fourth demographic variable assessed in this study was race/ethnicity. High school students perceptions based on race/ethnicity showed significance with “External Attribution” and “Affective Benefits”. “External Attribution” examined whether mandated assessments could predict the “school’s quality” and the student’s “future education/employment” Students who

identified “other” ($\bar{x} = 1.96$) as their race “disagreed” more with “External Attribution” than “white” ($\bar{x} = 2.63$), “black” ($\bar{x} = 3.16$) and “Hispanic/Latino” ($\bar{x} = 3.27$) students. “Affective Benefits” refers to mandated assessments being “enjoyable” and “helpful to classmates”. For “Affective Benefits”, “white” students ($\bar{x} = 2.08$) “disagreed” more than “black” ($\bar{x} = 2.66$), “Hispanic/Latino” ($\bar{x} = 2.65$) and “other” ($\bar{x} = 2.63$). In summary, “other” races did not see the purpose of mandated assessments and the ability to predict the school or student’s performance. White students were more adamant than other races that mandated assessments were not “enjoyable” or “helpful”.

The fifth demographic characteristic was GPA. A large percentage of the students in the study (78%) identified as having a GPA of “3.5 or higher”. This study encompassed a sizable number of students who generally did well academically at school. Statistically, no inferences could be drawn about this variable because it was so homogenous when assessing GPA.

The last demographic investigated was based on students plans after high school. According to the SCoA-VI, students who were planning to attend “college or technical school” ($\bar{x} = 2.07$) were more likely to show a greater disagreement toward “Affective Benefits”. These students disagreed that mandated assessments were “enjoying” and “helpful” than peers who sought a “job/gap year” ($\bar{x} = 2.50$) or were “not sure” ($\bar{x} = 2.48$) on plans after high school. Further analysis showed students who were committed to continuing their education after high school perceived to have a more negative outlook towards the enjoyment and helpfulness of mandated assessments.

In conclusion, demographic characteristics examined in the study (gender, language, race, parent/guardian education and plans after high school) had an impact on high school student perceptions towards mandated assessments. Three demographic factors (gender, race, and

parent/guardian education) affected two or more of the conceptions assessed in the study. The demographic characteristics (gender, race, parent/guardian education and plans after high school) had the greatest impact on the conception, “Affective Benefits”. Demographics will be corroborated by the literature in the next section.

Conclusions

Students may have a dislike for testing, but ultimately they want to improve. Despite their negative feelings towards the mandated testing process, students offered valuable insight on possible adjustments for creating a more conducive measure that is relevant, reliable and offers opportunities for improvement. This study draws several conclusion that include no benefits for students, the value of tests, lack of student motivation, minimal test feedback, student morale and demographic data.

First, this study aligns with the literature regarding unintended outcomes at the classroom level (Layton, 2015; Simon, 2010). According to the SCoA-VI, 9th and 10th students perceived mandated assessments were “unfair/bad” and often “ignored” by students. They expressed concern about the reliability of the results as many students guessed or were quickly “clicking” to finish the test. In addition to the survey data, a large percentage of students responded that there was “no benefits” for students. Also, concerns for the length and time these assessments consumed were described in the drawbacks as major contributors to the shortcomings of testing mandates.

Similar to these findings, Simon (2010) described the testing process as time-consuming due to preparing and administering the assessment. For example, Layton (2015) highlighted a study consisting of 66 city school districts that found on average PreK-12th grade students spent

18.4 hours a year on testing. Likewise, in this study, students reported, “a day taken away from learning”, “each test takes about 2 hours of class time which cuts into other classes”, and “it’s a waste of time that I could have spent actually learning.” Students’ responses substantiated the concerns educators report about the amount of time mandated assessments takes from instruction and other content areas (Mora, 2011; Au, 2009; Berliner, 2009). It is imperative that stakeholders at the Federal/state and district level assess the amount of time devoted to test-taking. Similar to what the students were advocating in this study, new approaches towards testing mandates that involve all stakeholders, including students, is greatly needed so that a more useful system might be created.

The second conclusion relates to the concern students had about mandated assessments and the value of the content being assessed. On the SCoA-VI, 9th and 10th grade students showed a strong “disagreement” for how these mandated exams improved learning. Students perceived testing mandates were not “helpful”, had “little impact on learning” and not seen as “an accurate measure” of their learning. In addition to the survey, students also responded negatively in the open-ended questions when describing test content. Students stated that it was “non-related to curriculum” and consisted of “irrelevant topics” that they were “not currently studying” and was just “about memorization”. Similar to these students’ perceptions, Berliner (2009) also concluded that mandated assessments tended to examine lower order thinking skills and further supported the need for more critical-thinking that was relevant to the real world.

Furthermore, the value of testing is important due to the money being spent on these initiatives (Topol et al., 2013). Given the current financial situation across the United States, new assessment designs need to be as cost-effective and efficient as possible, because funding and resources are very limited. As policymakers and administration make decisions about how to

design, conduct, and pay for assessments, they need to weigh and balance accurate information from all stakeholders, especially students. Topol, et al. (2013) stated that school districts spend on average between \$15 to \$20 per student on assessments and data management systems that are not always well-aligned and are unable to measure more advanced skills. Despite this, many experts agree that assessments are needed to measure not only what students have learned but can be used as learning tool for students, parents, and teachers (Topol et al. 2013). Given that testing mandates are costly, finding cost effective solutions that meet pedagogical and assessment needs is warranted. To be beneficial, federal/state and district systems need economically sound testing that accurately and reliably measures learning outcomes.

The third conclusion in this study is that students did not perceive the test's results provided motivation or helpful feedback. According to the SCoA-VI, students were more likely to "strongly disagree" with statements such as mandated assessments, "motivate me and my classmates to help each other" or "our class becomes more motivated to learn". As for feedback, students also "disagreed" on the survey that mandated assessments "identify what I need to study next", "tells my parents how well I am doing" and the "teacher uses assessments results to help me improve". In the open-ended questions students expressed, "we don't learn anything", "they don't help us learn or improve", "they don't give us the progress that we need", and "I don't think they show what we have learned". These findings were similar to a previous study that indicated teachers' concerns for how testing mandates did not support improvement for them and their students (Harris, Harnett, & Brown; 2009). However, this conclusion contradicts Nicholas (2007) rationale for teachers and students for having testing mandates because he argued that they are a motivation tool that provides higher levels of accountability and the feedback assists

with learning and teaching. This study, however, does not support Nicholas' rationale and instead suggest that the mandated tests provide little motivation for students.

Another conclusion drawn from the SCoA-VI indicated students disagreed mandated assessments measured "school quality". In other words, students felt that their mandated assessment scores did not represent the true value or worth of the school. This was further elaborated on by students in the open-ended questions. Although a majority of students found no benefit at all for having taken mandated assessments, a small percentage acknowledged non-student benefits. For example, students stated, "which schools need better funding and staff", "the overall improvement of the school as a whole", and "understanding achievement and wealth gaps". The fourth conclusion about non-student benefits was consistent with previous literature (Reback et al., 2011; Supovitz, 2009) describing testing mandates at the organizational level as a needed process to ensure all students have a fair and equal opportunity. Au & Hollar (2016) described that for many people testing can make sense because it provides clear data that can be used to hold people and schools accountable. Students described this notion that a non-student benefit of mandated assessments could be that it creates more equity among schools, districts, and states. However, since only a few of the students saw benefits for the school, it must be emphasized that a majority of the students do not see a need for mandated assessments to create equity or have an effect on the schools.

This study also confirmed prior research about mandated tests and their impact on students. The study showed that students perceived that mandated assessments were not an accurate measure of their learning. Students reported that they don't "ignore assessment information" but agree on the SCoA-VI that mandated assessments are "unfair to students", "not very accurate" and "value-less". In the open-ended questions, students described mandated

assessments did not “show how intelligent I am”, “have little impact” and “interfere with my learning”. These responses portrayed a “stressful” and “sometimes hostile learning environment” which resulted in students feeling “bad” or “stupid”. Previous research has shown that testing can create intense pressure and anxiety, which can produce negative results that potentially have harmful effects to the learner (Duncan & Stevens, 2011). In this study, students reported psychological symptoms and physical manifestations from having participated in mandated assessments. For example, students stated, “I get stressed out, I fear my grade, I can’t stay focused and feel drained after testing.”, “I don’t get enough sleep because I’m either nervous or studying”, and “Higher stress levels and being taught that a test score determines my whole future”. These findings substantiate the harmful effects that has been described in previous research (Duncan & Stevens, 2011).

Another conclusion connected to prior research are the effects mandated assessments have on student morale. In this study, the SCoA-VI, showed that 9th and 10th grade students expressed a lack of effort they had for their testing performance. In the qualitative section, students described the length of the exam caused them to want to “click random ones” or “guess”. Another explained that, “Having taken mandated test starting at a very young age you begin to lose interest and motivation for the same test year after year.” Previous literature advised that student motivation can often inflate or deflate a score, meaning the results may not always be an accurate measure of their knowledge and understanding of the content (Zerpa et al., 2011). Paris, Roth, and Turner (2000), discovered older students tended to put forth less effort and fewer test strategies compared to younger students being assessed. The results of this study, therefore, substantiates concerns about the impact on student morale with regards to the number of tests the students are required to take each year.

Finally, there were several conclusions related to the demographic data analyzed in this study. First, gender showed to have the most statistical significance regarding the four conceptions surveyed on the SCoA-VI. Three out of the four conceptions (External Attribution, Affective Benefits and Irrelevance) showed to have a relationship between the conceptions and the gender identified by students. Reardon et al. (2018) examined the relationship between test item formats and the gender achievement gap. They found that test item formats explained approximately 25 percent of the variation in gender achievement gaps in the United States. For example, multiple choice tests were shown to be favored more by males than females (Reardon et al. 2018). The mandated assessments referred to in this study were primarily in multiple choice formats which could explain why females tended to “disagree” more than males concerning the areas of “Improvement”, “External Attribution” and “Affective Benefits”. This was also reiterated in the open-ended responses by students when several recommendations were provided for other formats/designs to assess student learning/achievement.

Another conclusion was race/ethnicity and how it tied to previous literature concerning students of color. On the SCoA-VI, “white” students disagreed more strongly than “black”, “Hispanic/Latino” and “other” on the “Affective Benefits”. White students did not perceive testing assessments were helpful. Greenfield et al. (2003) noted that students of varying ethnic backgrounds can hold differing beliefs and these interpretations need to be taken into consideration when teaching and assessing students. Understanding how cultural influences shape students perceptions could have very easily affected how students perceived “Affective Benefits”. Therefore, white students could have very different meanings or values towards these exams when compared to students of color, based on their current situation/experiences.

The last conclusion drawn from the demographic data involved the level of parent/guardian education. In this study, students who identified as “high” parent/guardian education disagreed more on the “Improvement” and “Affective Benefits” from the SCoA-VI compared to “low” parent/guardian educational levels. Research has shown that an “opportunity gap” can lead to an achievement gap (Hardy, 2017; Carter & Welner, 2013). Socio-economic factors such as income levels, housing, educational attainment, employment rates, crime, and resources available to schools, are typically worse for African Americans and Hispanics than for Whites. These socio-economic factors can contribute to Black and Hispanic students having less access to opportunities than whites thus resulting in less educational advantages. In this study, students of color tended to disagree less than White Non-Hispanic students when asked about the “Improvement” and “Affective benefits” gained from completing testing mandates. Knowing that Black and Hispanic/Latino students are more likely to be exposed to an “opportunity gap” could influence how these students perceive testing mandates and the outcome these experiences may provide to them personally.

Implications for Practice

This study began with an interest in how students perceive mandated assessments. The findings showed students disliked testing mandates, but offered thoughts on how they might be more useful. Research has shown that students who believe they have a voice in school are seven times more likely to be academically motivated than students who do not believe they have a voice (Quaglia Institute for School Voice and Aspirations, 2016). Teachers, administrators, district level coordinators and policy makers should consider the students comments and feedback from this study. Why can't these exams be made more useful for educational purposes?

Are there ways these assessments might have more of an educational impact on students? If teachers are using test results for diagnostic purposes, should that be made clear to students? Also, what about the feedback provided to students after these exams?

For educators and administration, the main goal of our work is to help students learn. Several students in the study requested the opportunity to see what they got correct and what they got wrong. Listening to student voice can increase their level of effort and persistence, which is one of the most important factors that affect achievement (Toshalis & Nakkula, 2012). From this study, it is apparent students want to utilize these exams in a more beneficial way that supports their own learning. Considering the students' suggestions it seems to clearly be the next step needed to improve the testing process.

Implications for Future Research

This study is important for students, parents, and all stakeholders involved in the testing process because it offers an important perspective that has not been heavily emphasized within the current literature. The following areas are recommendations derived from the findings and conclusions.

The first recommendation for research would be the replication of this study using larger samples to further explore demographic characteristics. It would be important to make sure demographics such as gender, race/ethnicity, etc. are included to make the sample more representative of the population. This study only looked at students from one school district. The study's sample was primarily Non-Hispanic White (81%) which may have produced biased results. However, the demographic characteristics examined did highlight significant factors. Female and English speaking students were more likely to agree with the "Irrelevance" they feel

about mandated assessments. Also, females, whites, those planning to attend college or technical training, and those with higher levels of parent/guardian education disagreed with “Affective Benefits”, meaning they perceived the assessments were not enjoyable and helpful. Repeating the study with a larger representative sample size would be recommended, so that more information and understanding on how students demographic characteristics affect their perceptions towards mandated assessments.

The second recommendation for research would be to utilize the instrument with several tested grade levels at the primary and secondary levels rather than just high school students. This would provide varying viewpoints from students spanning grades 3rd-10th that are all required to participate in testing mandates at the state and district level. In this study, the effort and motivation high school students reported in regard to mandated assessments were concerning. One previous study did show a discrepancy between elementary and secondary students regarding the level of effort and motivation given during these exams (Paris, Roth, and Turner (2000)). Therefore, another recommendation would be to replicate this study using primary and secondary students or a longitudinal study to explore the relationship age/grade has on students perceptions towards mandated assessments.

The third recommendation would be to conduct a comparative study that examines the length of the tests. Many students described the time lost from the classroom as a major disruption from their current instruction. Students suggested recommendations such as shorter exams or having them taken in increments spread overtime to help alleviate these unwanted consequences. A study examining different length of days and time devoted to tests would add to the research on testing accuracy.

The fourth recommendation would be to investigate a specific mandated test rather than encompassing all of them at the federal/state and district level. Mandated assessments is a broad term that includes multiple tests at the state and district level. Narrowing down on a specific testing mandate and comparing it to other tests based on the test's purpose could provide additional insight on how students perceive this experience. It could also validate how students perceive the function or purpose for these exams which would be useful to those stakeholders at the district and building level initiating and assigning these tests.

The fifth recommendation would be to look at the differences in feedback provided by each test. Feedback referred to the information provided to students after completing the exam. Students reported concerns about "not showing a score" and wanting "reasons/explanations". Students also included recommendations with regards to "how to fix missed questions" or "more detailed information concerning their score". Taking a deeper look at mandated assessments by examining each test in regards to the level of feedback provided to students might distinguish some assessments. What information is given to students from the exam itself and what if any information is provided by teachers afterwards? It was evident that students want this process to be beneficial to their learning.

The sixth recommendation would be to look at privacy issues for students under 18. Hattie and Timperley (2007) described feedback as one of the most powerful influences on learning and achievement. In this study, students provided thoughtful comments and recommendations to improve the test to make it more beneficial to them and their learning. However, the sample size in this study was relatively small because of the difficulty to get permission to study students at the secondary level because they were under the age of 18. Having to go through multiple steps to attain parental consent, which was done through a third

party, limited the number of students participating in this study. I would recommend that when investigating student perceptions taking other avenues to attain parental consent to increase the sample size rather than the method used in this study going through a third party.

The last recommendation would be to study student perceptions after time had elapsed. Perhaps examining perceptions of students in college to determine, over time, if the tests had influence on their academic career or performance. As the educational environment strongly affects student achievement, satisfaction and success (Lo, 2010). It is important to gain feedback from students regarding their testing experience. Assessments are considered one of the most important aspects of higher education, because students' learning processes and learning outcomes are strongly influenced by the assessment system (Gibbs and Simpson 2004). Understanding how students perceive these testing mandates from individuals seeking higher levels of education after high school, could provide additional viewpoints that have not been explored in prior studies.

Appendix

A: Permission to Utilize SCoA-VI

Thursday, November 29, 2018 at 10:09:41 AM Central Standard Time

Subject: Permission to Use Survey: Kansas University (USA)
Date: Monday, October 10, 2016 at 12:24:09 PM Central Daylight Time
From: Jennifer Woolever
To: gt.brown@auckland.ac.nz

October 10, 2016

Dear Dr. Gavin Brown,

I am a doctoral student from the University of Kansas writing my dissertation tentatively titled Student Perceptions of State Mandated Assessments under the direction of my dissertation committee chaired by Dr. Rick Ginsberg, Dean of Education.

I would like your permission to utilize the Student Conceptions of Assessment Abridged Survey (SCoA) in my research study. I would like to utilize and print your survey under the following conditions:

- I will use this survey only for my research study and will not sell or use it with any compensated or curriculum development activities.
- I will include the copyright statement on all copies of the instrument.
- I will send my research study and one copy of reports, articles, and the like that make use of these survey data promptly to your attention.

If these are acceptable terms and conditions, please indicate so by signing one copy of this letter and returning it to me either through postal mail or e-mail:

Jennifer Woolever

Thursday, November 29, 2018 at 10:03:44 AM Central Standard Time

Subject: RE: Permission to Use Survey: Kansas University (USA)
Date: Monday, October 10, 2016 at 1:53:53 PM Central Daylight Time
From: Gavin Brown
To: Jennifer Woolever
Attachments: SCoA-VI 2008.doc, scoa-vi (2).xls

Hi Jennifer

Thank you for asking for permission

By all means you may use the SCoA as you have described. I am attaching the version I use so you can see the layout formatting and item order I use. The spreadsheet identifies the factor structure as identified in NZ studies.

I look forward to receiving your completed reports digitally when they are done.

Please use this citation for the SCoA

Brown, G. T. L. (2008). Students' Conceptions of Assessment Inventory (SCoA Version VI) [Measurement instrument]. Auckland, NZ: University of Auckland.

Please accept my signature space here as equivalent to signing.
sincerely

Prof. Gavin T L Brown, PhD

Director Quantitative Data Analysis and Research Unit

Faculty of Education & Social Work

The University of Auckland

Tel: +64 9 3737599 ext. 48602

Mob: +64 22 108 7253

Honorary Professor, Dept. of Curriculum & Instruction, Education University of Hong Kong

Affiliated Professor, Dept. of Applied Educational Sciences, University of Umea, Sweden

New: Handbook of Human and Social Conditions in Assessment.

<https://www.routledge.com/products/9781138811553>

B: IRB Approval**APPROVAL OF PROTOCOL**

March 7, 2018

Jennifer Woolever
wooleverj@ku.edu

Dear Jennifer Woolever:

On 3/7/2018, the IRB reviewed the following submission:

Type of Review:	Initial Study
Title of Study:	Student Perceptions Towards Mandated Assessments
Investigator:	Jennifer Woolever
IRB ID:	STUDY00142053
Funding:	None
Grant ID:	None
Documents Reviewed:	• 6.2 Update/Requested Items, • Assent Procedure & Student Survey, • HRPP Parent Guardian Consent Authorization Email & Form, • Human Research Protocol e-compliance, • RE_Permission to Use Survey_ Kansas University (USA).eml

The IRB approved the study from 3/7/2018 to 3/6/2019.

1. Before 3/6/2019 submit a Continuing Review request and required attachments to request continuing approval or closure.
2. Any significant change to the protocol requires a modification approval prior to altering the project.
3. Notify HRPP about any new investigators not named in original application. Note that new investigators must take the online tutorial at https://rgs.drupal.ku.edu/human_subjects_compliance_training.
4. Any injury to a subject because of the research procedure must be reported immediately.
5. When signed consent documents are required, the primary investigator must retain the signed consent documents for at least three years past completion of the research activity.

If continuing review approval is not granted before the expiration date of 3/6/2019 approval of this protocol expires on that date.

Please note university data security and handling requirements for your project:
<https://documents.ku.edu/policies/IT/DataClassificationandHandlingProceduresGuide.htm>

You must use the final, watermarked version of the consent form, available under the "Documents" tab in eCompliance.

Sincerely,

Jocelyn Isley, MS, CIP
IRB Administrator, KU Lawrence Campus

C: Parent/Guardian Email

Dear Parent/Guardian,

The Shawnee Mission School District Assessment & Research group is partnering with a doctoral student at the University of Kansas to obtain student perspectives on mandated assessments. We are seeking feedback from 9th and 10th grade students, and we need your permission to proceed.

If you are ok with your 9th or 10th grader completing this survey, please forward this message to your 9th or 10th grader's email address or have them access the link below from your email. Student responses are voluntary and anonymous. The survey consists of 42 questions. The survey should take no longer than 15 minutes to complete.

The researcher's name is Jennifer Woolever. She is a doctoral student at the University of Kansas, and she is also Principal at Roesland Elementary School in the Shawnee Mission School District. She is writing her dissertation for doctorate in Educational Leadership & Policy Studies at the University of Kansas. The district assessment office is also interested her analysis to perhaps better understand how we can make assessment results more relevant and meaningful to students.

The survey results will be used for Mrs. Woolever's dissertation. Individual results of the study will remain absolutely confidential and anonymous. Should this study be published, only the aggregate results will appear.

Your approval to conduct the study is greatly appreciated. I have attached additional information to the email concerning your consent/authorization in the study. Please know at any time you may contact Mrs. Woolever by email (jenniferwoolever@smsd.org) and/or her faculty supervisor (ginsberg@ku.edu). If you agree, please kindly forward the survey link below to your student or have them access the link below from your email.

If you do not want your child to participate, please disregard this email.

Student Survey Link: <https://www.surveymonkey.com/r/woolever>

Thank you for your support!

Jennifer Woolever
Doctoral Student, University of Kansas
Principal, Roesland Elementary, Shawnee Mission School District

Daniel Gruman, Ed.D.
Assessment & Research
Shawnee Mission USD 512
8200 W 71st St | Overland Park, KS 66204
913-993-8658

Attachment: Parent/Guardian Study Informati

D: Demographic Information

Student Perceptions Towards Mandated Assessments

Demographic Information

The information collected will be used for analysis purposes only. None of the information collected will be individually identifiable in the final study.

*** What is your gender?**

- Female
 Male

*** What race /ethnicity do you identify with the most?**

- Non-Hispanic White
 Black or African American
 Hispanic or Latino
 Other

*** What is the primary language spoken in your home?**

- English
 Spanish
 Other

*** What is the highest level of education that one of your parents or guardians have accomplished?**

- Did not Graduate High School
 Graduated High School
 Associates Degree (2 Years of College)
 Bachelor Degree (4 Years of College)

*** Which range best represents your middle and high school grade point average (GPA)?**

- 3.5 and Above (Mostly A's)
 2.5-3.5 (Mostly B's)
 1.5-2.5 (Mostly C's)
 Other

*** What do you plan to do after high school?**

- 4-year College
 2-year College or Technical Training
 Full time job
 Take a year off from school, what is referred to as a "Gap year" then attend college
 Not sure

2 / 7

Prev

Next

Powered by



See how easy it is to [create a survey](#).

*** Mandated tests help teachers track my progress**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input type="radio"/>							

*** Mandated tests are engaging and enjoyable experience for me**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input type="radio"/>							

*** I ignore mandated tests information**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input type="radio"/>							

*** Mandated tests are a way to determine how much I have learned from teaching**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input type="radio"/>							

*** A Mandated test is checking off my progress against achievement objectives or standards**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input type="radio"/>							

*** I make use of the feedback I get to improve my learning**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input checked="" type="radio"/>	<input type="radio"/>						

*** Mandated test provide information on how well schools are doing**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input checked="" type="radio"/>	<input type="radio"/>						

4 / 7



Prev

Next

*** Teachers use my mandated test results to see what they need to teach me next**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input type="radio"/>							

*** When we are assessed, our class becomes more motivated to learn**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input type="radio"/>							

*** I ignore or throw away my mandated test results**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input type="radio"/>							

*** Mandated tests show whether I can analyze and think critically about a topic**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input type="radio"/>							

*** I find myself really enjoying learning when I am assessed**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input type="radio"/>							

*** Mandated tests have little impact on my learning**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input type="radio"/>							

*** Mandated tests tell my parents/guardians how much I have learned**

Strongly Disagree							Strongly Agree
1	2	3	4	5	6	7	
<input type="radio"/>							

6 / 7



Prev

Next

F: Open-Ended Questions

Student Perceptions Towards Mandated Assessments

Overall, what do you see as the benefits to you as a high school student for having taken mandated tests in school?

What do you see as the drawbacks for you as a high school student having taken mandated tests in school?

Do you have any other opinions about mandated test you would like to share?

7 / 7

Prev

Done

Bibliography

- Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometrics issues. *Educational Assessment*, 8(3), 231-257.
- Abbott, S. (2014). Hidden Curriculum. The Glossary of Education Reform. 13 July 2015, <https://www.edglossary.org/>.
- Abrams, L. M., Pedulla, J. J., & Madaus, G. F. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory into Practice*, 42(1), 18-29.
- Adamson, F., Åstrand, B., & Darling-Hammond, L. (Eds.) (2016). *Global education reform: How privatization and public investment influence education outcomes*. New York: Routledge.
- Alexander, K. L., Entwisle, D. R., & Bedinger, S. D. (1994). When expectations work: Race and socioeconomic differences in school performance. *Social Psychology Quarterly*, 57, 283-299.
- Andrade, H. & Du, Y. (2005). Students Perspectives on Rubric Referenced Assessments . *Practical Assessment, Research and Evaluation*, 10(3), 1-11.
- Amrein-Beardsley, A. (2008). Methodological concerns about the Education Value-Added Assessment System (EV AAS). *Educational Researcher* 37(2), 65-75.
- Au, W., & Hollar, J. (2016). Opting out of the education reform industry. *Monthly Review: An Independent Socialist Magazine*, 67(10), 29-37.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative meta-synthesis. *Educational Researcher*, 36(5), 258-267.
- Au, W. (2009). Social studies, social justice: W(h)ither the social studies in high-stakes testing? *Teacher Education Quarterly*, 36(1), 23-58.
- Aviv, R. (2014). A Middle-School Cheating Scandal Raises Questions About No Child Left Behind. The New Yorker. Available at http://www.newyorker.com/reporting/2014/07/21/140721fa_fact_aviv?currentPage=all.
- Baker, M., & Milligan, K. (2013). Boy-Girl Differences in Parental Time Investments: Evidence from Three Countries. *Journal of Human Capital, University of Chicago Press*, 10(4), 399-441.

Black, P., & Wiliam, D. (2004). The formative purpose: Assessment must first promote learning. *The Yearbook of the National Society for the Study of Education*, 103(2), 20-50.

Berends, M., & Penaloza, R.V. (2010). Increasing Racial Isolation and Test Score Gaps in Mathematics: A 30-Year Perspective. *Teachers College Record*, 112(4), 978-1007.

Berliner, D. C. & Nichols, S. L. (2007). *Collateral Damage: How High-Stakes Testing Corrupts American Schools*. Cambridge: Harvard Education Press.

Berliner, D. C. (2009). MCLB (Much curriculum left behind): A U. S. calamity in the making. *Educational Forum*, 73(4), 284-296.

Borman, G. D., & D'Agostino, J. V. (2001). Title I and Student Achievement: A quantitative synthesis. Mahwah, NJ: *Lawrence Erlbaum Associates, Inc.*, 2, 65–75.

Brenchley, C.. *What is ESEA?* U.S. Department of Education. 8 April 2015, <https://blog.ed.gov/2015/04/what-is-esea/>.

Brown, A. B., & Clift, J. W. (2010).The unequal effect of adequate yearly progress. *American Educational Research Journal*, 47(4), 774–798.

Brown, C. (2010). Children of Reform; The Impact of High-Stakes Education Reform on Preservice Teachers. *Journal of Teacher Education*, 61, 477-291.

Brown, G. T. L., Lake, R. & Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education*, 27, 210-220.

Brown, G. T. L. (2008). Students' Conceptions of Assessment Inventory (SCoA Version VI) [Measurement instrument]. Auckland, NZ: University of Auckland.

Brown, G. T. L (2006). Teachers' conceptions of assessment: *Validation of an abridged version*. *Psychological Reports*, 99, 166-170.

Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy. *Assessment in Education: Principles, Policy & Practice*, 11(3), 301-318.

Brown, G. T. L., & Hirschfeld, G. H. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice*, 15(1), 3–17.

Brown, G. T. L., & Hirschfield, G. H. (2007). Students' conceptions of assessment and mathematics: Self-Regulation raises achievement. *Australian Journal of Educational and Developmental Psychology*, 7(1), 63–74.

Brown, G. T. L., Pishghadam, R., & Shayesteh Sadafian, S. (2014). Iranian university students' conceptions of assessment. *Assessment Matters*, 6(1), 5–33.

Brown, J. M. & Gaxiola, C. (2010). Why would they try? Motivation and motivating in low- stakes information skills testing. *Journal of Information Literacy*, 4(2), 23- 36.

Campbell, D. T., (1976). *Assessing the impact of planned social change*. Hanover, NH: Dartmouth College, Public Affairs Center.

Carnine, D. W., Silbert, J., Kame'enui, E. J., & Tarver, S. G. (2010). *Direct Instruction Reading* (5th ed.). Boston, MA: Merrill.

Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross- state analysis. *Educational Evaluation and Policy Analysis*, 24, 305-331.

Carter, P. & Welner, K. (2013). *Opportunity Gap: What America Must Do to Give Every Child and Even Chance*. Oxford University Press. Oxford, New York.

Cech, E. (2013). The Self-Expressive Edge of Occupational Sex Segregation 1. *American Journal of Sociology*, 119(3), 747-789.

Charles, M., & Bradley, K. (2009). Indulging Our Gendered Selves? Sex Segregation by Field of Study in 44 Countries. *American Journal of Sociology*, 114(4), 924-976.

Codero-Guzman, H. R. (2001). Cognitive skills, test scores and social stratification: The role of family and school-level resources on racial/ethnic difference in scores on standardized tests (AFQT). *The Review of Black Political Economy*, 28(4), 31-71.

Cowie, B. (2005). Pupil Commentary of Assessment for Learning. The Glossary of Education Reform. *Curriculum Journal*, 16(2), 137– 151.

Cullen, J.B. & Reback, R. (2006). Tinkering toward accolades: School gaming under a performance accountability system. *Advances in Applied Microeconomics*, 4, 1-34.

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

Darling-Hammond, L. & Rustique-Forrester, E. (2005). The Consequences of Student Testing for Teaching and Teacher Quality. *Yearbook of the National Society for the Study of Education*, 104(2), 289-319.

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession (Technical report)*. Washington, DC: National Staff Development Council.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and home environment. *Journal of Family Psychology, 19*(2), 294-304.

Dawson, H. (2013). Teachers' motivation and belief in a high-stakes testing context. *Humanities and Social Sciences, 74*(1), 115-137.

Dee, T. S., & Jacob, B. A. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management, 30*, 418-446.

Deming, D. J., Cohodes, S. Jennings, J., & Jencks, C. (2013). *School Accountability, Postsecondary Attainment and Earnings*. National Bureau of Economic Research. Cambridge, MA.

Destin, M., & Oyserman, D., (2009). From Assets to School Outcomes: How Finances Share Children's Perceived Possibilities and Intentions. *Psychological Science, 20*, 414-418.

Dubow, E. F., Boxer, P., & Huesmann, L. R. (2009). Long-term effects of parents' education on children's educational and occupational success: Mediation by family interactions, child aggression, and teenage aspirations. *Merrill-Palmer Quarterly, 55*(3), 224-249.

Duncan, B., & Stevens, A. (2011). High-stakes standardized testing: Help or hindrance to public education. *National Social Science Journal, 36*(2), 34-43.

Dutro, E., & Selland, M. (2012). I like to read, but I know I'm not good at it: Children's perspectives on high-stakes testing in a high poverty school. *Curriculum Inquiry, 42*(3), 340-

Dworkin, A. G., & Tobe, P. F. (2014). The effects of standards based school accountability on teacher burnout and trust relationships: A longitudinal analysis. *Trust & School Life, 121- 143*.

Edwards, N. T. (2006). The historical and social foundations of standardized testing: In search of a balance between learning and evaluation. *Shiken: JALT Testing & Evaluation SIG Newsletter, 10*(1), 8-16.

Eklof, H. (2007). Test-taking motivation and mathematics performance in TIMSS 2003. *International Journal of Testing, 7*, 311-326.

Eklof, H. (2010). Skill and will: Test-taking motivation and assessment quality. *Assessment in Education: Principles, Policy, & Practice, 17*, 345-356.

Elwood, J., and L. Lundy (2010). "Re-visioning Assessment through a Children's Rights Approach: Implications for Policy, Process and Practice." *Research Papers in Education, 25*(3), 335-353.

Eslinger, J. C. (2014). Navigating between a rock and a hard place: Lessons from an urban school teacher. *Education and Urban Society*, 46(2), 209-223.

Ewijk, R. V., & Slegers, P. (2010). The Effect of Peer Socioeconomic Status on Student Achievement: Meta-Analysis. *Educational Research Review*, 5(2), 1-20.

Ferguson, M., Kober, N., & Stark-Rentner, D. (2017). *What do Teachers and District Leaders Feel about State Standards and Assessments?* Center on Education Policy (CEP). The George Washington University: Graduate School of Education and Human Development.

Figlio, D. & Getzler, L. (2006). Accountability, ability, and disability: Gaming the system? Amsterdam: Elsevier. In *Advances in Microeconomics*, 14, 35–49.

Fletcher, A. (2005). Meaningful Student Involvement Guide to Students as Partners in School Change. *Soundout. The Free Child Project*. Olympia, WA.

Fletcher, D. "History: Standardized Testing". *Time*. 11 December 2009, <http://www.time.com/time/nation/article/0,8599,1947019,00.html>.

Francis, D. J., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). *Practical guidelines for the education of English language learners: Research-based recommendations for instruction and academic interventions*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Fryer R.G. & Levitt, Jr., S.D. (2009). An Empirical Analysis of the Gender Gap in Mathematics. *American Economic Journal: Applied Economics*, *American Economic Association*, 2(2), 210-240.

Gibbs, G., & Simpson, C. (2004). Does Your Assessment Support Your Students' Learning. *Journal of Teaching and Learning in Higher Education* 1(1): 1–30.

Guthrie, J. W. & Springer, M. G. (2004). A nation at risk revisited: Did "wrong" reasoning result in "right" results? At what costs? *Peabody Journal of Education*, 79(1) 7-35.

Hagopian, J. (2014). *More than a score: The new uprising against high-stakes testing*. Chicago: Haymarket Books.

Haladyna, T. M., & Downing, S. M. (2004). Construct irrelevant variance in high stakes testing. *Educational Measurement: Issues and Practice*, 23(1), 17–27.

Hallett, T. (2010). The myth incarnates: Recoupling processes, turmoil, and inhabited institutions in an urban elementary school. *American Sociological Review*, 75(1), 52–74.

Hamilton, L. M., Stecher, B. M., & Klein, S. P. (2002). *Making sense of test-based accountability in education*. Santa Monica, CA: RAND.

Hanushek EA, Kain JF, Rivkin SG. 2009. New evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement. *J. Labor Econ.* 27, 349-383.

Hanushek, E.A. & Raymond, M.E., (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, 24(2), 297–327.

Hardy, A. (2017). The Consequence of School Performance Measures: Inequality of Access and Opportunity. *UCL IOE Press*, 34(2), 39-43.

Harris, L., Harnnett, J., & Brown G., (2009). “Drawing” Out Student Conceptions of Assessment: Using Pupils’ Pictures to Examine their Conceptions of Assessment. *Information Age Publishing Inc.*, 77, 53-83.

Hartman, C. W. (2001). *Challenges to equality: Poverty and race in America*. Armonk, NY: M.E. Sharpe.

Hattie, J. & Timperley, H. (2007), The power of feedback, *Review of Educational Research*, 81-112.

Hao, L., & Yeung, W.J. (2015). Parental Spending on School-Age Children: Structural Stratification and Parental Expectation. *Demography*, 52(3), 835–860.

Heilig, J. V. (2011). Understanding the interaction between high-stakes graduation tests and English learners. *Teachers College Record*, 12, 2633-2669.

Heritage, M. (2013) *Formative Assessment in Practice: A Process of Inquiry and Action*. Harvard Education Press. Cambridge, MA.

Hirsch, Jr., E.D., & Moats, L.C. (2001). *Overcoming the Language Gap*. *Education Week*, 25(2), 4-9.

Hirschfeld, G. H. F., & von Brachel, R. E. (2008). Students’ conceptions of assessment predict strategy use in higher education. Paper presented at the ITC 2008 conference.

Hoffman, J.V., Assaf, L.C., & Paris, S.G. (2001). High- stakes testing in reading: Today in Texas, tomorrow? *The Reading Teacher*, 54(5), 482-494.

Holmes, K. & Papageorgiou, G. (2009). Good, Bad & Insufficient: Students’ Expectations, Perceptions and Uses of Feedback. *Journal of Hospitality, Leisure, Sport & Tourism Education*. 8(1), 85-96.

Horn, C. (2003). High-stakes testing and students: Stopping or perpetuating a cycle of failure? *Theory into Practice*, 42(1), 30-41.

Husband, T. (2014). Urban early childhood educators' perceptions of the impact of NCLB on their literacy practices. *Journal of Curriculum and Teaching*, 3(1), 47-53.

Hurst, D., Alexandra, T., Anne, M., Meek, A., & Sellers, J. (2003). Overview and Inventory of State Education Reforms: 1999 to 2000. National Center for Education Statistics. 5, January 2012, <http://nces.ed.gov/pubs2003/2003020.pdf>.

Hurst, D., Tan, A., Meek, A., & Sellers, J., (2003). *Overview and Inventory of State Education Reforms: 1990 to 2000*. U.S. Department of Education, National Center for Educational Statistics. Jessup, MD.

Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics*, 89, 761-796.

Jacob, B. A., & Levitt, S. D. (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics*, 118, 843-877.

Jacobs, J., Lanza, S., Osgood, D., Eccles, J., & Wigfield, A. (2002). Changes in Children's Self-Competence and Values: Gender and Domain Differences Across Grades one through twelve. *Child Development*, 73, 509-527.

Jennings, J. L., & Beveridge, A.A. (2009). How Does Test Exemption Affect Schools' and Students' Academic Performance? *Educational Evaluation and Policy Analysis*, 31(2), 153-175.

Johnson, L.B.. "Lyndon B. Johnson Inaugural Address 1965". UVA Miller Center Foundation. 20, January 1965, <https://millercenter.org/the-presidency/presidential-speeches/january-20-1965-inaugural-address>.

Jones, B.D., & Egley, R.J. (2006). Looking through different lenses: Teachers' and administrators' views of accountability. *The Phi Delta Kappan*, 87(10), 767-771.

Kansas State Department of Education. Kansas College and Career Ready Standards. 18, September 2016, <http://community.ksde.org/Default.aspx?tabid=4754>.

Kelly, S. (2009). The Black-White Gap in Mathematics Course Taking. *Sociology of Education*, 82(1). 47-69.

Koretz, D. (2008). *Measuring Up: What Educational Testing Really Tells Us*. Cambridge, MA: Harvard University Press.

Koretz, D.M., Hamilton, L.S., (2003). *Teachers' Responses to High-Stakes Testing and the Validity of Gains: A Pilot Study*. Center for Research on Evaluation, Standards, and Student Testing, CSE Technical Report 610.

Koretz, D.M., Linn, R.L., Dunbar, S.B., & Shepard, L.A. (1991). Effects of high-stakes testing on achievement: Preliminary findings about generalization across tests. Paper presented at the annual meeting of the American Education Research Association and the National Council of Measurement in Education, Chicago.

Layton, L.. “Majority of US Public School Students are in Poverty”. *The Washington Post*. Washington, D.C.. 31, March 2015,
<https://www3.nd.edu/~wschmitt/wapopovertyeducation.pdf>.

Lee, J., Moon, S., & Hegar, R. (2011). Mathematics skills in early childhood: Exploring gender and ethnic patterns. *Child Indicators Research* 4(3), 353–368.

Liu, L., Bridgeman, B., & Adler, R. M. (2012). Measuring learning outcomes in higher education: Motivation matters. *Educational Researcher*, 41, 352-362.

Lloyd-Smith, M., & Tarr, J. (2000). Researching children’s perspectives: A sociological dimension. *Buckingham Open University Press*, 39(3), 59-70.

Lo, C.C. (2010). How Student Satisfaction Factors Affect Perceived Learning. *Journal of the Scholarship of Teaching and Learning*, 10(1), 47 – 54.

Matos, D. A. S., Cirino, S. D., & Brown, G. T. L. (2009). *Students’ Conceptions of assessment in higher education in Brazil*. In D. M. McInerney, G. T. L. Brown, & G. A. D. Liem (Eds.), *Student Perspectives of assessment: What students can tell us about assessment for learning* (235–253). North Carolina: Information Age

McNeil, L. M. (2000). *Contradictions of School Reform Educational Cost of Standardized Testing*. Routledge. New York, New York.

Mitra, D. L. (2008). *Student voice in school reform: building youth-adult partnerships that strengthen schools and empower youth*. Albany, NY: State University of New York Press.

Mora, R. (2011). “School is so boring: High-stakes testing and boredom at an urban middle school”. *Penn GSE Perspectives on Urban Education*, 9(1), 1-9.

Neal, D. & Whitmore Schanzenbach, D. (2010). Left behind by design: Proficiency counts and test-based accountability. *Review of Economics and Statistics*. 92(2): 263-283.

Nelson, L. P., McMahan, S. K., & Torres, T. (2012). The impact of a junior high school community intervention project: Moving beyond the testing juggernaut into a community of creative learners. *School Community Journal*, 22(1), 125-144.

Nichols, S. L., & Berliner, D. C. (2007). *Collateral damage: How high-stakes testing corrupts America’s schools*. Cambridge, MA: Harvard Education Press.

Nichols, S. (2007). High-stakes testing: Does it increase achievement? *Journal of Applied School Psychology* 23(2), 47-64.

No Child Left Behind Act, (2002). U.S. Department of Education, Public Law 107-110. <https://www2.ed.gov/nclb/landing.jhtml>.

Olszewski-Kubilius, P., & Lee, S.-Y. (2011). Gender and other group differences in performance on off-level tests: Changes in the 21st century. *Gifted Child Quarterly*, 55(1), 54-73.

Oyserman, D. , Johnson, E., & James, L. (2011). Seeing the destination but not the path: Effects of socioeconomic disadvantage on school-focused possible self-content and linked behavioral strategies. *Self and Identity*, 10, 474-492.

Paris, S.G. Lawton, T. A., Turner, J.C. & Roth, J. L. (2000). Developing Disillusionment: Students' Perceptions of Academic Achievement Test. *Issues in Education*, 17-46.

Pedulla, J., Abrams, L., Madaus, G., Russell, M., Ramos, M., & Miao, J. (2003). *Perceived effects of state- mandated testing programs on teaching and learning: Findings from a national survey of teachers*. Chestnut Hill, MA: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.

Poulos, A., & Mahoney, M. J. (2007). Effectiveness of feedback: the students' perspective. *Assessment and Evaluation in Higher Education*, 33, 143-154.

Price, M., O'Donovan, B., & Rust, C. (2007). Putting a social-constructivist assessment process model into practice: Building the feedback loop into the assessment process through peer review. *Innovations in Teaching and Education International*, 44, 143-152.

Quaglia Institute for School Voice and Aspirations (2016). "School Voice Report 2016". Corwin Press. http://quagliainstitute.org/dmsView/School_Voice_Report_2016.

Raley, S., & Bianchi, S. (2006). Sons, Daughters, and Family Processes: Does Gender of Children Matter? *Annual Review of Sociology*, 32, 401-421.

Ravitch, D. (2010). *The death and life of the great American school system: How testing and choice are undermining education*. New York: Basic Books.

Reardon, S.F. (2008). *Thirteen Ways of Looking at the Black-White Test Score Gap*. Institute for Research on Education Policy & Practice. Stanford University.

Reardon, S.F., Kalogrides, D., Fahle, E.M., Podolsky, A., & Zárate, R.C. (2018). The Relationship Between Test Item Format and Gender Achievement Gaps on Math and ELA Tests in 4th and 8th Grade. *Education Researcher*. Retrieved from Stanford Center for Education

Policy Analysis: <https://cepa.stanford.edu/content/relationship-between-test-item-format-and-gender-achievement-gaps-math-and-ela-tests-4th-and-8th-grade>

Reback, R., Rockoff, J., & Schwartz, H. (2011). *Under Pressure: Job Security, Resource Allocation and Productivity in Schools Under NCLB*. National Bureau of Economic Research. Cambridge, MA.

Reeves, D. (2009). *Leading Change in Your School: How to Conquer Myths, Build Commitment, and Get Results*. Alexandria, VA: ASCD.

Rouse, C.E., Hannaway, J., Goldhaber, D. & Figlio, D. (2007). *Feeling the Florida Heat? How Low-Performing Schools Respond to Voucher and Accountability Pressure*. National Bureau of Economic Research. Cambridge, MA.

Rowe, A. (2011). The personal dimension in teaching: why students value feedback. *International Journal of Educational Management*, 25(4), 343-360.

Rudduck, J. & McIntyre, D. (2007) *Improving Learning Through Consulting Pupils (Improving Learning TLRP Series)*. London: Routledge.

Rumberger, R. W., & Palardy, G. J. (2005). Test Scores, Dropout Rates, Transfer Rates as Alternative Indicators of High School Performance. *American Educational Research Journal*, 42(1), 3-42.

Sauder, M. and Espeland, W. (2008). The Discipline of Rankings: Tight Coupling and Organizational Change. *American Sociological Review*, 74, 63-82.

Segers, M., & Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment? *Studies in Educational Evaluation*, 37(1), 49–54.

Shepard, L. A. (2002). The hazards of high stakes testing. *Issues of Science and Technology Online*. Retrieved at: www.issues.org/issues/19.2/shepard.htm.

Simon, M. (2010). Assessment versus achievement: Winner takes all! *Florida Journal of Educational Administration & Policy*, 3(2), 73-85.

Smith, M. L., Edelsky, C., Draper, K., Rottenberg, C., & Cherland, M. (1991). *The role of testing in elementary schools (CSE Technical Report 321)*. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Snyder, T., Brey, C., & Dillow, S. (2016). “*Digest of Education Statistics 2014*”. U.S. Department of Education, Washington, DC. *Institute of Education Sciences*. <http://nces.ed.gov/programs/digest>.

Spillane, J. P., Parise, L. M., & Sherer, J. Z. (2011) Organizational routines as coupling mechanisms: Policy, school administration, and the technical core. *American Educational Research Journal* 48(3), 586–619.

Stecher, B.M., Barron, S.L., Chun, T., Ross, K, (2000). *The Effects of the Washington State Education Reform on Schools and Classrooms*. Center for Research on Evaluations, Standards and Student Testing. RAND Research.

Steedle, J.T. (2010). Incentives, motivation, and performance on a low-stakes test of college learning. Paper presented at the 2010 Annual Meeting of the American Educational Research Association, Denver, CO.

Stiefel, L. Schwartz, A.E. & Ellen I.G. (2007). Disentangling the racial Test Score Gap: Probing the Evidence in a Large Urban School District. *Journal of Policy Analysis and Management*. Associate for Public Policy and Management.

Sundre, D. L., & Kitsantas, A. (2004). An exploration of the psychology of the examinee: Can examinee self-regulation and test-taking motivation predict consequential and non-consequential test performance? *Contemporary Educational Psychology*, 29, 6-26.

Supovitz, J. A. (2009). Can High Stakes Testing Leverage Educational Improvement? Prospects from the last decade of testing and accountability reform. *Journal of Educational Change*, 10(2), 211-227.

Topol, B., Olson, J., Roeber, E., & Hennon, P. (2013). *Getting to higher-quality assessments: Evaluating costs, benefits, and investment strategies*. Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education.

Triplett, C. F. & Barksdale, M. A. (2005). Third through sixth graders' perceptions of high- stakes testing. *Journal of Literacy Research*, 37(2), 237-260.

Tomasetto, C., Alparone, F.E., & Cadinu, M. (2011). Girls' Math Performance Under Stereotype Threat: The Moderating Role of Mothers' Gender Stereotypes. *Development Psychology*, 47(4), 943-949.

Toshalis, E. & Nakkula, M. (2012). *Motivation, Engagement, and Student Voice: The Students at the Center Series*. Jobs for the Future. Nellie Mae Education Foundation. <https://studentsatthecenterhub.org/>.

U.S. Department of Education (2016). “Elementary and Secondary Education Act of 1965”, Public Law 89-10. <https://www2.ed.gov/about/offices/list/oii/nonpublic/eseareauth.pdf>.

U.S. Department of Education (2015). “Every Student Succeeds Act (ESSA)”, Public Law 89-10; 79 Stat. 27. <https://www.ed.gov/esea>.

U.S. Department of Education (2002). “No Child Left Behind Act”, Public Law 107-110. <https://www2.ed.gov/nclb/landing.jhtml>.

U.S. Department of Education (1983). “National Commission on Excellence in Education”. A Nation at Risk: The Imperative for Educational Reform: A report to the Nation and the Secretary of Education. <https://www2.ed.gov/pubs/NatAtRisk/risk.html>.

Vanneman, A., Hamilton, L., Baldwin Anderson, J., & Rahman, T. (2009). Achievement Gaps: How Black and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress (NCES 2009-455). U.S. Department of Education, National Center for Education Statistics, Institute of Education Sciences. Washington, DC.

Wang, Z., & Brown, G. T. L. (2014). Hong Kong tertiary students' conceptions of assessment of academic ability. *Higher Education Research & Development*, 33(5), 1063–1077.

White, K. W., & Rosenbaum, J. E. (2008). Inside the black box of accountability: How high-stakes accountability alters school culture and the classification and treatment of students and teachers. *No Child Left Behind and the Reduction of the Achievement Gap: Sociological perspective on Federal Education Policy* (97-116). New York, NY: Routledge.

William, D., & Thompson, M. (2007). Integrating assessment with instruction: what will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: shaping teaching and learning* (53-82). Mahwah, NJ: Lawrence Erlbaum Associates.

Wise, S. L., & Cotten, M. R. (2009). Test-taking effort and score validity: The influence of student conceptions of assessment. In D. M. McInerney, G. T. L. Brown, G. Arief, & D. Liem (Eds.), *Student perspectives of assessment: What students can tell us about assessment for learning* (187–205). Charlotte, NC: Information Age Publishing.

Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment*, 10, 1-17.

Wise, S. L., & Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education*, 18(2), 163–183.

Wise, V. L., Wise, S. L., & Bhola, D. S. (2006). The generalizability of motivation filtering in improving test score validity. *Educational Assessment*, 11, 65-83.

Wolf, P. J. (2007). Academic improvement through regular assessment. *Peabody Journal of Education*, 82, 690-702.

Zerpa, C., Hachey, K., van Verneveld, C., & Simon, M. (2011). Modeling student motivation and students' ability estimates from a large-scale assessment of mathematics. *SAGE Open*, 1-9.