2018

The Relationship Between i-Ready Diagnostic and 10th Grade Students' High-Stakes Mathematics Test Scores Heath Andrew Thompson

Heath Andrew Thompson
Walden University

Walden University

College of Education

This is to certify that the doctoral study by

Heath Thompson

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Nicolae Nistor, Committee Chairperson, Education Faculty
Dr. Kathleen Claggett, Committee Member, Education Faculty
Dr. Markus Berndt, University Reviewer, Education Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2018

Abstract

The Relationship Between i-Ready Diagnostic and

10th Grade Students' High-Stakes Mathematics Test Scores

Heath Andrew Thompson

MA, Antioch University 2011

BM, University of Washington, 2005

Proposal Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Education

Walden University

March 2018

Abstract

Twenty percent of the 2013-2014 sophomore class at a Washington high school was failing high-stakes tests, making these students ineligible to graduate. In an attempt to help students identify their academic proficiency with respect to the Common Core Curricular Standards 9 months before the high-stakes exam, the high school recently introduced the adaptive diagnostic software i-Ready. Cognitive learning theories comprised the framework for this study, which posit that learning is dependent on previous knowledge and central to measuring performance levels. The purpose of this quantitative correlational project study was to examine whether 10^{th} grade students' achievement on i-Ready math scores ($N = 220$) could predict the subsequent high-stakes mathematics scores on the End of Course Exam while controlling for gender, ethnicity, and socioeconomic status. The i-Ready emerged as a statistically significant predictor of the End of Course Exam scores with $\beta = .64$ ($p < .001$), explaining $R^2 = .43$ of the criterion variance. Gender, ethnicity, and socioeconomic status had no significant moderating influence. The project deliverable as a result of this study was a position paper advising the use of the i-Ready as a predictor for the End of Course Exam at the high school under study. The implications for positive social change include allowing educators to use the i-Ready as an early warning system for students in danger of failing high-stakes exams. This study may help identify students at risk of not graduating who could benefit from instructional support.

The Relationship Between i-Ready Diagnostic and

10th Grade Students' High-Stakes Mathematics Test Scores

by

Heath Andrew Thompson

MA, Antioch University, 2011

BS, University of Washington, 2005

Proposal Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Education

Walden University

March 2018

Dedication

This study is dedicated to my Grandmother, Ann Thompson. Thank you for the countless times you inspired me, challenged me, and nurtured me. I also dedicate this work to my mom and dad, Bronwyn and Jeffrey Thompson. I would not have been able to achieve this dream without you always encouraging me to be my best. I also dedicate this to my wife Sarah Thompson, who has supported my crazy ambitions and loves me despite myself.

Acknowledgments

I acknowledge all the teachers who ever gave time in their day to help out a younger version of myself striving to be something more. In particular, I would like to acknowledge Jennifer Grajewski, Anthony Giles, Dr. Geoffrey Boers, Dr. Sean Ichiro Manes, Dr. Jonathan Saari, Dr. Rick Grimes, my project study chair Dr. Nicolae Nistor and my committee team Dr. Kathleen Claggett and Dr. Markus Berndt, for guiding me through my educational journey, pushing me beyond what I thought I could ever achieve, and affording me the privilege to learn from them.

Table of Contents

List of Tables

List of Figures

Section 1: The Problem

High-stakes testing has not been shown to be sufficient for its intended purpose of increasing academic success, providing equal access to quality education, or closing the achievement gap (Berliner, 2011). Since 2005, 4 out of 10 students have failed at least one section of their state's standardized tests nationally (Ravitch, 2016). High-stakes tests also have unintended consequences, including narrowing the curriculum, increased anxiety among students, and broadening the achievement gap among minority students in favor of White male students (Ruecker, 2013; Seymour & Garrison, 2015; von der Embse & Witmer, 2014). This study focused on high-stakes testing and attempted to predict which students were in danger of failing high-stakes tests at Evergreen High School (EHS, pseudonym). In the next section, I define the problem at a local school, provide justification for investigating the problem, and state the research question guiding this study. I also discuss the literature that (a) provides the theoretical framework for the study, and (b) include a description of the impact of high-stakes testing on schools and how it has impacted local communities. I also included implications for social change.

The Local Problem

Students are failing high-stakes tests at an alarmingly high rate (Nichols, Glass, & Berliner, 2012). At the state level, students' failure to pass these tests may significantly impact a student's ability to graduate. Students, parents, and educators need a method to identify students who are in danger of failing the high-stakes tests. In Washington State, 20% of the 2013-2014 sophomore class did not pass the mathematics portion, 23% failed the science section, and 14% failed the English section of the test (Office of

Superintendent of Public Instruction [OSPI], 2016). In the same year, 24% of the students in Washington State did not graduate. These students may have completed all coursework needed to graduate, obtained the necessary amount of credits, but due to the educational reform, the No Child Left Behind Act of 2001 (NCLB, 2002), they failed to demonstrate the knowledge required for a high school diploma (Daun-Barnett & St. John, 2012; Nichols et al., 2012). To assist students identify their academic proficiency with respect to the Common Core Curricular Standards at EHS, educators administered the i-Ready to all sophomore students in the fall, nine months before the End of Course Exam (EOCE). The i-Ready is a multiple-choice computerized adaptive diagnostic (CAD), a screening tool which assesses the same performance standards as the EOCE and measures the specific content knowledge and skills needed to pass the EOCE (Smarter Balanced Assessment Consortium, 2016).

Rationale

Students who fail high-stakes tests create a significant burden on the State of Washington: They are either delayed or prevented from graduating. Students must pass all portions of their high-stakes tests, including the mathematics EOCE, by the end of their sophomore year to be eligible for graduation. In the 2013-2014 year in the state of Washington, 16,000 sophomore students did not pass the mathematics portion of the EOCE, which meant 20% of the sophomore class had to retake at least one portion of their high-stakes tests to graduate. Students who failed will have one chance to retake the test in their junior year, and once again in their senior year if they fail a second time. This setback dramatically slows the process of graduation, reduces students' self-confidence, and strains schools and teachers. It means they cannot teach diverse, creative content because they need to focus their efforts on “teaching to the test” (Jennings & Bearak, 2014, p. 381). Students who do not pass these exams may be moving through the educational system without any method to determine if they lack the level of knowledge needed to succeed when taking certain high-stakes tests.

Therefore, educational stakeholders are currently in need of methods to detect students who are not likely to pass the EOCE. A critical tool that educators in Washington State use is the i-Ready, a form of progress monitoring. Though progress monitoring was created to evaluate students’ academic progress over time, diagnosis of learning disabilities, and students at risk of failure (Stevenson, 2015), no study was found that directly compares the relationship of progress monitoring measures and high-stakes tests at the high school level. Exploring the relationship between progress monitoring and high-stakes tests as a method to see whether success on the i-Ready is related to success on the high-stakes tests could be beneficial to educational stakeholders. The application of research-based mathematics support systems for 10^{th} grade students with low i-Ready scores might improve their mathematics skills, thus increasing their high-stakes test scores. Educators could frequently administer progress monitoring methods, such as the i-Ready. in class and quickly monitor student progress to determine which students might need intervention(s) before taking the EOCE (Hunley, Davies, & Miller, 2013).

If a student fails a high-stakes test, some research suggests that she will continue to fail the same exam without an intervention method to determine her level of knowledge (Miller, Bell, & McCallum, 2015; Singh, Märtsin, & Glasswell, 2015).

Currently, there is no such intervention method in one school district in the Northwestern United States, and researchers have not yet conducted such a study at the high school level in mathematics. In this study, I sought to determine if there was a relationship between a progress-monitoring measure and a high-stakes mathematics exam at one high school to establish a method that educators could use to provide students with the knowledge needed to pass their high-stakes tests, thereby reducing the probability of failure.

Previous studies have established that gender, ethnicity, and socioeconomic status (SES) could influence test scores as well (Berliner, 2011; Dixon-Román, Everson, & McArdle, 2013; Kearns, 2016). Researchers commonly refer to the moderation effect between test scores that often fall under this demographic data as the achievement gap (Freire, 1970). In this study, I sought to determine if the differences between gender, ethnicity, and SES might account for any differences between the test scores of the i-Ready and the EOCE.

Evidence of the Problem at the Local Level

At EHS, the passing rate for the mathematics EOCE was 60%, compared with 80% at the state level. EHS has consistently scored low, and the passing rate has decreased 5% since 2011 (OSPI, 2016). Approximately 30% of students fail to graduate on time. There could be many factors that cause EHS to perform lower than other schools, but student ethnicity and low SES could impact scores (Huang, 2015). At EHS, 70% percent of students live below the national poverty line; 25% are Hispanic, 20% Black, 20% Asian and 10% Pacific Islander or two or more other races (OSPI, 2016). In keeping with the national trend, these students consistently score at a 30% lower rate on high-stakes tests than White students above the poverty line (Nichols et al., 2012).

Evidence of the Problem from the Professional Literature

Many factors contribute to why students struggle with high-stakes testing. These exams and their consequences may lead to an increase in student testing anxiety, which can severely affect their emotional and cognitive ability to adequately perform on the tests (Kearns, 2016). Without accountability methods in place, these tests may be showing inadequacies between schools, lack of support for educators, and a failure to align curriculums to the tests (Mora, 2011; Nichols et al., 2011; Starr & Spellings, 2014). Though education reforms that mandate tests such as NCLB were designed to provide educational equality, minority and low SES students face negative impacts the strongest, showing that high-stakes testing has further excluded these groups of students rather than supported them (Kettler, Russel, & Puryear, 2015; Ruecker, 2013). High-stakes tests may also raise dropout rates, delay graduation, and encourage students to drop out (Huddleston, 2014).

Nationally, there is evidence of a substantial achievement gap (Huang, 2015). Research has shown White students have performed continuously better on standardized tests than Black and Hispanic students; high SES students have scored higher than low SES students; and test anxiety has a correlation to poverty (Huang, 2015; Nichols et al., 2011; Nichols & Valenzuela, 2015). As a result, teachers have begun narrowing the curriculum, eliminating elective classes, and limiting creative outlets for students specifically in low SES and minority-rich schools, thus creating further inequalities

(Birdwell, 2012; Jennings & Bearak, 2014; West, 2012). Currently, there is no method of determining which students may be in danger of failing high-stakes tests until after they have either passed or failed the exam in question (Hunley et al., 2013). In this study, I sought to compare the relationship between the i-Ready, a computerized adaptive diagnostic (CAD) screening tool and the mathematics EOCE. Since I found a correlation, it could provide teachers and administrators with a method that can track student progress in the critical skills needed for their high-stakes tests (Kirkham & Lamley, 2014).

Definition of Terms

End of Course Exam: Washington State's standardized assessment for 10^{th} graders in mathematics, English Language Arts, and Science. It is used to determine if a student is eligible to receive a diploma. This assessment measures student's progress toward meeting Washington's Common Core State Standards (OSPI, 2016).

Common Core Curricular Standards: A national educational initiative that specifies the grade-level content for English Language Arts, mathematics, and science and seeks to establish unified benchmarks of knowledge for every student in the United States (Common Core State Standards Initiative, 2010).

Progress Monitoring: A method of assessing student performance towards academic goals using curriculum-based assessments of computerized adaptive diagnostics that measure student progress on a regular basis (Hosp & Hosp, 2003).

Curriculum-Based Measures: Brief curriculum-based tests given to students regularly with the goal of monitoring their progress toward mastery of a content skill (Fuchs & Deno, 1981).

Computerized Adaptive Diagnostics: A computer-based assessment that adjusts to the level of knowledge of the student using item response theory (IRT). IRT is a method that predicts a student's ability to achieve a score based on unobservable latent traits such as aptitude, achievement, and personality (Rasch, 1960; Mao & Xin, 2013; Wainer et al., 2000).

The Achievement Gap: A lingering disparity of grade point average, drop-out rates, standardized test scores, and college enrollment that researchers define by gender, ethnicity and low SES groups (Freire, 1970). Many researchers argue that the achievement gap has broadened with educational reforms such as NCLB (Huang, 2015; Morris, 2015).

No Child Left Behind Act: A national initiative that governed K-12 education. NCLB required that educators evaluate every student through the use of standardized examinations, placing an emphasis on a narrowed curriculum at a national level by establishing sanctions for low-performing schools (NCLB, 2002).

Every Student Succeeds Act: The federal replacement of NCLB, the ESSA removes sanctions on schools but places a continued emphasis of annual statewide standardized assessments of all students (Every Student Succeeds Act [ESSA], 2015).

Significance of the Study

In this study, I supported Walden's mission of positive social change by recognizing the lack of high-stakes test failure prevention measures. Despite some promising research conducted at the elementary level, the research involving the utilization of progress monitoring measures for test prediction at the high school level is negligible, particularly in the field of math which has the highest failure rate both at the local and national level (Codding, Petscheder, & Truckenmiller, 2015). In this study, I attempted to give a valuable tool for educators at the secondary level by providing them with the ability to know which students might be in danger of failing these standardized exams before they take them, thereby giving educators the opportunity to employ additional academic support and remediation to students in danger of failing. The results of the study could have the potential to create positive social change by giving schools new tools to address the needs of students who may be in danger of failing high-stakes tests. This reduction could help reduce learner anxiety, prevent further narrowing of the school curriculum, lessen the achievement gap, and provide meaningful discussion on the effects of high-stakes testing.

In this project study, I focused on the lack of student failure prediction methods for the high-stakes tests that are required for graduation from high school. The significance of this research study is to establish a method of predicting performance on the EOCE required for graduation using progress monitoring measures. Progress monitoring measures offer diagnostic information to help understand student growth, create equitable learning practices, and close the achievement gap.

Research Questions and Hypotheses

The purpose of this study was to determine if there is a significant relationship between achievement in mathematics on the i-Ready and achievement in mathematics on the EOCE. The guiding research question for this study was as follows: “Does a statistically significant correlation exist between student achievement scores in mathematics on the i-Ready and student achievement scores in mathematics on the EOCE?" By analyzing demographic data in addition to the two tests, the goal was to determine the i-Ready's ability to predict success on the EOCE. Consequently, the following research questions guided this study. The following hypotheses were tested to answer these research questions:

RQ1: Does the i-Ready score predict mathematics scores on the EOCE?

H_{01} The i-Ready score does not predict EOCE scores.

Ha_1 The i-Ready score predicts EOCE scores.

RQ2: Do gender, ethnicity, and socioeconomic status moderate the relationship between i-Ready scores and mathematics scores on the EOCE?

H_{02} Gender, ethnicity, and socioeconomic status do not moderate the relationship between the i-Ready score and EOCE scores.

Ha_2 Gender, ethnicity, and socioeconomic status moderate the relationship between the i-Ready score and EOCE scores.

The next subsection of the study establishes the theoretical framework and includes a literature review for supporting the research questions for the study.

Review of the Literature

The Every Student Succeeds Act (ESSA, formerly NCLB) is an educational reform intended to improve student achievement (ESSA, 2015). According to this regulation, public schools must pay a significant amount of money and attention to administer high-stakes tests. Teachers must put forth substantial effort outside of classes and their curriculum to prevent students from failing tests and thus prevent sanctions to both the students and the school.

In the literature review section, I address these problems and discuss cognitive learning theories and how they relate to high-stakes testing and progress monitoring. Additionally, this literature review describes high-stakes testing and current resources related to high-stakes testing's effect on teaching. Lastly, I discuss progress monitoring as a means to predict test scores on standardized exams. To find information on these categories, I used the ERIC database. with the following terms: *high-stakes testing*, the *impact of high-stakes testing on teaching*, *curriculum-based measures, computerized adaptive diagnostics*, *progress monitoring*, and *cognitive learning theory*. Within ERIC, there were many relevant, peer-reviewed sources conducted within the last 5 years. I used information dating back to 1951 for foundational purposes. These sources provided this study with a more informed perspective of high-stakes testing and its relationship to teaching, as well as information about possible prevention measures and current gaps in the literature.

Theoretical Foundation

Piaget's (1954) theory of constructivism formed the basis for this study. Constructivism, according to Piaget, concerned how humans create meaning based on their relationship to their environments by building upon prior knowledge. Piaget's cognitive constructivism indicates that student activation of prior knowledge is necessary to solve new problems. Piaget's theories suggested that learning is structured as levels and is not random or a constant, which implied that these levels could be measured (Gray, 1978).

Piaget was a developmental psychologist who believed that knowledge is the ability to adapt to survive in an environment (Piaget, 1964). When humans face a challenge with information that contradicts prior knowledge, they must adapt prior knowledge to return to a state of balance with the environment. This adapting, or assimilation of knowledge, is an often unconscious, invisible action in which the mind constructs meaning. In relationship to this study, Piaget implies that students possess various levels of knowledge that differs depending on their experiences (Piaget, 1959). Cognitive structures of learners do not change without an external stimulus, which suggests that these structures can go a period without changing beyond certain thresholds (Piaget, 1951). In following with this theory, researchers have been able to measure levels of knowledge (Incantalupo, Treagust, & Koul, 2014; Kara, 2015), as well as use those levels to predict future outcomes of examinations (Campbell, Espin, & McMaster, 2013; Miller et al., 2015; Shapiro, Dennis, & Fu, 2015).

Piaget's theory of constructivism carries many implications for education and testing, and establishes a paradigm for teaching and learning (Slavin & Davis, 2006). In relationship to this study, constructivism implies that assessments such as curriculum-based measures (CBMs) hold measurable information about the learner's level of knowledge that could predict the outcome of high-stakes tests (Codding et al., 2015; McGlinchey & Hixson, 2004; O'Malley & Pierce, 1996). Constructivism posits that knowledge functions in levels and suggests that these levels can be measured, and it may be possible to use tools designed to identify their level of comprehension (such as the i-Ready) to predict proficiency on high-stakes tests (Shapiro, Solari, & Petsher, 2008).

In the same line, Vygotsky's contributions to constructivism form the second part of this study's theoretical framework through expanding cognitive constructivism to a learner's social environment. Vygotsky (1978) theorized that cognitive development was intrinsically related to a learner's social environment. Students learn based upon social constructs of meaning that are established by adults and peers. For Vygotsky, cultural forces cause the motivation to change or grow (Vygotsky, 1986). Vygotsky theorized the model of the Zone of Proximal Development, which establishes a structure of proximity among the most challenging activity a student can accomplish alone and the most challenging activity they can do with assistance (Vygotsky, 1978). In relationship to this study, Vygotsky's theories frame the notion that with a method of predicting student failure rates of standardized examinations, educators can utilize the Zone of Proximal Development to take measures to improve student performance on high-stakes test scores.

Similarly, the theories of Ausubel (1960) form the third part of this study's theoretical framework by applying constructivism to personal relevance. Ausubel theorized that for learning to be effective, it must be meaningful to the learner. Ausubel theorized that leaners used advanced organizers, or relevant, abstract ideas that allow students to draw upon prior knowledge to help relate it to new material (Ausubel, 1960; Ausubel & Fitzgerald, 1962). Meaningful learning, as opposed to rote learning, is the integration of new thoughts into current cognitive configurations, anchored by prior knowledge. Ausubel believed new material should relate to the learner's existing knowledge and be presented clearly to best connect new material into preexisting mental structures (Ausubel, 1960). Like Piaget and Vygotsky, Ausubel felt that students were not blank slates and that learning progressed in a structure, or level, of knowledge before connecting to a new level. In relationship to this study, Ausubel demonstrates a linear instructional method using advanced organizers, suggesting that prior learning and new knowledge connect to one another through assimilation. Developing a method of predicting the performance on high-stakes tests for students at the high school level would provide educators with the opportunity to develop the advance organizers necessary to connect new concepts needed to achieve success for students, thereby reducing the failure rate on high-stakes tests.

The theories of Anderson (1984), Gagné (1979), and Rumelhart (1980) comprised the fourth part of this study's theoretical framework. These theorists all gave significant contributions to schema theory, which provides a mental representation of knowledge, often drawing comparisons of the human mind to the processing of a computer, storing data into memory. According to schema theory, schemas (or schemata) are units of knowledge that are stored onto empty files, or shelves, in the brain. Schema theory explains how prior knowledge is organized into long-term memory and focuses on the job prior knowledge plays in processing new knowledge. In relationship to this study, schema theory is similar to constructivism, which implies that knowledge works in levels, and these levels remain constant until they are changed much in the manner a computer processes information.

The teachings of Bloom completed this study's theoretical framework. Bloom (1956) integrated the teachings of Piaget and Vygotsky in part to form Bloom's taxonomy, which is a visual, tiered structure of organizing levels of thinking that progress in complexity. According to this theory, having established prior knowledge is an essential step before springing to the next level. Bloom's taxonomy illustrates that students who fail high-stakes tests sometimes have not achieved the sufficient knowledge levels needed to achieve success on these tests while providing a design for instructional planning for educators to guide students through the processes necessary to progress to the required knowledge levels (Krathwohl, 2002). Bloom's theory argues that achievement is contingent on prior knowledge and that knowledge works in levels that will not change before assimilation of new information, which implies that these levels can be measured to predict the outcomes of high-stakes test scores.

Review of the Broader Problem

High-stakes tests are a method of assessment that students, teachers, school districts, and states use to measure a learner's ability and enforce academic accountability (Brennan, 2015). Data found from high-stakes tests can impact decisions at the high school level, such as grade retention, school curriculum, and student graduation (Hursh, 2013). There are many ways to assess, monitor, and evaluate students, however educational reforms such as No Child Left Behind (NCLB, 2002), and the more recent Every Student Succeeds Act (ESSA, 2015) place importance on high-stakes testing as the proprietary method of educational progress (Neely, 2015). High-stakes testing is conducted using standardized tests that are administered uniformly to students throughout every state in the United States in the 4th and 8th grades as well as the high school level (ESSA, 2015). These examinations are summative and are typically administered at the culmination of a course to assess if a student has mastered the content of a corresponding class, subject, or unit (Lit & Lotan, 2013). Many scholars agree that assessment is a vital element of education (Dessoff, 2012; Finn & Roediger, 2013; Marzano & Heflebower, 2011). The discussion among scholars has recently shifted towards how high-stakes tests play a role in education, and how they have impacted the classroom (Conley, 2015; Harland, McLean, Wass, Miller, & Sim, 2015; Huang, 2015).

To meet federal and state legislation, Washington State has initiated the use of high-stakes tests including EOCE in the elementary and middle school grades, and at the end of certain classes in 9th and 10th grade (OSPI, 2016). The EOCE is a high-stakes test designed to provide summative data and is used determine if a student is eligible for graduation. If students do not pass the EOCE in the 10th grade, they are expected to re-take it and pass by 12th grade (OSPI, 2016). Teachers administer the EOCE at the end of four courses including 9th grade Math I, 10th grade Math II, Biology, and English Language Arts. These high-stakes tests provide assessment data used to fulfill adequate yearly progress requirements by ensuring that students meet established performance standards (OSPI, 2016). While high-stakes testing has brought about a means of monitoring student performance, they have been fraught with extensive amounts of criticism for many reasons (Minarechová, 2012).

The Impact of High-Stakes Tests on Teaching

Current research notes several adverse effects of high-stakes tests. Many scholars argue educational reforms such as NCLB and ESSA have endangered the democratic mission of education, destabilized the agency of school curriculums, and created undue amounts of anxiety upon school communities (Birdwell, 2012; Brennan, 2015). One of the most common results of standardized tests is curriculum narrowing, or the displacement of other subjects including the arts, technical education programs, physical education, and literature (Backer & Lewis, 2015; Berliner, 2011; Jennings & Bearak, 2014; Polesel, Rice, & Dulfur, 2014). Excessive test preparation for standardized exams can lead to a lack of connection with the content, and boredom with school (Mora, 2011). With accumulative pressure to only teach what is on each test, the classroom becomes increasingly scripted, less creative, and provides little opportunity for diversity, personal exploration, and discovery (Au, 2011; Becker, 2015; Thompson & Allen, 2012). High-stakes testing determines academic success (Erskine, 2014), and schools that do not meet adequate yearly progress requirements have had music, art, poetry, and physical education classes either greatly diminished or entirely removed from their schools (Dymoke, 2012; West, 2012). Other classes more traditionally thought of as core subjects such as history have also had their curriculums significantly narrowed, rendering a curriculum lacking in engaging content (Maranto, 2015; Starr, 2012). In numerous states, as many as five or more high-stakes tests are required to graduate (Hursh, 2013). High-stakes testing can create many logistical issues that can impede the learning process, such as requiring time outside of classes to study and take the tests. If failure to pass occurs, students may repeat this cycle of testing without academic intervention (Duan-Barnett & St. John, 2012; Newhouse & Terricone, 2014; Nichols & Valenzuela, 2013). Some researchers argue that high-stakes testing has failed to generate evidence that it improves student grades or dropout rates. (Nichols et al., 2012).

The National Association of Educational Progress (NEAP) suggested that minority and low SES students have consistently been scoring 30% lower than White male students (Nichols et al., 2012). One of NCLB initiatives was to offer an equal chance for all students to succeed academically (NCLB, 2002). Due to the threat of sanctions schools could receive based on test performance, and the reallocation of funds to schools that perform better than others, the resulting effect has been further school inequity (Lewis & Hardy, 2015; Starr, 2012). Some scholars contend that high-stakes testing assumes an environment of meritocracy or the notion that every student taking the tests is afforded the same opportunities to learn the required material (Cornelius, 2011; Ydesen, 2014). Tests are not entirely equal and can reflect biases that favor specific ethnicities, cultures, or genders (Ford & Yelms, 2012). There are also socioeconomic factors that influence a student's access to education. Schools that initially perform low will often face sanctions in the form of a decrease in funding, staffing, and supplemental support (Young & Cox, 2014). This deficit reproduces a lack of resources to our most vulnerable students (Huddleston, 2014; Yell, Katsiyannis, Collins, & Losinski, 2012). Achievement gaps in test scores have shown to correspond with higher drop-out rates among minority students versus White males, indicating that these tests may be adding to the inequality that they measure (Au, 2013). Having to re-test students, along with higher dropout and re-enrollment rates create additional challenges to school funding and distribution of resources for all students (Barrat, Berliner, & Fong, 2012). From 2005 to 2015, minority students graduated at a rate of 48% to 75%, consistently 10-30% lower than White male students (Thompson & Allen, 2012). Minority-dense schools, such as schools in urban settings, are correlated with less student funding which in effect has led to teacher shortages for minority students (Dixon-Román et al., 2013, Kettler et al., 2015). In relationship to the minority achievement gap, standardized test scores show that low SES also affects student performance (Kettler et al., 2015; Reucker, 2013). While the ESSA enacted re-authorization of mandated standardized exams, researchers suggest that the achievement gap is widening at a national level (ESSA, 2015; Huang, 2015). Standardized testing, particularly in low SES and minority-rich schools are thought by some to be a failure in educational policy, and research has demonstrated a link between inequality in income and inequality in education (Berliner, 2013; Dixon-Román et al., 2013; Kearns, 2016).

The lack of ability and resources to pass high-stakes tests could create an inordinate amount of pressure on students and educators. Without methods to intervene, measure, or prevent the continued testing failure, students at schools with low tests scores often feel an excessive amount of test anxiety (von der Embse & Witmer, 2014). Test anxiety has been reported to be significantly high in response to high-stakes tests (Segool, Carlson, Goforth, von der Embse, & Barterian, 2013). The daunting prospect of having to pass five or more high-stakes tests, without the ability to predict their performance, creates low levels of student engagement, motivation, and student self-efficacy (Tempel

& Neuman, 2014; Ünal-Karagüven, 2015). These behavioral changes are of concern because they affect more than just a student's performance on a test, but also their view of what education is and the role it plays in their lives (Brennan, 2015). Some scholars argue that high-stakes testing impairs learning rather than improving it (Davis & Chan, 2015; Finn & Roediger, 2013).

Standardized, high-stakes testing has become the chief determinant of educational performance and success both for students and schools (Mitra, Mann, & Hlavacik, 2016). At the high school level, these tests determine the eligibility for student graduation. At EHS, there is currently no method to predict how students will perform on these exams. This study sought to mitigate the disparity of research by determining if progress monitoring methods will identify students that may need interventions to avoid failing these tests, thereby increasing the graduation rate and providing better educational equity.

Progress Monitoring

The repeated collection of student data over time, or progress monitoring, has allowed educators to identify and address gaps in learning. Educators created progress monitoring as a way to monitor teaching (Deno, Fuchs, Marston, & Shin, 2001; Hosp, Hosp, & Howell, 2007; Johnson, Jenkins, & Petscher, 2010). Progress monitoring has more recently developed into a method of collecting data for a broader range of applications, such as response to intervention, special education referral, curriculum alignment, and some forms of test prediction (Maier et al., 2016; Van Norman & Christ, 2016).

A CBM is one method used for progress monitoring. In the 1970s and 1980s, the Institute for Research on Learning Disabilities developed measurement to detect alignment of instructional effectiveness and curricular instruction for use with elementary students in reading, written expression, and mathematics, resulting in CBMs (Espin et al., 2013). CBMs are assessments that measure specific skills demonstrated by students in the classroom, as opposed to high-stakes tests which measure a student's broad understanding of a curriculum (Hauerwas, Brown, & Scott, 2013). CBMs are designed to indicate general performance in a given academic area and must be short in length, easy to dispense and grade, and repeatable (Campbell et al., 2013; Deno, 1985). CBMs must also be valid and reliable to provide meaningful indicators for educators (Hosp, Hensley, Huddle, & Ford; 2014). Researchers demonstrated that CBMs could be useful for identifying students with learning disabilities (Jenkins, Schulze, Marti, & Harbaugh, 2017; Marston, Mirkin, & Deno, 1984), and was a significant factor in creating interventions for remedial learning and alternative testing (Fuchs & Fuchs, 1992). Early research found that CBMs were useful in working with students who struggled with basic concepts (Fuchs, Hamlett, & Stecker, 1991) and that educators could use the tool with general student populations (Marston, Deno, Kim, Diment, & Rogers, 1995). The reading CBM is the most common version of this instrument (Ball & Christ, 2012). Typically, reading CBMs are timed assessments of oral reading fluency that can be administered either individually or in group settings.

CBMs, though developed in the 1970s, have come back to the forefront as a means to help meet the requirements mandated by high-stakes testing (Hosp & Hosp, 2003). Scholars conducted CBM research in reading, writing, and mathematics at the elementary level (Codding, Mercer, Connell, Fiorello, & Kleinert, 2016; Hensley, Rankin, & Hosp, 2017; McMaster et al., 2017). CBM indicators such as oral reading fluency were initially not found to be suitable methods of writing proficiency for high school students (Parker, Tindal, & Hasbrouck, 1991). Over time, researchers found that more comprehensive CBMs could predict competence at the high school levels (Espin, 1999). More recently, educators have used mathematics CBMs for the progress monitoring that measures computations and concepts (Fuchs & Fuchs, 2006). A growing amount of research shows CBMs in math are reliable and valid screening measures for predicting performance on high-stakes tests in mathematics (Fuchs & Fuchs, 2007; Keller-Margulis, Shapiro, & Hintze, 2008).

An emerging form of progress monitoring is the use of computer adaptive diagnostics (CAD). CADs are the testing of students using a computer with each successive question changing depending on the accuracy of each student’s answer to the previous question (Shapiro et al., 2015). CADs are built using item response theory (IRT), a theory based on item-level responses which continually adjusts the test items, as opposed to static tests such as traditional CBMs (Petscher, Mitchell, & Foorman, 2015). A CAD is structured on learning progressions of skill competencies as defined by curriculum standards that span across grade levels (Shapiro, 2014).

CADs adapt on student ability, which can then be modeled to demonstrate student growth levels and can measure the accuracy of response as alongside the timing of a response in contrast to a test or traditional CBM. CADs can increase test efficiency, optimize test length, and increase the security of testing (Latu & Chapman, 2002). CADs can be used to estimate more extensive ranges of ability since IRT demonstrates a relationship between an item and ability or trait of the student (Lazendic & Martin, 2018). CADs might provide better testing methods and procedures for predicting what students may do on test scores by using the IRT ability scales. Some researchers suggest that CADs have better validity, reliability, and flexibility than other large testing instruments.

The results of some studies suggest that CADs have demonstrated strong relationships alongside traditional CBMs with high-stakes test scores as a method of prediction (Shapiro et al., 2015). Some researchers analyzed the value of data on progress monitoring, both for CBMs and CADs. Many of the studies that attempt to predict student outcomes depend on the degree to which slope or rate of change predicts outcomes of student exams (Miller et al., 2015). In this study, I sought to extend to the review of literature related to progress monitoring and high-stakes testing, which I summarized in the next section.

Using Progress Monitoring for Test Prediction

In this study, I examined the relationship of student performance on the i-Ready and performance on the EOCE. Researchers have conducted tests on the validity of CADs as predictors of high-stakes tests (Good, Simmons, & Kameenui, 2001; Hintze & Silberglitt, 2005; Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008; Stage & Jacobsen, 2001). Very few studies have been conducted at the secondary levels (Speer, King, & Howell, 2015), and there is a gap in the literature regarding this field of research in mathematics scores. There are no published studies on the relationship between CBMs or CAD's and performance on Washington State's EOCE.

Researchers conducted one of the first prediction studies using CBMs in 2001 (Crawford, Tindal, & Stieber, 2001). A sample of 51 students from a rural school in the Northwest United States comprised the study. Researchers administered CBM for two years and then collected for second and third grade students. Moderate but significant correlations were demonstrated using standardized test scores in reading and CBM scores at the second and third grades, at $r = .66$ and $r = .60$ correspondingly.

Additionally, researchers conducted a study between CBMs and high-stakes test scores in 2001 using 173 fourth grade students (Stage & Jacobsen, 2001). A growth curve analysis was conducted to establish a relationship between reading CBMs and high-stakes test scores in the Northwestern United States. The study found that CBMs reliably predict test scores by using CBMs administered in September to predict test scores in May. All ANOVAs were statistically significant at $F(2,170) \geq 29$, $p < .0001$. Results indicated that researchers could use prediction methods using CBMs for providing additional academic support to students in need.

Concurrently, researchers conducted a study in 2001 examining the strength of the relationship of CBMs and the high-stakes test in reading for grades K-3 (Good et al., 2001). The researchers explored predictive validity using a correlational study. The researchers found a significant correlation, between CBMs and the exam results, finding 96% of 198 students that met the CBM benchmark of proficiency met or exceeded expectations on the exam. Forty-six percent of students scored below the level of proficiency on the CBM and 28% of those students met expectations on the exam. These studies serve to suggest that there may be a relationship between CBMs and standardized exams.

Alongside these studies, in 2007 researchers evaluated the relationship between CBMs, a standardized exam in English Language Arts and the Stanford Achievement Test for third grade English students (Roehrig et al., 2008). The researchers found high correlations between winter CBM scores and both outcome measures (.70-.71). When recalibrated with ROC curves, the cut scores increased by 1.7%, signifying that a large number of the student population had increased scores. While the prominent research was made establishing a relationship between CBMs and high-stakes tests, these early studies focused on elementary students, and in the field of reading only.

In another recent study, researchers assessed CBM scores as predictors for high-stakes tests in reading for 448 third grade students in the Southeastern United States using a correlational method (Miller et al., 2015). The results indicated strong predictive performance with a correlation coefficient of $r = .58$, $p < .001$. The results of the study suggested that researchers could use progress monitoring CBMs to predict the scores on high-stakes tests. Additionally, researchers have conducted more extensive studies with CBMs and high-stakes test scores in math and reading at the elementary level.

Researchers conducted a study at the elementary level comparing CBMs to high-stakes math and reading tests with 1,242 third grade students across three years in a correlational study in the Southeastern United States (Bell, Taylor, McCallum, Coles, & Hays, 2015). The results indicated that all interrelations among the CBMs were statistically significant, with moderate correlations for reading (*r* = .55) and math (*r* = .47). Implications concluded that comprehensive techniques should be used to identify at-risk students.

Additionally, some researchers have assessed elementary English Language Learner (ELL) students. In 2009, a study was conducted using correlations between CBMs and standardized assessments for fifth grade ELL students (Muyskens, Betts, Lau, & Marston, 2009). The study utilized logistic regression analysis. The study found that the CBM was a substantial predictor of ELL student performance on high-stakes tests, as well as some individual language groups with an approximate classification percentage power of 75%.

In contrast with the work assessing reading and writing comprehension, researchers examined mathematics test prediction using CBMs. In 2014 researchers conducted one of the first studies comparing math CBMs and high-stakes tests in mathematics (Jitendra, Dupuis, & Zaslofsky, 2014). The study used 136 third grade math students in a correlational study. Results showed that the CBM demonstrated significant predictive validity with correlations ranging from *r* = .67 to .71. Implications for practice suggest further assessment of this type of data, as well as using interventions for students.

In another recent study, researchers investigated CADs and CBMs in relationship to an assessment in mathematics for third, fourth, and fifth grade students in the Northeastern United States (Shapiro et al., 2015). In this study, researchers attempted to determine the degree to which slope or rate of change predicted student outcomes for both CADs and CBMs. The researchers assessed scores of 250 students using two hierarchical linear models and compared both methods of progress monitoring to high-stakes test scores. The results indicated significant positive linear slopes throughout measurement and found both progress monitoring methods to be significant predictors of success on mathematics test scores, with a slight advantage for CADs. Implications included the need for addressing the literature gap in research regarding progress monitoring for mathematics and finding that both methods were capable of predicting performance in early grades. This study also addressed the gap in the literature regarding the emerging use of CADs, noting that more research is needed.

In conjunction with these studies, researchers assessed CBMs with standardized exams for students that have low SES. In 2014, a study investigated the relationship between a fall, winter, and spring CBMs and a standardized exam for third graders in reading in the Southern United States using multiple regression analysis (Kirkham & Lampley, 2014). Roughly 47% were of low SES. The study found a robust predictive relationship between the CBM and exam scores. The zero correlations in this study ranged from r = .70 to .74, indicating a strong positive relationship.

Alongside studies that assess elementary students, there has been some work done in the intermediate grades. In 2013, researchers conducted one of the first studies that investigated the relationship between CBMs and test scores for middle school students (Hunley et al., 2013). The researchers used 75 middle school students that were assessed with a CBM before they took a standardized exam in reading using correlational analysis. The authors found a strong positive and significant correlation between the two measures, with $r = .76$, $p < .001$.

Additionally, researchers conducted studies with students in middle school grade levels. In in 2015 a researcher attempted to predict the relationship between three different CBMs and a standardized exam for seventh and eighth graders in the Midwestern United States using logistic regression (Stevenson, 2015). The finding in the study indicated that one of the three CBMs was a stronger predictor than the other two, for both seventh and eighth graders with grade 7 $e^{\beta} = 1.75$, grade 8 $e^{\beta} = 1.68$.

In another study, a researcher examined scores for a CAD in reading and mathematics for middle schoolers (Kettler, 2011). The researcher examined 463 students at the middle school level in the Midwestern United States using a correlational method. The results found medium correlations for both math (r = .62) and language (r = .61), indicating CADs can be efficient, valid predictors for middle school students.

Furthermore, in one study researchers investigated the relationship of CBMs and standardized writing exams for 83 middle school students in the Northwestern United States using a two-way mixed model of analysis of variance (López & Thompson, 2011). The researchers found strong moderate correlations in sixth grade (*ICC* (2,1) = .950, p < .01), and stronger correlations in seventh grade (*ICC* (2,1) = .961, p < .01) and eighth grade (*ICC* (2,1) = .987, p < .01). Implications of the study found that CBMs may give educational stakeholders material that would aid them in supplementing the learning of students.

There has also been some work done for standardized test score prediction in science at the middle school level. In a study in 2013, researchers investigated the relationship of science CBMs with the corresponding standardized test for 198 seventh grade students in the Midwestern United States using a correlational method (Espin et al., 2013). Results provided initial support of CBMs being an indicator of progress in science with r = .55 to .76. Results indicated that some CBMs do provide technical adequacy in predicting performance on high-stakes test scores.

Despite a gap in the literature regarding CBMs as predictors of standardized tests at the high school level, there has been some research in this area. In one study researchers compared the validity of CBMs predicting test scores in reading, mathematics, and writing for 41 high school students with learning disabilities using descriptive statistics and bivariate correlations (Hosp et al., 2014). The findings suggested that CBMs have strong correlations for reading (r = .71), math (r = .67), and writing (r = .53). The results suggested that students with learning disabilities would benefit from test prediction and intervention using CBMs. The studies discussed to this point have focused primarily on predicting performance on standardized tests in reading, with some exceptions in mathematics and science. This review discussed the various ways to determine the exactitude of screening measures.

In this study, I attempted to fill a part of the gap for CBMs ability to predict scores on high-stakes tests for 10th graders. I also utilized the response to intervention method for high school students. While response to intervention is focused on early detection and prevention, secondary students at risk of failing their high-stakes tests required for graduation would receive significant benefit to identification and intervention. Many scholars agree that educators should address this gap (Hosp et al., 2014; Hudson, Browder, & Wood, 2013; Shapiro et al., 2015).

Learning Analytics

Learning analytics (LA), an emerging form of collecting, analyzing, and reporting data about learners, has emerged with the ever-growing use of technology in educational settings (Shum & Ferguson, 2012). LA is a form of extrapolation of relevant data from massive amounts of digital information to optimize learning (Baker & Inventado, 2014). Small sets of validated data taken from CADs and other digital learning tools could contribute to larger populations of LA data. LA could then perform quick, unobtrusive projections of behavior or outcomes in educational contexts (Ferguson & Shum, 2012). As more educators contribute learner data from the use of CADs, LA could use this data to improve learning. Likewise, educational stakeholders seek to develop and apply new tools that can discover patterns in data and allow predictions. Through the use of regression analysis, I sought to find a relationship between a CAD and a high-stakes test. On a larger scale, tools for learning prediction that function like this study could be used to build models that could enhance online learning and assessment systems (Bienkowski, Feng, & Means, 2012, Cheng, 2010).

While prevalent in higher education, the use of technology to facilitate educational tasks, such as diagnostics and assessment, has only been slowly adopted at the high school level (Slavin, Cheung, Holmes, Madden, & Chamberlain, 2013). If CADs, such as the i-Ready, can be shown to be effective predictors of student proficiency on high-stakes tests, it may encourage educational stakeholders to adopt additional technology that would allow enhanced access to large amounts of student data. This advancement would allow systems, such as LA, to calculate and detect various ways to improve access to student learning at the high school level (Wagner & Ice, 2012).

Implications

In this study, I examined the relationship between the i-Ready scores and the EOCE scores of 10^{th} grade students at an urban high school in the Pacific Northwest. Research on the use of CBMs as well as CADs for prediction at the high school level is limited. The findings of this study could encourage other researchers to contribute to the present gap in the literature on predicting high-stakes test scores. At the local level, the results of this study could prompt the use of CADs to predict proficiency on high-stakes tests. Additionally, the findings of this study could reduce student test anxiety, prevent further narrowing of the curriculum, and narrow the achievement gap. Lastly, the findings of this study could help inspire schools to embrace the use of electronic forms of assessment, such as CADs. If they did, it would necessitate a move toward computerized gathering, analyzing, and interpreting learning data, thereby further opening the door for LA to improve student learning at the high school level.

Based on the review of the literature, one future project incorporating the research of this study would be a policy recommendation paper that would include the findings of the study and additional suggestions for a test prediction and intervention plan based on the statistical outcomes. Another possible project would be a professional development unit teaching administrators about an implementation and evaluation plan for test prediction. If there were a connection to i-Ready and EOCE scores, staff education and training that showed the relationship between the i-Ready and the EOCE could also show the need to develop new methods of intervention for students at risk of failing their high-stakes tests.

Summary

Students who do not pass the EOCE who may have met all other requirements to graduate, including a passing GPA, may not have developed the skills and knowledge currently necessary for graduation. Many of these students have had to endure multiple years of an inability to pass high-stakes tests. The result of this has been students having to retake these tests until they graduate or until they drop out of high school without any method of intervention. Current research has explained how these tests may have unintended consequences, such as creating test anxiety, the widening of the achievement gap, and the narrowing of school curriculums. Against this background, I examined school archival data for a population of students registered in a public urban high school to conclude if students who did not pass a specific part of the i-Ready CAD also were unable to pass the same content area of the EOCE.

In the literature review of this study, I explained the role of prior knowledge in cognitive learning as an essential link for all learning, and how cognitive researchers theorized that new knowledge must assimilate with prior knowledge before progressing to the next level. In the review of the literature, I described how researchers suggested that these levels may last for periods of time, allowing them to be measured. Additionally, in the literature review I described studies concerning current legislation regarding high-stakes testing, the effect these tests have on teaching, and current measures taken to remediate the consequences of testing using progress monitoring measures. In section 1, I discussed the low mathematics performances on the EOCE at EHS and emphasized the significance of this problem as phrased by the guiding research question, explained the local problem, provided a critical review of the literature addressing this problem, and discussed implications of the study. In Section 2, I discuss the methodology and the results of the study.

Section 2: The Methodology

In this study, I used a quantitative, correlational design to examine the relationship between mathematics scores on the i-Ready and on the EOCE for 10th graders at EHS. I examined how far test scores on the i-Ready predict test scores on the EOCE. The setting and sample subsection comprise the population, the sampling technique used, and the participants' eligibility criteria. In the instrumentation and materials subsection, I describe the i-Ready and EOCE data collection instruments, the validity and reliability of the instruments, the raw data source, and I explain the data used for measuring variables in the study. In Section 2, I also discuss the study's assumptions, limitations, delimitations, and measures taken for the protection of participants' rights.

Research Design and Approach

In this correlational study, I examined the relationship between i-Ready scores (predictor variable) and EOCE scores (criterion variable), while controlling for gender, ethnicity, and SES. A correlational design allowed the measurement of a relationship between variables, but it did not determine the cause of a phenomenon, that is, why a variable has specific values and by which other variables it is influenced (Triola, 2006). Since I sought to find a potential relationship between variables, I chose a correlational design as the most appropriate method (Creswell, 2012).

Setting and Sample

EHS enrolled approximately 2,000 economically and culturally diverse students, with 70% receiving free and reduced-price meals. The student ethnicity was as follows: 25% Hispanic, 18% Black, 23% White , 20% Asian, and 4% Native Hawaiian/Pacific

Islander. Seventeen percent of the population was categorized as ELL. In the 2014-2015 school year, females comprised 47% of students and males 53% (OSPI, 2016).

In this study, I used a nonrandom collection of the entire target population because the school district had already collected the data and made it available to me for the study. Creswell (2012) stated that to avoid confirmation bias, researchers should use the largest amount of data possible.

The target population consisted of archival records for students in the 10th grade (N = 512) who took the i-Ready and the EOCE in 2014-2015. The statistical power analysis software, G* Power 3.1, was used to calculate the minimum sample size for linear multiple regression by using a slope of .15, an alpha error probability of .05, and power set to .95, which indicated a minimum of N = 129 participants. Only students who were administered the i-Ready and the EOCE during the period of 2014-2015 were eligible to participate in the study. After removing the incomplete data, the actual sample consisted of N = 220 students, which exceeded the calculated minimum. I present demographic frequencies and percentages of the sample in Table 1.

Table 1

Demographic Frequencies and Percentages of the Sample

Distribution		N	%
Gender	Male	123	56
	Female	97	44
Ethnicity	Asian	40	18
	Black	65	29
	Hispanic	63	29
	Native Hawaiian/Other	8	4
	Two or More Races	11	5
	White	33	15
School Lunch Status	Free Lunch	144	66
	Reduced Lunch	15	7
	Not Eligible	61	28
Total Overall		220	100

Instrumentation

In this correlational study, I utilized archival i-Ready and EOCE data. The i-Ready was used to measure the independent (predictor) variable, while the EOCE was used to measure the dependent (criterion) variable. Eligible students in 10^{th} grade took the i-Ready in September, while all students were administered the EOCE in May. These assessments provided the data for the study.

The i-Ready

The i-Ready was the tool used in this study and was considered the predictor variable. The i-Ready is a computerized adaptive diagnostic (CAD) published in 2013, developed by the Smarter Balanced Assessment Consortium, and is used by many states to measure grade-level proficiency in the Common Core curriculum (SBAC, 2016). The i-Ready is a 54-72 item multiple choice exam. The test was developed to give students, parents, and teachers an accurate assessment of the grade level at which a student performs in reading and mathematics. Student scores aligned with a grade-level of performance. To complete the i-Ready, teachers administered the multiple choice exam on the laptop computer of each student during a testing session. Since the i-Ready is a CAD, comprised of thousands of questions, and is different for every student, there is no singular version of the test. Therefore, it is not attached to the appendix of this study.

For the validity of the i-Ready content, experts in education, reading, and mathematics reviewed and evaluated of each of the test items. Each test item was then considered for pilot testing and field testing, and then revised in adherence with the test construction principals prescribed by the AERA. The i-Ready was found to be reliable, with an r = .77-.85 across all subjects (SBAC, 2016). The i-Ready provided scores that are meant to categorize students into grade levels of algebraic thinking that range from the Kindergarten level the 12th grade level. Table 2 displays the scale scores for each grade-level in mathematics for the i-Ready.

Table 2

i-Ready Algebraic Thinking Placements for the Students in the 10th Grade

Grade level	Grade 10
Level K	0-415
Level 1	416-437
Level 2	438-452
Level 3	453-463
Level 4	464-472
Level 5	473-482
Level 6	483-492
Level 7	493-502
Level 8	503-514
Level 9	515-555
Level 10	556-610
Level 11	611-629
Level 12	630-800

The EOCE

Washington State has used the EOCE since 2011. In this study, I used these existing instruments with no modifications. Students must take a written version of the test during a testing session in May to complete the exam. Due to the testing policies administered by the local school regarding the EOCE, there were no additional copies of the test given without names and codes attached to specific students, and thus the school is not able to provide additional copies of the exam. Therefore, there is no copy of this test attached to the appendix of this study. Both the i-Ready and the EOCE have been labeled valid and reliable (OSPI, 2016). Creswell (2012) deemed the content validity of test questions as a representation of all possible test questions available. Criterion validity rests on the ability for the i-Ready and EOCE and to offer an adequate prediction. The scores of the i-Ready and EOCE have offered reliable prediction since their first administration (OSPI, 2016). The construct validity is determined by whether scores are significant and i-Ready and EOCE scores allow stakeholders to understand the population.

The Office of Superintendent of Public Instruction has implemented measures to ensure high content validity on the EOCE. Validity is a value of a test's strength to measure the content of a given construct through a collection of evidence (American Psychological Association, 2002). Washington educators, curriculum specialists, and grade-level administrators wrote line item specification for each subject assessed. A collective agreement was then attained to clarify learning standards through the development of benchmark indicators. Test makers then prepared the specifications of the test. The test item writers prepared test questions and scoring rubrics. Educational stakeholders reviewed the test items for bias and fairness. Once item writers approve a test question from the pilot test, a sample of students from across the state pilot the test items. Test developers perform two types of analyses for all items: classical item analysis, and IRT scaling using the Rasch (1960) model for multiple choice items.

After statistical analyses of pilot items have completed, data review committees evaluated item quality and appropriateness for inclusion in the general test item pool and candidacy for potential use. The tests items are then accepted or rejected upon re-evaluation of alignment to learning standards. Test developers then assessed a statistical review of item means, the IRT difficulties, and item-test correlations. Test makers removed items from the test item pool if items scored poorly. Then, test makers assembled the test items with alignment to state specifications. Refinements have been made annually since 2000.

The initially developed test underwent internal and external validation studies over the course of four years by cut score examination, formative assessments, student grade comparison, and test score correlation with other tests (SBAC, 2014). The test is considered a reliable measure of the items tested, with a Cronbach's alpha coefficient of .90. The standardized mathematics assessment (EOCE) provided scores categorized into achievement levels ranging from Level 1 to Level 4. Achievement levels represent specific scale scores that range from 200-675 and provided the bases for the criterion available. The EOCE achievement levels are as follows:

Level 1: Below Basic, or well-below standard. Represents little to no mastery of a subject's skills and knowledge.

Level 2: Basic, or below standard. Represents some mastery of a subject's skills and knowledge.

Level 3: Proficient, or meets standard. Represents mastery of a subject's skills and knowledge at grade level.

Level 4: Advanced, or exceeds standard. Represents advanced mastery of a subject's skills and knowledge.

Each EOCE achievement level associated with a specific range of achievement scores. Table 3 displays the scale scores for each level in mathematics on the EOCE.

Table 3

EOCE Mathematics Levels and Associated Scale Scores

EOCE Mathematics Level	Scale Scores Range
Level 1	200-374
Level 2	375-399
Level 3	400-442
Level 4	443-675

Variable Scaling

The EOCE scores measured the criterion variable. The i-Ready scores were measured as continuous. Gender was measured as categorical. Grade level and SES (determined via free and reduced lunch codes) were measured as categorical. Table 4 displays the criterion and predictor variables, measurement, and measurement scales.

Table 4

Displaying the Variables, Measurement, and Scale

Variable type	Variable	Measurement scale
Criterion	Math scores	Continuous
Predictor	i-Ready scores	Continuous
Predictor	Gender	Categorical
Predictor	Socioeconomic status	Categorical
Predictor	Ethnicity	Categorical

Gender was a categorical variable as 1 for female and 0 for male. SES was a categorical variable as determined as 1 for high SES (ineligible for free lunch), 2 for middle SES (receiving reduced lunch), and 3 for low SES (receiving free lunch). I was able to provide the raw data from both instruments by request.

Data Collection

In this study, I examined the relationship between a predictor variable, the i-Ready scores, and a criterion variable, the EOCE scores, while controlling for influences of gender, ethnicity, and SES. After receiving IRB approval, I retrieved archival data from the school district for the assessment period of 2014-2015. The district published EOCE and i-Ready data reports independently, but both reports displayed the required information for predictor and criterion variables. The data were accessible as developmental scale scores for each student. These scores gauged a student's level of knowledge growth over each EOCE administration from the elementary levels to high school. If a score increased, this could suggest an increase in a level of knowledge and subsequent performance increase on the EOCE. A letter of permission from the school district with permission to access the data was included in the appendix of this study (p. 117).

Data Analysis

How far the i-Ready scores predicted mathematics scores on the EOCE was the central research question for this study. Whether gender, ethnicity, and SES moderated the relationship between the test scores was the second research question for this study. To answer these research questions, I obtained archival data was from the school archive through the principal at EHS. I obtained the data from the archive via email after the Walden University IRB gave permission. I then imported the data into IBM SPSS Statistics software, version 24, for Mac OS. I assessed data normality using a Kolmogorov-Smirnov test. If the test was nonsignificant ($p > .05$), the data would be normally distributed and further statistical tests, such as regression analysis, would be applied. I would then perform descriptive statistical analysis including minimal maximal and mean values, and standard deviation.

To answer the research questions, I examined the relationship between the predictor variable (the i-Ready scores) and the criterion variable (the EOCE scores), under the moderating influences of gender, ethnicity, and SES. I used multiple linear regression analysis to calculate the beta weights for each predictor variable. Gender as a nominal variable was dummy-coded (1) for female and (0) for male. Ethnicity was a categorical variable and was coded as (0) for White, and (1) for others, i.e., Black, Asian, Pacific Islander, Hispanic, or mixed ethnicity. School lunch codes provided the basis for SES. SES as a categorical variable was dummy coded (0) for high SES (not eligible for free lunch), and (1) for low SES (receiving free or reduced lunch). To test the moderating effects, I added the products of the potentially moderating variables and the i-Ready scores to the multiple regression analysis as additional predictors.

Conducting regression analysis, I evaluated how well the i-Ready predicted EOCE scores, and if gender, ethnicity, and SES moderated any impact of i-Ready on EOCE scores. Using unstandardized (*b*) and standardized (ß) coefficients, I assessed the each predictor variable's moderating impact of the i-Ready on the EOCE. I used R^2 values to explain the amount of variance predicted in the criterion variable. Entering variables using simultaneous regression allowed SPSS to consider all variables simultaneously. Choosing a confidence level of 99% implied a statistically significant level of $p < 0.05$ to guide decisions in rejecting or failing to reject the null hypotheses.

Assumptions, Limitations, Scope and Delimitations

This study was based on several assumptions. The first assumption was that archival data was accurate. The second was that the EOCE mathematics item specifications and administrative procedures remained consistent throughout the study. The third is that the i-Ready administrative procedures remained consistent throughout the study.

A limitation of this quantitative study was that this urban and diverse population may not reflect the general population of American high school students. Another limitation was that certain variables, such as teacher quality, parental involvement, and school climate might exist as relevant factors for this study but were not available for measurement in the archival data.

The scope of this study delimited to 10^{th} grade students enrolled during the 2014-2015 school year in one high school with a total population of approximately 2,000 students. This study was delimited to the 2014-2015 sophomores who had completed both the i-Ready and the EOCE during their sophomore year.

Protection of Participants' Rights

To help ensure the ethical treatment of research participants, I adhered to U.S. Federal Government Department of Health and Human Services (2016) regulation 45 CFR § 7246.10. The purpose of this regulation was the protection of the participants by ensuring that the study was completely above reproach in all aspects to which the information used could extend its reach beyond normal academic boundaries, thereby possibly causing even the smallest bit of harm to the participants. Because this study is only using public archival data, I also followed this regulation by completely avoiding any action that subjected students to experimentation, or any manipulation of their normal daily life. I performed this study with the proper boundaries established by Walden University's Institutional Review Board (IRB). Before data collection occurred, I obtained approval from the Walden University's IRB. The IRB approval number for this study was 06-23-17-0409277.

For this study, I did not involve interaction with any participants, and data was not publicly made available. The school principal approved the data for use. In this study, I took steps to ensure that the data was de-identified before releasing it. Additionally, the data was safeguarded in a locked cabinet and on a password-protected computer to prevent any sensible individual lacking specific knowledge of the procedures to identify any particular element of the information used in this study.

Results

The descriptive statistics showed that students scored on average 484 points (*SD* = 34) at i-Ready, and 389 points (*SD* = 45) at the EOCE test. Table 5 displays demographic frequencies and percentages for each variable used in the study.

Table 5

Descriptive Statistics for the Criterion and Predictor Variables

Variable	*Min*	*Max*	*M*	*SD*
i-Ready	375	569	484.0	34.0
EOCE	245	560	389.0	45.0

Note. $N = 220$

In response to RQ1, the i-Ready scores significantly and positively predicted the EOCE test scores with $\beta = .64$ ($p < .001$), explaining $R^2 = .43$ of the criterion variance (Table 6). As a result of the findings, I rejected the null hypothesis. Figure 7 displays the linearity between the i-Ready and the EOCE.

Table 6

Multiple Regression Analysis Results for i-Ready, Gender, Ethnicity and SES

Model	ß	*p*	R^2
i-Ready	.64	.00	.43
i-Ready x gender	.03	.58	
i-Ready x ethnicity	-.08	.15	
i-Ready x SES	-.04	.48	

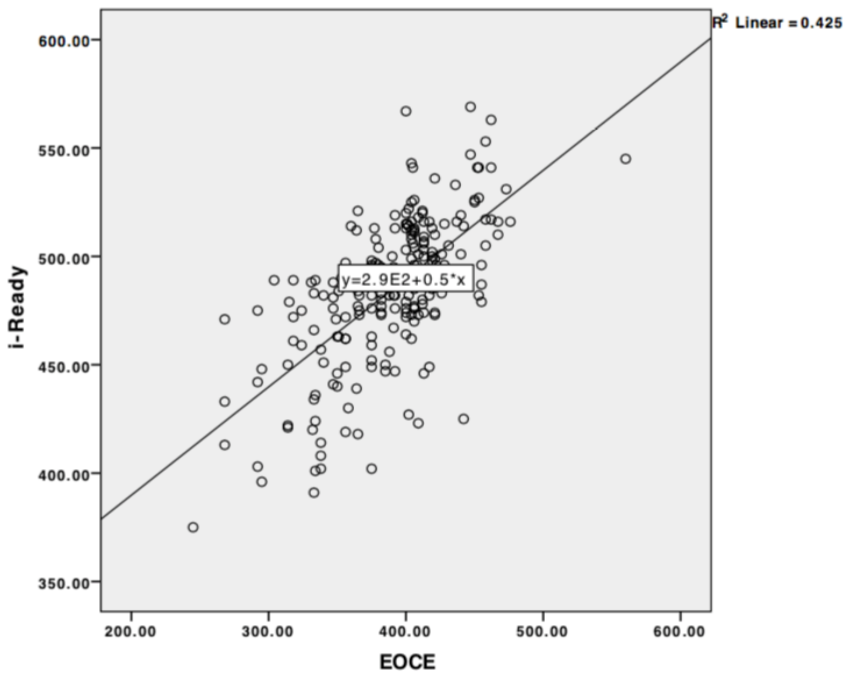

Figure 7. Scatterplot showing linearity between i-Ready and the EOCE scores.

In response to RQ2, the moderator analysis revealed that the demographic variables i-Ready x gender (β = .03, p = .58), i-Ready x ethnicity (β = -.08, p = .15) and i-Ready x SES (β = -.04, p = .48) did not predict the EOCE scores, meaning that the moderating effect was not significant. Thus, the findings confirmed the null hypothesis that suggested that gender, ethnicity, and SES did not moderate the relationship between the i-Ready and EOCE scores.

Discussion

In this study, I investigated the predictive power of CAD tools in mathematics education, and their effects on student achievement as moderated by gender, ethnicity, and SES. Using multiple regression analyses, I assessed the research questions to determine whether to reject or accept the null hypotheses. The next section is a discussion of the regression analysis findings.

i-Ready

I used multiple regression analysis to reveal that the i-Ready scores were a strong predictor of EOCE test scores. The findings supported Shapiro, Dennis, and Fu (2015), who observed that CADs could function as CBMs in in test prediction, but this study provides research for a CADs ability to predict proficiency at the high school level. From a broader perspective, the findings of this study supported the assertions of Piaget (1951), Vygotsky (1978), Ausubel (1960), Bloom (1956), and other cognitive researchers (Anderson, Spiro, & Anderson, 1978; Gagné, 1979; Rumelhart, 1980) who implied learning occurs from level to level and grows through experience and necessity.

Gender, Ethnicity, and Socioeconomic Status

I used multiple regression analysis to reveal that gender, ethnicity, and SES did not moderate the relationship between the i-Ready and EOCE scores. These findings are supported by the theories of Piaget (1959), who argued that knowledge will be different dependent on experience, but not necessarily on demographic traits. While demographic information did vary regarding years of study in mathematics, all of these students have had similar levels of mathematics experience. The results of this study suggested that the i-Ready would provide a fair assessment without bias by gender, ethnicity, or SES.

Conclusion

The purpose of this study was to determine whether the i-Ready, a CAD, could successfully predict proficiency on a high-stakes test, the EOCE, required for graduation. Previous studies have sought to determine if CADs and CBMs can be effective tools for test prediction (Kettler, 2011; Shapiro et al., 2015) but no researchers have conducted studies at the high school level involving the i-Ready and the EOCE. With this study I intended to fill this gap in research. In this study, I provided data to show a statistically significant correlation exists among scores in the i-Ready and EOCE, while gender, ethnicity, and SES did not moderate the relationship between the i-Ready and EOCE scores. In accordance with the reviewed theories, these data supported the notion that students lacking the level of knowledge needed to perform well on the i-Ready will also not do well on the EOCE.

The i-Ready scores were made to be instantly accessible to school administration. Unlike written tests which can still take a substantial amount of time for processing, the data from the i-Ready would come back as soon as the students complete the exam in the fall (SBAC, 2016). With the results of this study, I identified 10th grade students who may be in danger of failing the EOCE, which educators could use to provide additional academic support as an effort to prevent them from failing the EOCE (Brookhart, 2017;

Gibbs & Simpson, 2005). In Section 3, I discuss the project as a product deliverable from the results of this study.

Section 3: The Project

Introduction

As schools and districts continue to become more socially and economically diverse, the need has grown to find methods to help students identify the danger of failing high-stakes tests, which are required to graduate (Ravitch, 2016). With the findings of this study, I demonstrated a positive relationship between i-Ready and EOCE scores, and explained a significant variance in the EOCE scores. Gender, ethnicity, and SES had no significant moderating influence on prediction. CAD tools, such as the i-Ready, offered an accurate measurement and prediction of EOCE scores. Based on the results of this study, a policy recommendation in the form of a position paper was the most appropriate genre to share data and recommendations with the administrative team, based on the results and their implications (Stelzner, 2007). A policy paper allowed results and recommendations to be available to the school administration and could lead to a revision of its policies regarding preventative measures for students in danger of failing high-stakes tests. With the results of this study, I provided insight into the i-Ready's ability to account for a substantial variance in test scores, regardless of gender, ethnicity, or SES. Three recommendations are presented in the position paper: (1) collect and analyze i-Ready data (2) provide additional resources for students in danger of failing the EOCE based on the i-Ready data and (3) use learning analytics tools to gain empirical insights into teaching and learning. In the position paper, I summarized the results of this study and provided recommendations for high-stakes test failure prevention by using the i-Ready.

Rationale

The audience for this position paper was the administration responsible for assessment, curriculum, and instruction at EHS. The problem was whether the i-Ready predicted high-stakes test performance for 10^{th} graders at one high school in the Northwestern part of the United States. Since the i-Ready positively predicted the EOCE scores, this position paper was grounded in use of the i-Ready to help school administrators implement plans to predict high-stakes tests. A position paper directly addressed the problem in the study, the results of the study, and the immediate implication for the local site (Gordon & Graham, 2003). The position paper will provide the administrative team with the results of the study in hand, along with a list of recommendations for the use of the i-Ready with the EOCE. To be effective, administrators would have to make a change in school policy. A position paper was the most suitable method to enact appropriate social change to the school's administration regarding this study (Powell, 2012; Seltzner, 2010).

The following section, a literature review, relates the themes in the position paper.

Review of the Literature

The literature review covered policy recommendations/position papers, formative feedback, and learning analytics. The search was conducted to locate peer-reviewed scholarly articles that were within five years of the study completion date. Some foundational studies outside of the five-year period were also used to establish a rationale for position papers, as well as provide a connection to more recent studies. The following databases used were: Academic Research Complete, ERIC, and Education Source. The search terms used were: *position papers, educational policy, policy papers, formative feedback,* and *learning analytics*.

Policy Recommendations and Position Papers

A policy paper (also known as a white paper), began as an official government publication that makes a strong argument and resolution to an issue (Stelzner, 2010; Taylor & Bragg, 2015). Position papers are founded in empirical fact, grounded in both opinion and strong logic, and contain information useful to the reader (Gordon & Graham, 2003). In the simplest of terms, position papers are documents that describe a problem and then provide a compelling solution (Kantor, 2009). In education, the specific purpose of a position paper may vary depending on the author and its audience. Position papers are often used for sales of products and goods, or as marketing tools, while at other times they are used to influence thought non-commercially (Stelzner, 2007).

Formats and themes vary for position papers as a reflection of their specific target audience. However, some common themes have arisen in the modern era (Powell, 2012). Position papers commonly comprise a cover page, executive summary, a description of the problem, a targeted solution to that problem, a conclusion with an argument to implement the solution, and a reference list (Kantor, 2009). Modern position papers are described as a mixture of a high-concept advertisement blended with a fact-based report that is immediately accessible to the reader, uses appropriate charts and graphs, and sometimes uses the integration of color, designs, bullets, headers and footers, and pictures (Sakamuro, Stolley, & Hyde, 2010).

Formative Assessment

The second part of this literature review presents evidence from the recent literature that focuses on the best practices and applications of formative assessment. In education, the documentation and measurement of empirical data of student knowledge is defined as assessment (Shute, Leighton, Jang, & Chu, 2016). Formative assessment is a measurement of learning that is usually not graded and provides useful feedback to the students and teachers regarding content area knowledge (Sato & Atkin, 2007). Formative assessment usually occurs before summative assessment (Dixson & Worrell, 2016). Formative assessment can come in many forms, ranging from oral question and answers to written work, tests and diagnostics (Vie, Popineau, Bruillard, & Bourda, 2017). Many scholars posit that the natural result of formative assessment is feedback to aid the learner in constructing knowledge (Evans, 2013; Nicol & Macfarlane-Dick, 2006). One of the primary roles of formative assessment is to deliver meaningful feedback to learners within the classroom to make a more effective learning environment (Clarke, 2001). Another critical aspect of formative assessment is delivering feedback for the learner before a summative assessment (Starkman, 2006).

Furthermore, another challenge for formative assessments is the amount of time and resources it takes from other classroom activities (Fonseca et al., 2015). It is often the case that formative feedback can take a substantial amount of time, effort, and resources which take away from classroom instruction that is normally focused on instruction and summative assessments (Irving, 2007). Recently, research has been done to improve the timeliness, reliability, and validity or formative assessments by utilizing modern technology (Owen, 2016).

In addition to making evaluation more efficient and impartial, computerized assessments generate data that computers can collect and analyze. As new computer-based technologies have changed the overall educational landscape, researchers have begun studying how computer-based instruction and technology can integrate with formative feedback (Floratos, Guasch, & Espasa, 2015; Gikandi & Morrow, 2016). Using foundational feedback research, scholars conducted studies to identify ways to make feedback meaningful and attainable when using CADs (Timmers, Braber-van den Broek, & van den Berg, 2013). Bangert-Drowns, Kulik, Kulik, and Morgan (1991) found that feedback based on one CBM positively impacted the level of effort students were willing to do to improve upon prior knowledge. With more and more schools using computer-based education, as well as implementing CADs into schools, new ways of collecting data to improve education have been made possible (Reigeluth, 2016).

Learning Analytics

With the use of CADs in the classroom, schools could begin to bring more learning analytics (LA) instruments in public schools for formative assessment. In the past, traditional educational structure was to provide a level of knowledge benchmark in the form of a high-stakes test, and then to present the educational content in a traditional educational approach which catered to the median student, with nothing other than hope that as many students as possible will understand the material (Tempelaar, Heck, Cuypers, van der Kooij, & van de Vrie, 2013). Online diagnostics of LA have emerged as a means to collect, analyze, and synthesize data for the improvement of student learning experiences (Adejo & Connolly, 2017). Data-rich technologies such as some CAD applications allow a much greater understanding for both the teacher and student. Educators can better understand precisely where students are in their levels of knowledge through online formative assessment, regardless of the educational setting, which is one of the critical functions of LA (Baker & Inventado, 2014). This knowledge allows educators to better focus on the specific, empirical challenges any student may have without having to guess as to whether they meet the appropriate level of knowledge (Arastoopour et al., 2014).

Since the i-Ready tracks all student work through an online system, it is possible to go back and examine all the work that students have done in a formative manner and see what level a student is performing at, and whether their knowledge changes over time. Before these online systems, all teachers and educators had for a record of learning was the grade they put into the grade book (Knight, Shum, & Littleton, 2013). CADs can be very well integrated into the LA systems and have the potential of a recommendable LA instrument (Thompson, 2017).

The further adoption of LA instruments in public schools could create a monumental shift in the educational landscape for formative assessment (Deakin Crick, 2017). Instead of a feedback loop that occurs in the traditional sense of test, score and result, feedback could be given in real-time and shift towards formative assessment becoming monitoring function (Knight, Buckingham Shum, & Littleton, 2014; Young & Muller, 2017). The more teachers and students adopt technology like CADs, the more aware they will become of their level of knowledge at all times. In this way, LA can generate data that establishes an evidence-basis for school, rather than teaching purely on theory and speculation (Lonn, McKay, & Teasley, 2017).

The question remained as to why more had not been done to integrate LA technology into the classrooms. Formative assessment through LA has faced many challenges in being integrated into the public school system due to ethical concerns, such as student privacy and the control of data for children (Ferguson et al., 2015). However, at the university level, LA has been used as a profile analysis tool to predict student dropout rates, and behavior tracking tools monitoring of student user data in virtual online groups (Nistor et al., 2014; Nistor, Derntl, & Klamma, 2015). Slowly, K-12 public schools have begun adopting the use of LA as continuing research demonstrates the potential impact LA can make to education through predictive modeling, personalization, and automated guidance systems (Conde & Hernández-García, 2015; Daniel, 2015; Nistor et al., 2015). With educational data mining (Kop, 2012) and learning analytics (Chatti, Dyckhoff, Schroeder, & Thüs, 2012), research has continued to find ways LA can enhance the classroom for all students. CADs are one such platform that has the potential to bring LA further into public schools (Deakin Crick, 2017; Macfadyen, Dawson, Pardo, & Gašević, 2014).

School Adoption of Learning Analytics Tools

The impact of both LA and CADs have been shown to increase the value of student learning and instruction (Daniel, 2015; Siemens, Dawson, & Lynch, 2013). The ability for schools to adopt LA tools such as some CADs, for the use of educational improvement practices, such as test prediction have also been suggested (Baker & Invitado, 2014). Researchers have identified the fields that inhibit the use of LA tools in schools into the following: technological readiness (Arnold et al., 2014), leadership (Arnold, Lonn, & Pistilli, 2014), organizational culture (Carbonell, Dailey-Hebert, & Gijselaers, 2013), and staff capacity (Norris & Baer, 2013).

While obstacles to LA tools like CADs use may still exist, many high schools have been able to address many of these barriers, such as finally achieving a level of technological readiness that would facilitate the use of LA and CADs (Roberts-Mahoney, Means, & Garrison, 2016). Pedagogical approaches have addressed many of the issues to support student use of LA tools (Wise, 2014). The most prominent challenge remains both the elements of leadership and school culture in adopting the use of LA (Kennedy et al., 2014). Therefore, the final recommendation of this position paper will be an emphasis of literature that supports the use of LA tools like the i-Ready for the data-rich benefits such tools provide, to move schools towards a more empiricism-based approach to education (Brasiel et al., 2016; Lockyer, Heathcote, & Dawson, 2013; Siemens, 2013; Tempelaar, Rienties, & Giesbers, 2015).

Project Description

In the position paper, I recommended strategies to design and implement the use of the i-Ready as a tool for prediction of the EOCE and made recommendations for the implementation of the i-Ready as predictive tool for the EOCE at EHS. The needed resources to present the project would be a meeting with educational stakeholders at EHS. This meeting would include the building principal, assistant principals as well as the testing coordinator for the high school. EHS has many existing supports to aid in the implication of the project because they already take both the i-Ready and the EOCE. These existing supports include testing space, the time allotted to conduct the testing, and test materials. Potential barriers to the project could include the unwillingness to use the i-Ready as a prediction tool. However, the evidence provided in this study that demonstrates the i-Ready's ability to predict EOCE scores will be a potential solution to this barrier. The position paper will encourage the use of LA tools, which include the i-Ready, to help EHS discover ways to use data to gain a better understanding of each student. While formative feedback is one application of LA, the position paper will recommend that online dashboards, student risk indexes, and CADs be used to better personalize educational experiences to each student, and to create a more efficient formative feedback cycle. The position paper will also explain how the use of LA can also track student work and progress of their levels of knowledge, which is now available to for schools to use at the high school level.

The proposal for implementation of this project recommended that schools and administrators and the testing coordinator use the i-Ready in the Fall of 2018 to predict which students are in danger of failing the EOCE and provide them additional support. Then, students would take the EOCE in the Spring of 2019. When the school receives the EOCE scores, an assessment will be done to see whether the EOCE scores improved as a result of the i-Ready.

The roles and responsibilities of the project would rely on the school administration and the testing coordinator at EHS regarding the project and implementation. In the position paper, I demonstrated, via this project study, the method I used to conduct a predictive analysis. From this study, the administration and testing coordinator should be able to replicate the analysis. The students involved in the study play no formal role, but EHS may provide some students with additional testing support if they are found to be in danger of failing the EOCE. I included more information regarding the evaluation of the effectiveness of the project in the next section.

Project Evaluation Plan

The evaluation for this position paper will be a standard summative outcome-based evaluation to determine if there was positive impact based on test score improvement resulting from test prediction and intervention provided to students. Since the purpose of the position paper was to encourage EHS to use the i-Ready to predict performance on the EOCE, statistical analysis would be used to determine the effect of test intervention from the i-Ready would best provide empirical evidence of its effectiveness (Creswell, 2012). The overall evaluation goal would be to determine if the i-Ready prediction and intervention could improve student test scores. Therefore, test scores of the EOCE will be the outcome measure of this study. The key stakeholders in this evaluation will be the school principal, administration, and the testing coordinator at EHS.

The project evaluation could begin in the 2018-2019 year. The entire 10th grade class at EHS and the entire 10th grade class at a nearby high school where intervention is not available as part of their curriculum could comprise the population of this study. Both high schools could then take i-Ready as their pre-test in the fall of 2018. The experimental group (EHS) would receive intervention for students in danger of failing the EOCE. The control group (neighboring high school) would receive no intervention for those students in danger of failing the EOCE because it is not part of their curriculum. In the spring of 2019, all students would take the EOCE as their post-test. ANOVA analysis could be used to determine if there was a significant difference between the experimental and control groups. If there is a significant improvement in the experimental group in comparison to the control group, then the intervention will be determined to have a positive impact.

Project Implications

In the position paper, I explained the lack of prediction methods used for the high-stakes test at EHS. I provided recommendations to fix this using the i-Ready CAD. I also demonstrated ways CADs could function as a formative assessment to further the use of LA tools in the school system.

For administrators, these recommendations could be used to help aid students in danger of failing high-stakes exams like the EOCE before they take them, thereby lowering the failure rate, reducing dropouts, and narrowing the achievement gap. The use of CADs may also lead to an increase in overall academic monitoring or real-time formative assessment that will allow teachers, students, and administrators a more precise representation of the level of knowledge students possess. On a larger scale, the use of CADs and other LA related instruments in the classroom could open the door to larger LA platforms in schools. LA in schools would provide data that allows for a more accurate assessment of student learning to approach education and learning from an empirical standpoint, rather than a philosophical pedagogy. In Section 4, I discuss the reflections and conclusions of the project study.

Section 4: Reflections and Conclusions

In this section, I share my reflections and conclusions about the creation of the project study. This section includes project strengths and limitations, alternative approaches, scholarship, implications, and suggestions for future research. I conclude with an analysis of my work and what was learned.

Project Strengths

The strength in the use of a position paper for this project was that it was aligned with the administration of EHS, the target audience. Additionally, the project design of a position paper allowed me to address the specific problems the school in this study faced, present evidence and data, as well as recommendations for the administration based on the findings. Since the administration was ultimately the team that would be able to make changes based on the findings of this study, the position paper was the most reliable choice to reach them in a scholarly and convincing manner. A position paper could target my intended audience, include both evidence of my research and peer-reviewed literature that supports the study, best practices, and possible outcomes of my findings to help accurately solve the problem at the school of study.

The position paper allowed me to present literature supporting the use of CADs within the school and to suggest the use of LA as a form of feedback monitoring for students. LA could present some solutions, but one of the problems facing schools is that they are slow to trust in emerging LA technologies even, if they are already in place in the schools, as educators seek to navigate the various ways LA might change the role of feedback in schools (Deakin Crick, 2017).

I based my project on cognitive theories that frame learning in levels of knowledge. The target audience of the position paper was often concerned with students meeting certain levels of knowledge in both standardized assessments and CADs, and this framework allowed educators to better understand the benefits of this study without having to apply a framework that did not readily relate to their school mission. Having a clear purpose for the project, in alignment with the framework, helped guide the recommendations towards a more precise path of adoption by the administration.

Project Limitations

Since the position paper was meant to be a convincing argument for the use of CADs to predict high-stakes tests, it was reliant on the school administration to implement such changes outlined in the position paper. If the administration deemed the project to be outside of their time and resources, the position paper would continue to be a strong argument with empirical data demonstrating the effectiveness of CADs; however, it would be limited in its ability to improve student scores in the school where the original study took place. Another limitation of this project paper was that, like the original study, the arguments in this position paper may have reduced its generalizability. However, the position paper of this study could be informative for other research in the field of CADs and high-stakes tests, and use of LA, particularly at the high school level. With the position paper, I hoped to inspire future research that would add to the body of literature on CADs, LA and high-stakes tests, thereby broadening the scope of data related to high-stakes test prediction.

Recommendations for Alternative Approaches

Researchers could address the problem of low achievement rates among 10th grade students on the EOCE by researching the SES barriers that impact student success. Another approach would be identifying specific elements that create the achievement gap such as teacher perceptions or early childhood education structures (Pianta, Downer, & Hamre, 2016). New research has suggested students with culturally diverse backgrounds perform better academically when taught by teachers with similar backgrounds to their own (Egalite, Kisida, & Winters, 2015). Researchers could explore trying to establish a quantitative connection to this phenomenon.

Scholarship, Project Development and Evaluation, and Leadership and Change

Creating the project has allowed me to reflect on my growth as a scholar. The doctoral process has been the most challenging endeavor I have ever faced, as it forced me to grow in ways that no other educational experience has. The project asked me to look at education through the eye of a scholar and a practitioner for social change. I began my journey questioning why the education system struggled so often to produce better results. These questions evolved to asking how technology could help improve education. Then, after many trials and errors, I came to the research questions presented in this project. Finally, I submitted new research and suggested opportunities for the continuation of the research presented in this study. Through this process, I have gone from wondering about how the education system works from a broad scope, synthesized the literature and have gained a very close understanding of themes that I reviewed for this project, and it has helped me to comprehend the challenges I face as a practitioner.

This process has also humbled me and taught me that as a scholar, the search for knowledge is never finished. Since I began this project as a practitioner, the research I have done for this project has helped me to understand the broader world of educational research and its impact on education.

Reflection on Importance of the Work

There is a need for the students and educators to have access to formative assessment data to better give transparency, equity, and progress. Every student, when faced with a high-stakes exam, should have the ability to be aware of what they currently know in relation to what they must know. Teachers and educators should also have this data, so that they may provide the appropriate remediation. Parents, administrators, and educational stakeholders should be able to provide accurate support and accountability based on empirical facts about student knowledge.

This work has helped me as a writer, scholar, teacher, and practitioner for social change. I have developed my ability to think critically, and discovered a new passion for research and inquiry. Rather than taking things at face value, I now often question the validity of claims or statements presented as fact. I have become more aware of the vast amounts of learning I could continue to pursue, and have learned to accept criticism more readily. I have learned that research is an iterative process, and that becoming a good scholar can require many attempts before it is ever achieved.

Implications, Applications, and Directions for Future Research

While technology has developed, education has struggled to keep up. Currently, there are systems in place that can generate incredible amounts of data for educators. In addition to this, analytic systems can create profound meaning for this data. However, as an academic institution, public education struggles in understanding, navigating, and harnessing data as an educational tool. I was able to successfully predict the performance of 10^{th} grade students on one high-stakes exam through the use of a CAD. As research in this utilization of adaptive diagnostic tools increases, hopefully so will our understanding of the potential applications of technology in the classroom. Technology is not only an instructional tool, but it could eventually shift how educators practice formative assessment. Rather than taking assessments at metered intervals, students, parents, and educators could eventually be able to have real-time monitoring of their learning as well as systems in place to help students in danger failing, while continuing to challenge the students who excel. It is my opinion that this use of data is inevitable, but the question remains as to who will end up in control of it. With scholarly research and inquiry, we can give this data meaning to the people for whom it matters most: our students.

The use of technology for formative assessment could help education become more efficient and equitable. Currently, our most underserved and under-assessed students are lower income and minority students. With students having access to the same diagnostic and prevention measures as students in areas with high SES, we may finally begin to bridge the academic achievement gap and provide appropriate levels of education and remediation to all students regardless of their background.

Conclusion

Despite the vast wealth of technology that exists in education, work must be done to help make it more useful for the learner. The massive amount of educational technology and the data it generates can be overwhelming to school administrations, teachers, and students. Our goal should be to harness educational data to make it easy, intuitive and useful for the learner. While there is no easy solution to this task, in this study, I presented one way to help students in danger of failing high-stakes mathematics tests. If we can use tools that already exist to help make learners more aware of what they know in relation to high-stakes tests, it could significantly benefit high school students who need to pass these tests to graduate.

The use of LA and data tracking will continue to impact public school systems, despite some schools' reluctance to utilize it. Even without software that schools purchase to help track student data, businesses in the private sector, both small and large, make use of learner data for their purposes. While public schools might be trying to use traditional testing as means of measuring student knowledge, they should instead embrace the use of learner data, and make efforts to keep the power of LA-based tools in the hands of the teachers, students, and educational stakeholders. A practical first step would be to use the information at their disposal to help students in danger of failing required high-stakes tests. Beyond this immediate and urgent need, the possibilities of online formative assessment systems, data tracking, and LA are seemingly endless.

References

Adejo, O., & Connolly, T. (2017). Learning analytics in higher education development: A roadmap. *Journal of Education and Practice, 8*(15), 156-163.

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, *57*(12), 1060-1073.

Anderson, R. C. (1984). Some reflections on the acquisition of knowledge. *Educational Researcher*, *13*(9), 5-10.

Anderson, R. C., Spiro, R. J., & Anderson, M. C. (1978). Schemata as scaffolding for the representation of information in connected discourse. *American Educational Research Journal*, *15*(3), 433-440.

Arastoopour, G., Buckingham Shum, S., Collier, W., Kirschner, P., Knight, S. K., Shaffer, D. W., & Wise, A. F. (2014, 23–27 June). Analytics for learning and becoming in practice. In J. L. Polman et al. (Eds.), *Learning and becoming in practice: Proceedings of the International Conference of the Learning Sciences* (ICLS ’14), Boulder, CO (Vol. 3, pp. 1680–1683). Boulder, CO: International Society of the Learning Sciences.

Arnold, K. E., Lonn, S., & Pistilli, M. D. (2014). An exercise in institutional reflection: The learning analytics readiness instrument (LARI). *Proceedings of the 4th International Conference on Learning Analytics and Knowledge* (LAK ’14), 24-28 March 2014, Indianapolis, IN (pp. 163-167). New York, NY: Association for Computing Machinery.

Arnold, K. E., Lynch, G., Huston, D., Wong, L., Jorn, L., & Olsen, C. W. (2014, March).

Building institutional capacities and competencies for systemic learning analytics initiatives. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge* (LAK ’14), 24-28 March 2014, Indianapolis, IN (pp. 257-260). New York, NY: Association for Computing Machinery.

Au, W. (2011). Teaching under the new Taylorism: High-stakes testing and the standardization of the 21st century curriculum. *Journal of Curriculum Studies*, *43*(1), 25-45.

Au, W. (2013). Hiding behind high-stakes testing: Meritocracy, objectivity and inequality in US education. *The International Education Journal: Comparative Perspectives*, *12*(2), 7-19.

Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, *51*(5), 267-272.

Ausubel, D. P., & Fitzgerald, D. (1962). Organizer, general background, and antecedent learning variables in sequential verbal learning. *Journal of Educational Psychology*, *53*(6), 243-249.

Backer, D. I., & Lewis, T. E. (2015). Retaking the test. *Educational Studies*, *51*(3), 193-208.

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), *Learning analytics* (pp. 61-75). New York, NY: Springer.

Ball, C. R., & Christ, T. J. (2012). Supporting valid decision making: Uses and misuses of assessment data within the context of RTI. *Psychology in the Schools*, *49*(3), 231-244.

Bangert-Drowns, R. L., Kulik, C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213-238.

Barrat, V. X., Berliner, B., & Fong, A. B. (2012). When dropping out is not a permanent high school outcome: Student characteristics, motivations, and reenrollment challenges. *Journal of Education for Students Placed at Risk (JESPAR)*, *17*(4), 217-233.

Bell, S. M., Taylor, E. P., McCallum, R. S., Coles, J. T., & Hays, E. (2015). Comparing prospective twice-exceptional students with high-performing peers on high-stakes tests of achievement. *Journal for the Education of the Gifted*, *38*(3), 294-317.

Berliner, D. (2011). Rational responses to high stakes testing: The case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*, *41*(3), 287-302.

Berliner, D. (2013). Effects of inequality and poverty vs. teachers and schooling on America's youth. *Teachers College Record*, *115*(12), 1-26.

Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. *US Department of Education, Office of Educational Technology, 1*, 1-57.

Birdwell, S. D. (2012). School leadership: Lessons from the lived experiences of urban teachers. *Journal of Ethnographic & Qualitative Research*, *7*(2), 52-63.

Bloom, B. S. (1956). *Taxonomy of educational objectives, handbook I: The cognitive* *domain*. New York, NY: McKay.

Brasiel, S., Jeong, S., Ames, C., Lawanto, K., Yuan, M., & Martin, T. (2016). Effects of educational technology on mathematics achievement for K-12 students in Utah. *Journal of Online Learning Research, 2*(3), 205-226.

Brennan, R. L. (2015). Testing for accountability: A balancing act that challenges current testing practices and theories. *Measurement: Interdisciplinary Research & Perspectives*, *13*(1), 35-38.

Brookhart, S. M. (2017). *How to provide effective feedback to your students*. Alexandria, VA: Association for Supervision and Curriculum Development.

Campbell, H., Espin, C. A., & McMaster, K. (2013). The technical adequacy of curriculum-based writing measures with English learners. *Reading and Writing*, *26*(3), 431-452.

Carbonell, K. B., Dailey-Hebert, A., & Gijselaers, W. (2013). Unleashing the creative potential of faculty to create blended learning. *The Internet and Higher Education, 18*, 29-37. doi:10.1016/j.iheduc.2012.10.004

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology-Enhanced Learning, 4*(5), 318-331.

Cheng, Y. (2010). Improving cognitive diagnostic computerized adaptive testing by balancing attribute coverage: The modified maximum global discrimination index method. *Educational and Psychological Measurement, 70*(6), 902-913.

Clarke, S. (2001). *Unlocking formative assessment*. London, UK: Hodder & Stoughton.

Codding, R. S., Mercer, S., Connell, J., Fiorello, C., & Kleinert, W. (2016). Mapping the relationships among basic facts, concepts and application, and common core curriculum-based mathematics measures. *School Psychology Review*, *45*(1), 19-38.

Codding, R. S., Petscher, Y., & Truckenmiller, A. (2015). CBM reading, mathematics, and written expression at the secondary level: Examining latent composite relations among indices and unique predictions with a state achievement test. *Journal of Educational Psychology*, *107*(2), 437-450.

Common Core State Standards Initiative. (2010). *Common core state standards for mathematics.* Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers. Retrieved from http://www.corestandards.org

Conde, M. Á., & Hernández-García, Á. (2015). Learning analytics for educational decision making. *Computers in Human Behavior, 47*, 1-3.

Conley, D. (2015). A new era for educational assessment. *Education Policy Analysis Archives*, *23*, 1-41.

Cornelius, D. (2011). The education and skills gap: A global crisis. *Techniques: Connecting Education and Careers*, *86*(4), 50-55.

Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral reading rate to predict student performance on statewide achievement tests. *Educational Assessment*, *7*(4), 303-323.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating* *quantitative and qualitative research* (Laureate Education ed.). Boston, MA: Pearson Education.

Daniel, B. (2015). Big data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology, 46*(5), 904-920.

Daun-Barnett, N., & John, E. P. S. (2012). Constrained curriculum in high schools: The changing math standards and student achievement, high school graduation and college continuation. *Education Policy Analysis Archives, 20*(5), 1-24.

Davis, S. D., & Chan, J. C. (2015). Studying on borrowed time: How does testing impair new learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1-56.

Deakin Crick, R. (2017). Learning analytics: Layers, loops, and processes in a virtual learning infrastructure. In C. Lang, G. Siemens, A. Wise, & D. Gašević (Eds.), *Handbook on Learning Analytics* (pp. 291-308). Society for Learning Analytics Research. doi:10.18608/hla17.025

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*(3), 219-232.

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurements to establish growth standards for students with learning disabilities. *School Psychology Review, 30*(4), 507-524.

Dessoff, A. (2012). Education reformer: Robert J. Marzano. Models for education reform, part 1. *District Administration*, *48*(3), 78-84.

Dixon-Román, E. J., Everson, H. T., & McArdle, J. J. (2013). Race, poverty and SAT scores: Modeling the influences of family income on black and white high school students' SAT performance. *Teachers College Record*, *115*(4), 1-19.

Dixson, D. D., & Worrell, F. C. (2016). Formative and summative assessment in the classroom. *Theory Into Practice*, *55*(2), 153-159.

Duan-Barnett, N., St. John, P. E. (2012). Constrained curriculum in high schools: The changing math standards and student achievement, high school graduation and college continuation. *Education Policy Analysis Archives*, 20(5), 1-25.

Dymoke, S. (2012). Opportunities or constraints? Where is the space for culturally responsive poetry teaching within high-stakes testing regimes at 16+ in Aotearoa New Zealand and England? *English Teaching*, *11*(4), 19-35.

Egalite, A. J., Kisida, B., & Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, *45*, 44-52.

Erskine, J. L. (2014). It changes how teachers teach: How testing is corrupting our classrooms and student learning. *Multicultural Education*, *21*(2), 38-40.

Espin, C. A. (1999). Criterion-related validity of curriculum-based measures in writing for secondary school students. *Reading & Writing Quarterly*, *15*(1), 5-27.

Espin, C. A., Busch, T. W., Lembke, E. S., Hampton, D. D., Seo, K., & Zukowski, B. A. (2013). Curriculum-based measurement in science learning: Vocabulary-matching as an indicator of performance and progress. *Assessment for Effective Intervention*, *38*(4), 203-213.

Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of*

Educational Research, *83*(1), 70-120.

Every Student Succeeds Act, S.1117, 114th Cong. (2015). Retrieved from https://www.congress.gov/bill/114th-congress/senate-bill/1177/text

Ferguson, R., Macfadyen, L., Clow, D., Tynan, B., Alexander, S., & Dawson, S. (2015). Setting learning analytics in context: Overcoming the barriers to large-scale adoption. *Journal of Learning Analytics*, *1*(3), 120-144.

Ferguson, R., & Shum, S. B. (2012). Towards a social learning space for open educational resources. *Collaborative Learning*, *2*, 309-327.

Finn, B., & Roediger III, H. L. (2013). Interfering effects of retrieval in learning new information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(6), 1665-1681.

Floratos, N., Guasch, T., & Espasa, A. (2015). Recommendations on formative assessment and feedback practices for stronger engagement in MOOCs. *Open Praxis*, *7*(2), 141-152.

Fonseca, J., Carvalho, C., Conboy, J., Valente, M. O., Gama, A. P., Salema, M. H, & Fiúza, E. (2015). Changing teachers' feedback practices: A workshop challenge. *Australian Journal of Teacher Education*, *40*(8), 59-82.

Ford, D. Y., & Helms, J. E. (2012). Overview and introduction: Testing and assessing African Americans: "Unbiased" tests are still unfair. *The Journal of Negro Education*, *81*(3), 186-189.

Freire, P. (1970). *Pedagogy of the oppressed*. New York, NY: Seabury Press.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it?. *Reading Research Quarterly*, *41*(1), 93-99.

Fuchs, L. S., & Deno, S. (1981). *The relationship between curriculum-based mastery measures and standardized achievement tests in reading* (Research Report No. 57). Minneapolis, MN: University of Minnesota, Minneapolis Institute for Research on Learning Disabilities.

Fuchs, L. S., & Fuchs, D. (1992). Identifying a measure for monitoring student reading progress. *School Psychology Review*, *21*, 45-58.

Fuchs, L. S., & Fuchs, D. (2007). A model for implementing responsiveness to intervention. *Teaching Exceptional Children*, *39*(5), 14-20.

Fuchs, L. S., Hamlett, D. F. C. L., & Stecker, P. M. (1991). Effects of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research Journal*, *28*(3), 617-641.

Gagné, R. M. (1979). *The conditions of learning.* New York, NY: Holt, Rinehart and Winston.

Gareis, C. R., & Grant, L. W. (2015). *Teacher-made assessments: How to connect curriculum, instruction, and student learning.* Larchmont, NY: Routledge.

Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education, 1*, 3-31.

Gikandi, J. W., & Morrow, D. (2016). Designing and implementing peer formative feedback within online learning environments. *Technology, Pedagogy and Education*, *25*(2), 153-170.

Good, R. H. III, Simmons, D. C., & Kameenui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high stakes outcomes. *Scientific Studies of Reading*, *5*, 257–288.

Gordon, M., & Graham, G. (2003). *The art of the white paper.* Retrieved from http://www.gordonandGordon.com/downloads.html.

Gray, W. M. (1978). A comparison of Piagetian theory and criterion-referenced measurement. *Review of Educational Research*, *48*(2), 223-249.

Harland, T., McLean, A., Wass, R., Miller, E., & Sim, K. N. (2015). An assessment arms race and its fallout: High-stakes grading and the case for slow scholarship. *Assessment & Evaluation in Higher Education*, *40*(4), 528-541.

Hauerwas, L. B., Brown, R., & Scott, A. N. (2013). Specific learning disability and response to intervention: State-level guidance. *Exceptional Children*, *80*(1), 101-120.

Hensley, K., Rankin, A., & Hosp, J. (2017). Comparing student performance on paper- and computer-based math curriculum-based measures. *Assistive Technology*, *29*(3), 140-145.

Hintze, J. M., & Silberglitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review*, *34*(3), 372-386.

Hosp, J. L., Hensley, K., Huddle, S. M., & Ford, J. W. (2014). Using curriculum-based measures with postsecondary students with intellectual and developmental disabilities. *Remedial and Special Education*, *35*(4), 247-257.

Hosp, M. K., & Hosp, J. L. (2003). Curriculum-based measurement for reading, spelling, and math: How to do it and why. *Preventing School Failure: Alternative Education for Children and Youth*, *48*(1), 10-17.

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2007). *The ABCs of CBM: A practical guide to curriculum-based measurement.* New York, NY: Guilford Press.

Huang, H. (2015). Can students themselves narrow the socioeconomic-status-based achievement gap through their own persistence and learning time? *Education Policy Analysis Archives*, *23*, 2-36.

Huddleston, A. P. (2014). Achievement at whose expense? A literature review of test-based grade retention policies in US schools. *Education Policy Analysis Archives*, *22*(18), 1-34.

Hudson, M. E., Browder, D. M., & Wood, L. A. (2013). Review of experimental research on academic learning by students with moderate and severe intellectual disability in general education. *Research and Practice for Persons with Severe Disabilities*, *38*(1), 17-29.

Hunley, S. A., Davies, S. C., & Miller, C. R. (2013). The relationship between curriculum-based measures in oral reading fluency and high-stakes tests for seventh grade students. *RMLE Online*, *36*(5), 1-8.

Hursh, D. (2013). Raising the stakes: High-stakes testing and the attack on public education in New York. *Journal of Education Policy*, *28*(5), 574-588.

Incantalupo, L., Treagust, D. F., & Koul, R. (2014). Measuring student attitude and knowledge in technology-rich biology classrooms. *Journal of Science Education and Technology*, *23*(1), 98-107.

Ingersoll, R. M., & May, H. (2011). The minority teacher shortage: Fact or fable? *Phi Delta Kappan*, *93*(1), 62-65.

Irving, K. (2007). Formative assessment improves student learning. *NSTA Reports, 18*(7), 6-8.

Jenkins, J., Schulze, M., Marti, A., & Harbaugh, A. G. (2017). Curriculum-based measurement of reading growth: Weekly versus intermittent progress monitoring. *Exceptional Children*, *84*(1), 42-54.

Jennings, J. L., & Bearak, J. M. (2014). “Teaching to the test” in the NCLB era: How test predictability affects our understanding of student performance. *Educational Researcher*, 43(8), 381-389. doi: http://10.3102/0013189X14554449

Jitendra, A. K., Dupuis, D. N., & Zaslofsky, A. F. (2014). Curriculum-based measurement and standards-based mathematics: Monitoring the arithmetic word problem-solving performance of third-grade students at risk for mathematics difficulties. *Learning Disability Quarterly*, *37*(4), 241-251.

Johnson, E. S., Jenkins, J. R., & Petscher, Y. (2010). Improving the accuracy of a direct route screening process. *Assessment for Effective Intervention*, *35*(3), 131-140.

Kantor, J. (2009). *Crafting white paper 2.0: Designing information for today’s time and attention challenge business leader*. Raleigh, NC: Lulu Enterprises.

Kara, F. (2015). Use of the drawing-writing technique to determine the level of knowledge of pre-service teachers regarding renewable energy sources. *Journal of*

Education and Practice, *6*(19), 215-225.

Kearns, L. L. (2016). The construction of 'illiterate' and 'literate' youth: the effects of high-stakes standardized literacy testing. *Race Ethnicity and Education*, *19*(1), 121-140.

Kennedy, G., Corrin, L., Lockyer, L., Dawson, S., Williams, D., Mulder, R.,…Copeland, S. (2014). Completing the loop: Returning learning analytics to teachers. In B. Hegarty, J. McDonald, & S.-K. Loke (Eds.), Rhetoric and reality: Critical perspectives on educational technology. *Proceedings Ascilite Dunedin*, 436-440.

Keller, L. A., Keller, R., Cook, R. J., & Colvin, K. F. (2016). Impact of accumulated error on item response theory pre-equating with mixed format tests. *Applied Measurement in Education*, *29*(1), 65-82.

Keller-Margulis, M. A., Shapiro, E. S., & Hintze, J. M. (2008). Long-term diagnostic accuracy of curriculum-based measures in reading and mathematics. *School Psychology Review*, *37*(3), 374-390.

Kettler, R. J. (2011). Computer-based screening for the new modified alternate assessment. *Journal of Psychoeducational Assessment*, *29*(1), 3-13.

Kettler, T., Russell, J., & Puryear, J. S. (2015). Inequitable access to gifted education: Variance in funding and staffing based on locale and contextual school variables. *Journal for the Education of the Gifted*, *38*(2), 99-117.

Kirkham, S., & Lampley, J. H. (2014). Predicting performance on the Tennessee comprehensive assessment for third grade reading students using reading curriculum based measures. *SRATE Journal*, *24*(1), 38-46.

Knight, S., Shum, S. B., & Littleton, K. (2013). Collaborative sensemaking in learning analytics. *CSCW and Education Workshop* (2013,23 February*): Viewing education as a site of work practice, co-located with the 16th ACM Conference on Computer Support Cooperative Work and Social Computing*, San Antonio, Texas. Retrieved from http://oro.open.ac.uk/36582/

Knight, S., Shum, S. B., & Littleton, K. (2014). Epistemology, assessment, pedagogy: Where learning meets analytics in the middle space. *Journal of Learning Analytics, 1*(2), 23-47.

Kop, R. (2012). The unexpected connection: Serendipity and human mediation in networked learning. *Educational Technology & Society*, *15*(2), 2-11.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, *41*(4), 212-218.

Latu, E., & Chapman, E. (2002). Computerised adaptive testing. *British Journal of Educational Technology, 33*(5), 619-622.

Lazendic, G., & Martin, A. J. (2018). Computer-adaptive testing: Implications for students' achievement, motivation, engagement, and subjective test experience. *Journal of Educational Psychology*, *110*(1), 27-45.

Lewis, S., & Hardy, I. (2015). Funding, reputation and targets: The discursive logics of high-stakes testing. *Cambridge Journal of Education*, *45*(2), 245-264.

Lit, I. W., & Lotan, R. (2013). A balancing act: Dilemmas of implementing a high-stakes performance assessment. *The New Educator*, *9*(1), 54-76.

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action:

Aligning learning analytics with learning design. *American Behavioral Scientist*, *57*(10), 1439-1459.

Lonn, S., McKay, T. A., & Teasley, S. D. (2017). Cultivating institutional capacities for learning analytics. *New Directions for Higher Education*, *2017*(179), 53-63.

López, F. A., & Thompson, S. S. (2010). The relationship among measures of written expression using curriculum-based measurement and the Arizona Instrument to Measure Skills (AIMS) at the middle school level. *Reading & Writing Quarterly*, *27*(1-2), 129-152.

Macfadyen, L. P., Dawson, S., Pardo, A., & Gasevic, D. (2014). Embracing big data in complex educational systems: The learning analytics imperative and the policy challenge. *Research & Practice in Assessment, 9*(2), 17-28.

Maier, M. P., Pate, J. L., Gibson, N. M., Hilgert, L., Hull, K., & Campbell, P. C. (2016). A quantitative examination of school leadership and response to intervention. *Learning Disabilities Research & Practice, 31*(2), 103-112.

Mao, X., & Xin, T. (2013). The application of the Monte Carlo approach to cognitive diagnostic computerized adaptive testing with content constraints. *Applied Psychological Measurement*, *37*(6), 482-496.

Maranto, J. H. (2015). The effect of standardized testing on historical literacy and educational reform in the US. *Academic Leadership Journal in Student Research*, *3*, 1-8.

Marston, D., Deno, S. L., Kim, D., Diment, K., & Rogers, D. (1995). Comparison of reading intervention approaches for students with mild disabilities. *Exceptional*

Children, *62*(1), 20-37.

Marston, D., Mirkin, P., & Deno, S. (1984). Curriculum-based measurement: An alternative to traditional screening, referral, and identification. *The Journal of Special Education*, *18*(2), 109-117.

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

McGlinchey, M. T., & Hixson, M. D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review*, *33*(2), 193-203.

McMaster, K. L., Shin, J., Espin, C. A., Jung, P. G., Wayman, M. M., & Deno, S. L. (2017). Monitoring elementary students' writing progress using curriculum-based measures: Grade and gender differences. *Reading and Writing*, *30*(9), 2069-2091.

Miller, K. C., Bell, S. M., & McCallum, R. S. (2015). Using reading rate and comprehension CBM to predict high-stakes achievement. *Journal of Psychoeducational Assessment*, *33*(8), 707-718.

Minarechová, M. (2012). Negative impacts of high-stakes testing. *Journal of Pedagogy/Pedagogický Casopis*, *3*(1), 82-100.

Minor, E. C., Desimone, L. M., Spencer, K., & Phillips, K. J. (2014). A new look at the opportunity-to-learn gap across race and income. *American Journal of Education*, *121*(2), 241-269.

Mitra, D., Mann, B., & Hlavacik, M. (2016). Opting out: Parents creating contested spaces to challenge standardized tests. *Education Policy Analysis*

Archives, *24*(31), 1-23.

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

Morris, D. S. (2015). Actively closing the gap? Social class, organized activities, and academic achievement in high school. *Youth & Society*, *47*(2), 267-290.

Muyskens, P., Betts, J., Lau, M. Y., & Marston, D. (2009). Predictive validity of curriculum-based measures in the reading assessment of students who are English language learners. *The California School Psychologist*, *14*(1), 11-21.

Neely, S. R. (2015). No Child Left Behind and administrative costs: A resource dependence study of local school districts. *Education Policy Analysis Archives*, *23*(26), 1-22.

Newhouse, C. P., & Tarricone, P. (2014). Digitizing practical production work for high-stakes assessments. *Canadian Journal of Learning and Technology*, *40*(2), 1-17.

Nichols, S. L., Glass, G., & Berliner, D. (2012). High-stakes testing and student achievement: Updated analyses with NAEP data. *Education Policy Analysis Archives*, *20*(20), 1-35.

Nichols, S. L., & Valenzuela, A. (2013). Education policy and youth: Effects of policy on practice. *Theory Into Practice, 52*(3), 152-159.

Nicol, D. J., & MacFarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Nistor, N., Baltes, B., Dascălu, M., Mihăilă, D., Smeaton, G., & Trăuşan-Matu, Ş. (2014).

Participation in virtual academic communities of practice under the influence of technology acceptance and community factors. A learning analytics application. *Computers in Human Behavior, 34,* 339-344.

Nistor, N., Derntl, M., & Klamma, R. (2015). Learning analytics: Trends and issues of the empirical research of the years 2011-2014. *Proceedings of the 10th European Conference on Technology Enhanced Learning: Design for Teaching and Learning in a Networked World* (EC-TEL 2015), 15-17 September 2015, Toledo, Spain (pp. 453-459). Springer.

No Child Left Behind Act of 2001, 20 U.S.C. § 6319 (2002).

Norris, D. M., & Baer, L. L. (2013). *Building organizational capacity for analytics*. Louisville, CO: EDUCAUSE.

Office of Superintendent of Public Instruction (2016). *Washington state report card* [Data file]. Retrieved from http://reportcard.ospi.k12.wa.us/summary

O'Malley, J. M., & Pierce, L. V. (1996). *Authentic assessment for English language learners: Practical approaches for teachers.* Boston, MA: Addison-Wesley.

Owen, L. (2016). The impact of feedback as formative assessment on student performance. *International Journal of Teaching and Learning in Higher Education*, *28*(2), 168-175.

Parker, R., Tindal, G., & Hasbrouck, J. (1991). Countable indices of writing quality: Their suitability for screening-eligibility decisions. *Exceptionality: A Special Education Journal*, *2*(1), 1-17.

Petscher, Y., Mitchell, A. M., & Foorman, B. R. (2015). Improving the reliability of student scores from speeded assessments: an illustration of conditional item response theory using a computer-administered measure of vocabulary. *Reading and Writing*, *28*(1), 31-56.

Piaget, J. (1951). *The child's conception of the world* (No. 213). Washington, DC: Rowman & Littlefield.

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

Piaget, J. (1954). Language and thought from a genetic perspective. *Acta Psychologica, 10*, 51-60.

Piaget, J. (1959). *The language and thought of the child* (Vol. 5). New York: Psychology Press.

Piaget, J. (1964). Part I: Cognitive development in children: Piaget: Development and learning. *Journal of Research in Science Teaching*, *2*(3), 176-186.

Pianta, R., Downer, J., & Hamre, B. (2016). Quality in early education classrooms: Definitions, gaps, and systems. *The Future of Children*, *26*(2), 119-137.

Polesel, J., Rice, S., & Dulfer, N. (2014). The impact of high-stakes testing on curriculum and pedagogy: A teacher perspective from Australia. *Journal of Education Policy*, *29*(5), 640-657.

Powell, V. (2012). Revival of the position paper: Aligning curricula and professional competencies. *Communication Teacher*, *26*(2), 96-103.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: The University of Chicago Press.

Ravitch, D. (2016). *The death and life of the great American school system: How testing and choice are undermining education.* New York, NY: Basic Books.

Reigeluth, C. M. (2016). Instructional theory and technology for the new paradigm of education. *Revista de Educación a Distancia*, 50.

Roberts-Mahoney, H., Means, A. J., & Garrison, M. J. (2016). Netflixing human capital development: Personalized learning technology and the corporatization of K-12 education. *Journal of Education Policy, 31*(4), 405-420.

Roehrig, A. D., Petscher, Y., Nettles, S. M., Hudson, R. F., & Torgesen, J. K. (2008). Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology, 46*(3), 343-366.

Ruecker, T. (2013). High-stakes testing and Latina/o students: Creating a hierarchy of college readiness. *Journal of Hispanic Higher Education, 12*(4), 303-320.

Sakamuro, S., Stolley, K., & Hyde, C. (2010). White paper: Organization and other tips. *Purdue University Online Writing Lab.* Retrieved from http://owl.english.perdue.edu/owl/resource/546/1

Sato, M. & Atkin, J. (2007). Supporting change in classroom assessment. *Educational Leadership, 64*(4), 76-79.

Schunk, D.H. (1991). *Learning theories: An educational perspective.* New York, NY: Macmillan.

Segool, N. K., Carlson, J. S., Goforth, A. N., von der Embse, N., & Barterian, J. A. (2013). Heightened test anxiety among young children: Elementary school students' anxious responses to high-stakes testing. *Psychology in the*

Schools, *50*(5), 489-499.

Seymour, C., & Garrison, M. (2015). Is the physical being taken out of physical education? On the possible effects of high-stakes testing on an embattled profession's curriculum goals. *Quest*, *67*(4), 400-408.

Shapiro, E. S. (2014). Tiered instruction and intervention in a response-to-intervention model. *RTI Action Network, 381*, 1-11.

Shapiro, E. S., Dennis, M. S., & Fu, Q. (2015). Comparing computer adaptive and curriculum-based measures of math in progress monitoring. *School Psychology Quarterly*, *30*(4), 470-487.

Shapiro, E. S., Solari, E., & Petscher, Y. (2008). Use of a measure of reading comprehension to enhance prediction on the state high stakes assessment. *Learning and Individual Differences, 18*(3), 316-328.

Shum, S. B., & Ferguson, R. (2012). Social learning analytics. *Educational Technology & Society, 15*(3), 3-26.

Shute, V. J., Leighton, J. P., Jang, E. E., & Chu, M. W. (2016). Advances in the science of assessment. *Educational Assessment*, *21*(1), 34-59.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, *57*(10), 1380-1400.

Siemens, G., Dawson, S., & Lynch, G. (2013). *Improving the quality and productivity of the higher education sector: Policy and strategy for systems-level deployment of learning analytics*. Sydney, Australia: Australian Government Office for Teaching and Learning.

Singh, P., Märtsin, M., & Glasswell, K. (2015). Dilemmatic spaces: High-stakes testing and the possibilities of collaborative knowledge work to generate learning innovations. *Teachers and Teaching*, *21*(4), 379-399.

Slavin, R. E., Cheung, A., Holmes, G., Madden, N. A., & Chamberlain, A. (2013). Effects of a data-driven district reform model on state assessment outcomes. *American Educational Research Journal*, *50*(2), 371-396.

Slavin, R. E., & Davis, N. (2006). *Educational psychology: Theory and practice.* Needham Heights, MA: Allyn and Bacon.

Smarter Balanced Assessment Consortium. (2016). *2013-2014 Technical Report.* North Billerica, MA: Curriculum Associates.

Snelbecker, G. E. (1974). Learning theory, instructional theory, and psychoeducational design. New York, NY: McGraw-Hill.

Speer, N. M., King, K. D., & Howell, H. (2015). Definitions of mathematical knowledge for teaching: Using these constructs in research on secondary and college mathematics teachers. *Journal of Mathematics Teacher Education*, *18*(2), 105-122.

Stage, S. A., & Jacobsen, M. D. (2001). Predicting student success on a state-mandated performance-based assessment using oral reading fluency. *School Psychology Review, 30,* 407-419.

Starkman, N. (2006). Building a better student. *T H E Journal*, *33*(14), 41-46.

Starr, J. D. (2012). A lack of depth: One preservice teacher's experiences in a post-NCLB world. *The Social Studies*, *103*(6), 241-246.

Starr, J., & Spellings, M. (2014). Examining high-stakes testing. *Academic Journal*, *14*(1), 70-77.

Stelzner, M. A. (2007). *Writing white papers: How to capture readers and keep them engaged.* Poway, CA: WhitePaperSource Publishing.

Stelzner, M. A. (2010). *How to write a white paper*. Retrieved from http://www.stelzner.com

Stevenson, N. A. (2015). Predicting proficiency on statewide assessments: A comparison of curriculum-based measures in middle school. *The Journal of Educational Research*, *108*(6), 492-503.

Taylor, J. L, & Bragg, D. D. (2015). *Optimizing reverse transfer policies and processes: Lessons from twelve CWID states*. Retrieved from http://occrl.illinois.edu/files/Projects/CWID/optimizing-reverse-transfer.pdf

Tempel, T., & Neumann, R. (2014). Stereotype threat, test anxiety, and mathematics performance. *Social Psychology of Education*, *17*(3), 491-501.

Tempelaar, D. T., Heck, A., Cuypers, H., van der Kooij, H., & van de Vrie, E. (2013, April). Formative assessment and learning analytics. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 205-209). Association for Computing Machinery.

Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior*, *47*, 157-167.

Thompson, G. L. (2017). Computer adaptive testing, big data and algorithmic approaches to education. *British Journal of Sociology of Education*, *38*(6), 827-840.

Thompson, G. L., & Allen, T. G. (2012). Four effects of the high-stakes testing movement on African American K-12 students. *The Journal of Negro Education*, *81*(3), 218-227.

Timmers, C. F., Braber-van den Broek, J., & van den Berg, S. M. (2013). Motivational beliefs, student effort, and feedback behaviour in computer-based formative assessment. *Computers & Education*, *60*(1), 25-31.

Tomlinson, C. (2001). *How to differentiate instruction in mixed ability classrooms* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Triola, M. F. (2006). *Elementary statistics*. Reading, MA: Pearson/Addison-Wesley.

Ünal-Karagüven, M. H. (2015). Demographic factors and communal mastery as predictors of academic motivation and test anxiety. *Journal of Education and Training Studies*, *3*(3), 1-12.

van den Broek, P., & Espin, C. A. (2012). Connecting cognitive theory and assessment: Measuring individual differences in reading comprehension. *School Psychology Review*, *41*(3), 315-326.

Van Norman, E. R., & Christ, T. J. (2016). How accurate are interpretations of curriculum-based measurement progress monitoring data? Visual analysis versus decision rules. *Journal of School Psychology*, *58*, 41-55.

VanPatten, B., Trego, D., & Hopkins, W. P. (2015). In-class vs. online testing in university-level language courses: A research report. *Foreign Language Annals*, *48*(4), 659-668.

Vie, J. J., Popineau, F., Bruillard, É., & Bourda, Y. (2017). A review of recent advances in adaptive assessment. In *Learning Analytics: Fundaments, Applications, and Trends* (pp. 113-142). New York, NY: Springer International Publishing.

von der Embse, N. P., & Witmer, S. E. (2014). High-stakes accountability: Student anxiety and large-scale testing. *Journal of Applied School Psychology, 30*(2), 132-156.

Vygotsky, L. S. (1986). *Language and thought*. Cambridge, MA: MIT Press.

Vygotsky, L. S. (1978). Interaction between learning and development. *Readings on the Development of Children*, *23*(3), 34-41.

Wagner, E., & Ice, P. (2012). Data changes everything: Delivering on the promise of learning analytics in higher education. *Educause Review, 47*(4), 32-36.

Wainer, H., Dorans, N., Eignor, D., Flaugher, R., Green, B. F., Mislevy, R. J., & Steinberg, L. (2001). Computerized adaptive testing: A primer. *Qual Life Res*, *10*(8), 733-734.

Wang, C. (2013). Mutual information item selection method in cognitive diagnostic computerized adaptive testing with short test length. *Educational and Psychological Measurement*, *73*(6), 1017-1035.

Wang, C., Chang, H. H., & Huebner, A. (2011). Restrictive stochastic item selection methods in cognitive diagnostic computerized adaptive testing. *Journal of Educational Measurement*, *48*(3), 255-273.

West, C. (2012). Teaching music in an era of high-stakes testing and budget reductions. *Arts Education Policy Review*, *113*(2), 75-79.

Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. In *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge*, 203-211. New York, NY: Association for Computing Machinery. doi:10.1145/2567574.2567588

Wolf, M. A., Jones, R., Hall, S., & Wise, B. (2014). Capacity enablers and barriers for learning analytics: Implications for policy and practice. *Alliance for Excellent Education.* Retrieved from http://all4ed.org/reports-factsheets/capacity-enablers-and-barriers-forlearning-analytics-implications-for-policy-and-practice/

Ydesen, C. (2014). High-stakes educational testing and democracy-antagonistic or symbiotic relationship? *Education, Citizenship and Social Justice*, *9*(2), 97-113.

Yell, M. L., Katsiyannis, A., Collins, J. C., & Losinski, M. (2012). Exit exams, high-stakes testing, and students with disabilities: A persistent challenge. *Intervention in School and Clinic*, *48*(1), 60-64.

Young, I. P., & Cox, E. P. (2014). A better leveled playing field for assessing satisfactory job performance of superintendents on the basis of high-stakes testing outcomes. *Board of Editors*, 4-15.

Young, M., & Muller, J. (2015). Curriculum and the specialization of knowledge: Studies in the sociology of education. New York, NY: Routledge.

Appendix A: The Project

"Predicting High-Stakes Tests Using the i-Ready"
Heath Thompson

Executive Summary

Twenty percent of the 2013-2014 sophomore class at Evergreen High School (EHS, pseudonym) are failing the Mathematics EOCE (EOCE), making these students ineligible to graduate. To help monitor the progress of students, EHS has recently introduced the i-Ready, a computerized adaptive Diagnostic (CAD) designed to assess the grade level at which sophomores are performing in mathematics based on student test scores. The purpose of this paper was to present the findings and recommendations of a doctoral research study of whether 10th grade students' achievement i-Ready math scores could predict the subsequent high-stakes mathematics scores on the EOCE while controlling for gender, ethnicity, and socioeconomic status (SES). Cognitive learning theories comprised the framework for this study, which posits that learning and development are cognitive processes dependent on previous knowledge, and central to measuring performance levels.

The primary research questions for this study: Does the i-Ready score predict mathematics scores on the EOCE? The second research question: Do gender, ethnicity, and SES moderate the relationship between i-Ready scores and the EOCE? Results showed a statistically significant correlation between the i-Ready and the EOCE and that gender, ethnicity, and SES did not moderate the relationship between the i-Ready and EOCE scores.

These results along with theory, best practices, and review of the literature resulted in the following recommendations to aid students in passing high-stakes tests:

1. Collect and analyze i-Ready data.
2. Provide additional resources for students in danger of failing the EOCE based on the i-Ready data.
3. Use Learning Analytics tools to gain empirical insights into teaching and learning

The Problem

Students are failing high-stakes tests at an alarmingly high rate (Nichols, Glass, & Berliner, 2012). At both the state and national level, students' failure to pass these tests may significantly impact a student's ability to graduate. Students, parents, and educators need a method to identify students who are in danger failing the high-stakes tests. In Washington State, twenty percent of the 2013-2014 sophomore class did not pass the mathematics portion, 23% failed the science section, and 14% failed the English section of the test (Office of Superintendent of Public Instruction [OSPI], 2016). In the same year, twenty-four percent of the students in Washington state did not graduate. These students may have completed all coursework needed to graduate, obtaining the necessary amount of credits required, but due to educational reform No Child Left Behind Act of 2001 (No Child Left Behind [NCLB], 2002), they failed to demonstrate learning required for a high school diploma (Daun-Barnett & St. John, 2012; Nichols et al., 2012).

The problem at EHS is that in the 2013-2014, year only 60% of high school students met standard one the EOCE, a graduation requirement. EHS has consistently scored low and has decreased 5% since 2011. Approximately 30% of students fail to graduate on-time. There could be many factors that cause EHS to perform lower than other schools, but one reason could be the school's high ethnicity rate and low SES (Huang, 2015). At EHS, 70% percent of students live below the national poverty line and are 25% Hispanic, 18% Black, 23% White, 20% Asian, and 4% Native Hawaiian/Pacific

Islander (OSPI, 2016). In keeping with the national trend, these students consistently score at a 30% lower rate on high-stakes tests than White students above the poverty line.

If a student fails a necessary required part of a high-stakes test, such as the mathematics portion of the EOCE, some research says that they will continue to fail the same exam without an intervention method to determine what their knowledge level is, and then provide the new knowledge necessary to pass (Miller, Bell, & McCallum, 2015; Singh, Märtsin, & Glasswell, 2015). Currently, there is no such intervention method at EHS, and before this study, no researchers have compared the i-Ready to the EOCE at the high school level in mathematics.

Constructivism

Constructivism is the theoretical foundation of this position paper. Constructivism, according to Piaget (1964), explains how humans create meaning based on their relationship to their environments by building upon prior knowledge. Piaget implies that students possess various levels of knowledge that differ depending on their experiences. Cognitive structures of learners can change only at specific points in the cognitive development, which suggests that these structures can go a period without changing beyond certain thresholds (Piaget, 1964). Cognitivist research has come to a similar conclusion. Anderson, Spiro, and Anderson (1978), Gagné (1979), and Rumelhart (1980) explained that knowledge stored in the long-term memory stores as schemata, i.e., mental structures that influence attention and absorb new knowledge that fits into existing patterns. Schemata tend to remain unchanged for more extended periods of time, thus defining relatively stable levels of knowledge, such as those described by Piaget.

Formative assessment and CADs open several possibilities of intervention and improvement of student achievement (Dunn & Mulvenon, 2009; Kingstone & Nash, 2011). Formative assessment builds the groundwork for adaptive teaching at the teacher level (Scriven, 1967). Taking learning prerequisites and student performance into account, teachers can compensate deficits, support and foster student strengths, or allow students to follow their preferences in open learning environments (Klieme & Warwas, 2011; Salomon, 1972). At the student level, formative assessment can support students' self-assessment. According to Boud (2013), "self-assessment provides the fundamental link with learning" (p. 15) since it requires students to consider useful characteristics and strategies that they can apply to their work, which in turn promote self-directed learning and lifelong learning skills (Knowles, 1975). Within this framework, students will find formative feedback most applicable when perceived as effective to in both providing useful information that benefits their learning.

Purpose and Design

The purpose of this study was to determine if there was a significant relationship between achievement in mathematics on the i-Ready and achievement in mathematics on the EOCE. This correlational study examined the relationship between i-Ready scores (predictor variable) and EOCE scores (criterion variable) while controlling for gender, ethnicity, and SES. A correlational design allows the measurement of a relationship between variables, but it does not determine the cause of a phenomenon, i.e., why a variable has certain values and by which other variables it is influenced (Triola, 2006). I chose a correlational design since the study sought to find a potential relationship between variables (Creswell, 2012).

In this study, I used a non-random collection of the entire target population because the school district already collected the data and made it available to the researchers for the study with permission. Creswell (2012) stated that to avoid confirmation bias, researchers should use the largest amount of data possible. The target population consisted of archival records for students in the 10th grade (N = 512) taking the i-Ready and the EOCE in 2014-2015. The statistical power analysis software G* Power 3.1 was used to calculate the minimum sample size for linear multiple regression by using a slope of .15, an alpha error probability of .05, and power set to .95, which indicated a minimum of n = 129 participants. Only students who were administered the i-Ready and the EOCE during the period of 2014-2015 were eligible to participate in the study. After removing the incomplete data, the actual sample consisted of N = 220 students, which was over the calculated minimum. I presented demographic frequencies and percentages of the sample in in Table 1.

Table 4

Demographic Frequencies and Percentages of the Sample

Distribution		*N*	%
Gender	Male	123	56
	Female	97	44
Ethnicity	Asian	40	18
	Black	65	29
	Hispanic	63	29
	Native Hawaiian/Other	8	4
	Two or More Races	11	5
	White	33	15
School Lunch Status	Free Lunch	144	66
	Reduced Lunch	15	7
	Not Eligible	61	28
Total Overall		220	

In this correlational study, I utilized archival i-Ready and EOCE data. The i-Ready measured an independent (predictor) variable, while the EOCE measured the dependent (criterion) variable. Eligible students in 10th grade took the i-Ready in September and took the EOCE in May. These assessments provided the data for the study.

Findings

The students scored on average 484 points (*SD* = 34.00) at i-Ready, and 389 points (*SD* = 45.00) at the EOCE test. In response to RQ1, the i-Ready scores significantly and positively predicted the EOCE test scores with β = .64 ($p < .001$), explaining $R^2 = .43$ of the criterion variance. In response to RQ2, the moderator analysis revealed that the demographic variables i-Ready x gender (β = .03, $p = .58$), i-Ready x ethnicity (β = -.08, $p = .15$) and i-Ready x SES (β = -.04, $p = .48$) did not predict the EOCE scores, meaning that there was no significant moderating effect. Thus, the findings confirmed the null hypothesis that suggested that gender, ethnicity, and SES did not moderate the relationship between the i-Ready and EOCE scores. Figure 2 displays the linearity between the i-Ready and EOCE scores.

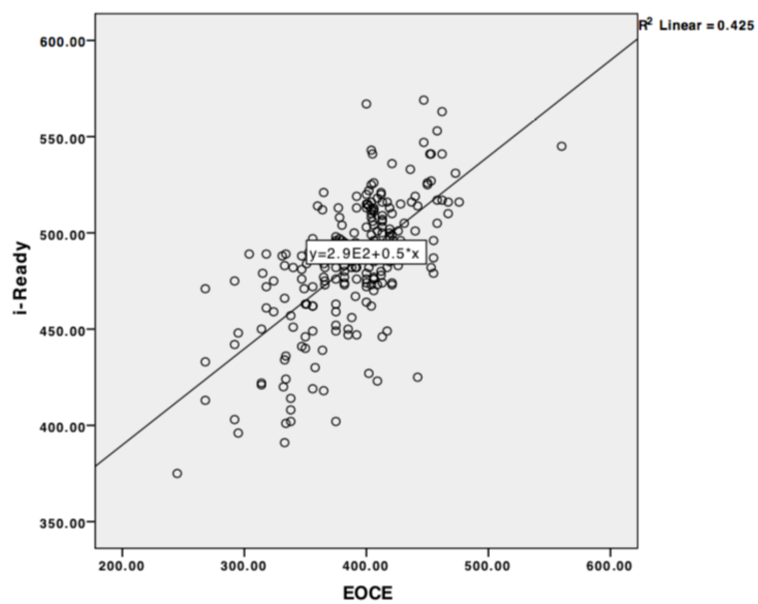

Figure 2. Scatterplot showing linearity between i-Ready scores and the EOCE.

i-Ready

Regression analysis revealed that the i-Ready scores were a strong predictor of EOCE test scores in this study. The findings support the notion of support the research Shapiro, Dennis, and Fu (2015) who observed that CADs could function as CBMs in test prediction, but this study provides research for a CAD's ability to predict proficiency at the high school level. From a broader perspective, the findings of this study support the assertions of Piaget (1951), Vygotsky (1978), Ausubel (1960), Bloom (1956), and other cognitive researchers (Anderson, Spiro, & Anderson, 1978; Gagné, 1979; Rumelhart, 1980) who implied learning occurs from level to level and is built through experience and necessity. The strength of the correlation suggested that i-Ready scores increased simultaneously with increases in EOCE test scores or decreased simultaneously with decreases in EOCE test scores.

Gender, Ethnicity, and SES

Regression analysis statistics revealed that gender, ethnicity, and SES had a weak correlation to EOCE scores, and were not statistically significant. The theories of Piaget (1959) supported these findings, who argued that knowledge will be different depending on experience. While demographic information did vary in regard to years of study in mathematics, all of these students have had similar levels of mathematics experience. As a result of this study, I suggested that the i-Ready, used as a tool determine the precise level of mathematics experience a student possessed was able to predict whether or not they would pass the EOCE successfully. While some studies have said that the achievement gap is widening (ESSA, 2015; Huang, 2015), the findings of this study suggested that tools like i-Ready could be useful in helping predict proficiency for high-stakes tests, particularly in areas with similar demographics of gender, ethnicity, and SES. Teachers do not control student's gender, ethnicity, or SES, but teachers do control the instruction and mediation, and overall quality of education that students receive. The i-Ready has shown to be an accurate predictor or EOCE scores. Through formative assessment using the i-Ready, teachers should give students additional support to those who demonstrate low performance in mathematics through remediation, regardless of their background through scaffolding instruction, differentiated learning, and enhanced instruction. Educators can measure students in mathematical terms in specific levels in knowledge, and in the case of the i-Ready, grade levels of knowledge.

Recommendation #1

Collect and analyze i-Ready data.

The findings of this study confirmed i-Ready's ability to predict test scores on the EOCE, establishing a need to collect data to predict which students may be in danger of failing their high-stakes exams at EHS. The i-Ready data and analysis will allow EHS to make data-based decisions to aid students in passing high-stakes exams. EHS currently does not have any data that could help predict which students may be in danger of failing. When EHS begins to collect and analyze i-Ready data, it can more accurately determine which students may be in need of additional support.

Data collection and analysis steps

- Administer the i-Ready in September
- Collect and analyze the i-Ready data
- Determine which students are in danger of failing the i-Ready

Side Box
Additional Costs - $0
Additional Resources - 0
Additional Time Spent – 1 day

Recommendation #2

Provide additional resources for students in danger of failing the EOCE based on the i-Ready data.

Since EHS administers the i-Ready in September, Recommendation #1 will provide teachers and administrators with seventh months of time to provide additional support and training for students in need before they take the EOCE. Research has been done connecting the effectiveness of formative feedback to computerized adaptive tests (Timmers, Braber-van den Broek & van den Berg, 2013). Understanding which level of knowledge a student is at will help them be able to take preventative steps to potentially increasing their scores before they take the EOCE (Incantalupo, Treagust, & Koul, 2014; Kara, 2015). Because of the i-Ready, we are now able to predict if students may be in danger of failing the EOCE. Without high-stakes test prevention, students may not be aware they are in danger of failing high-stakes exams until after they have failed it. Students, parents, and teachers deserve empirical, unbiased feedback of what level of knowledge students currently possess in relationship to what level of knowledge they need to be at to pass high-stakes tests. Recommendation #2 is that the school provide additional support to students that are identified to be in danger of failing the EOCE.

Recommendation # 3

Use Learning Analytics tools to gain empirical insights into teaching and learning.

The impact of both LA and CAD have been shown to increase the quality of student learning and education (Daniel, 2015; Siemens, Dawson, & Lynch, 2013). Some researchers suggest that schools adopt LA tools such as some CADs, for the use of educational improvement practices, such as test prediction (Baker & Invitado, 2014). Researchers have identified the fields that inhibit the use of LA tools in schools into the following: technological readiness (Arnold et al., 2014), leadership (Arnold, Lonn, & Pistilli, 2014), organizational culture (Carbonell, Dailey-Hebert, & Gijselaers, 2013), and staff capacity (Norris & Baer, 2013). This study recommends that EHS, along with other public schools, adopt LA as a means of improving student learning. EHS could consider other LA tools, such as online dashboards and student risk indexes.

With the present ability for public school systems to use LA tools, educators should consider the ethical concerns of LA. Tools that use LA give schools the potential to expedite feedback to teachers and students (Deakin Crick, 2017). While we gain a greater empirical understanding of student knowledge, we have a heightened responsibility to protect ownership of learning for the student and teacher (Knight, Shum, & Littleton, 2014). Additionally, LA raises ethical concerns when using data that is either tied to specific groups of students or when entities collect or use data for reasons not related to learning (Slade & Prinsloo, 2013). At EHS, administers can address some of these ethical concerns by making sure that learners have equal access to LA tools, keeping the agency of learning in the hands of students and teachers, and making sure data is used to help improve student learning.

Conclusions

Despite the vast wealth of technology that exists in education, work must be done to help make it more useful for the learner. The vast amount of educational technology and the data it generates can be overwhelming to school administrations, teachers, and students. Our goal should be to harness educational data to make it easy, intuitive and useful for the learner. While there is no easy solution to this task, this study presents one way to help students in danger of failing high-stakes mathematics tests. If we can use tools that already exist to help make learners more aware of what they know in relationship to high-stakes tests, it could significantly benefit high school students who need to pass these exams to graduate.

The use of LA and data tracking will continue to impact public school systems, despite some schools' reluctance to utilize it. Even without software that schools purchase to help track student data, businesses in the private sector, both small and large, make use of learner data for their purposes. While public schools might be trying to use traditional testing as means of measuring student knowledge, they should instead embrace the use of learner data, and make efforts to keep the power of LA-based tools in the hands of the teachers, students, and educational stakeholders. A practical first step would be to use the information at their disposal to help students in danger of failing required high-stakes examinations. Beyond this immediate and urgent need, the possibilities of online formative assessment systems, data tracking, and LA are seemingly endless.

References

Anderson, R. C., Spiro, R. J., & Anderson, M. C. (1978). Schemata as scaffolding for the representation of information in connected discourse. *American Educational Research Journal*, *15*(3), 433-440.

Arnold, K. E., Lonn, S., & Pistilli, M. D. (2014). An exercise in institutional reflection: The learning analytics readiness instrument (LARI). *Proceedings of the 4th International Conference on Learning Analytics and Knowledge* (LAK '14), 24-28 March 2014, Indianapolis, IN (pp. 163-167). New York, NY: Association for Computing Machinery.

Arnold, K. E., Lynch, G., Huston, D., Wong, L., Jorn, L., & Olsen, C. W. (2014, March). Building institutional capacities and competencies for systemic learning analytics initiatives. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge* (LAK '14), 24-28 March 2014, Indianapolis, IN (pp. 257-260). New York, NY: Association for Computing Machinery.

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In *Learning analytics* (pp. 61-75). New York, NY: Springer.

Bloom, B. S. (1956). *Taxonomy of educational objectives, handbook I: The cognitive domain*. New York, NY: David McKay Co Inc.

Boud, D. (2013). *Enhancing learning through self-assessment*. New York, NY: Routledge.

Carbonell, K. B., Dailey-Hebert, A., & Gijselaers, W. (2013). Unleashing the creative potential of faculty to create blended learning. *The Internet and Higher*

Education, *18*, 29-37.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (Laureate Education ed.). Boston, MA: Pearson Education.

Daniel, B. (2015). Big data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, *46*(5), 904-920.

Daun-Barnett, N., & John, E. P. S. (2012). Constrained curriculum in high schools: The changing math standards and student achievement, high school graduation and college continuation. *Education Policy Analysis Archives*, *20*(5), 1-24.

Deakin Crick, R. (2017). Learning analytics: Layers, loops, and processes in a virtual learning infrastructure. In *Handbook on Learning Analytics* (pp. 291-308). Society for Learning Analytics Research. doi:10.18608/hla17.025

Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation, 14*(7), 1-11.

Every Student Succeeds Act, S.1117, 114th Cong. (2015). Retrieved from: https://www.congress.gov/bill/114th-congress/senate-bill/1177/text

Ferguson, R., & Shum, S. B. (2012). Towards a social learning space for open educational resources. *Collaborative Learning*, *2*, 309-327.

Gagné, R. M. (1979). *The conditions of learning.* New York, NY: Holt, Rinehart and Winston.

Huang, H. (2015). Can students themselves narrow the socioeconomic-status-based achievement gap through their own persistence and learning time? *Education Policy Analysis Archives*, *23*, 2-36.

Incantalupo, L., Treagust, D. F., & Koul, R. (2014). Measuring student attitude and knowledge in technology-rich biology classrooms. *Journal of Science Education and Technology*, *23*(1), 98-107.

Kara, F. (2015). Use of the drawing-writing technique to determine the level of knowledge of pre-service teachers regarding renewable energy sources. *Journal of Education and Practice*, *6*(19), 215-225.

Kingston, N., & Nash, B. (2011). Formative assessment: A meta- analysis and a call for research. *Educational Measurement: Issues and Practice*, *30*(4), 28-37.

Klieme, E., & Warwas, J. (2011). Konzepte der individuellen Förderung [Concepts of individual student performance fostering]. *Zeitschrift für Pädagogik*, 57(6), 805-818.

Knight, S., Shum, S. B., & Littleton, K. (2014). Epistemology, assessment, pedagogy: Where learning meets analytics in the middle space. *Journal of Learning Analytics, 1*(2), 23-47.

Knowles, M. (1975). *Self-directed learning*. Chicago, IL: Follet.

Miller, K. C., Bell, S. M., & McCallum, R. S. (2015). Using reading rate and comprehension CBM to predict high-stakes achievement. *Journal of Psychoeducational Assessment*, *33*(8), 707-718.

Nichols, S. L., Glass, G., & Berliner, D. (2012). High-stakes testing and student achievement: Updated analyses with NAEP data. *Education Policy Analysis*

Archives, 20(20), 1-35.

No Child Left Behind Act of 2001, 20 U.S.C. § 6319 (2002).

Norris, D. M., & Baer, L. L. (2013). *Building organizational capacity for analytics*. Louisville, CO: EDUCAUSE.

Office of Superintendent of Public Instruction (2016). *Washington state report card* [Data file]. Retrieved from http://reportcard.ospi.k12.wa.us/summary

Piaget, J. (1951). *The child's conception of the world*. Washington, DC: Rowman & Littlefield.

Piaget, J. (1959). *The language and thought of the child* (Vol. 5). New York, NY: Psychology Press.

Piaget, J. (1964). Part I: Cognitive development in children: Piaget: Development and learning. *Journal of Research in Science Teaching, 2*(3), 176-186.

Rumelhart, D. E. (1980) Schemata: The building blocks of cognition. In *Theoretical issues in reading comprehension* (pp. 33-58). Hillsdale, NJ: Lawrence Erlbaum.

Salomon, G. (1972). Heuristic models for the generation of aptitude-treatment interaction hypotheses. *Review of Educational Research, 42*(3), 327-343.

Scriven, M. (1967). The methodology of evaluation. In *Perspectives of Curriculum Evaluation, Vol. I* (pp. 39-83). Chicago, IL: Rand McNally.

Shapiro, E. S., Dennis, M. S., & Fu, Q. (2015). Comparing computer adaptive and curriculum-based measures of math in progress monitoring. *School Psychology Quarterly, 30*(4), 470-487.

Siemens, G., Dawson, S., & Lynch, G. (2013). *Improving the quality and productivity of* *the higher education sector: Policy and strategy for systems-level deployment of learning analytics*. Sydney, Australia: Australian Government Office for Teaching and Learning.

Singh, P., Märtsin, M., & Glasswell, K. (2015). Dilemmatic spaces: High-stakes testing and the possibilities of collaborative knowledge work to generate learning innovations. *Teachers and Teaching*, *21*(4), 379-399.

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, *57*(10), 1510-1529.

Timmers, C. F., Braber-van den Broek, J., & van den Berg, S. M. (2013). Motivational beliefs, student effort, and feedback behaviour in computer-based formative assessment. *Computers & Education*, *60*(1), 25-31.

Triola, M. F. (2006). *Elementary statistics*. Reading, MA: Pearson/Addison-Wesley.

Vygotsky, L. (1978). Interaction between learning and development. *Readings on the Development of Children*, *23*(3), 34-41.

Appendix B: Letter of Cooperation from a Research Partner

Community Research Partner Name: [redacted] High School
Contact Information: Dr. [redacted]

10/24/2017

Dear Heath Thompson,

Based on my review of your research proposal, I give permission for you to conduct the study entitled "Does the i-Ready Diagnostic Predict High School Students' High-Stakes Mathematics Test Scores?" within the [redacted] School District. As part of this study, I authorize you to utilize archival i-Ready and EOCE test score data in mathematics for 10th graders at [redacted] High School.

We understand that our organization's responsibilities include provision of these records. We reserve the right to withdraw from the study at any time if our circumstances change.

I understand that the student (Heath Thompson) will not be naming our organization in the doctoral project report that is published in ProQuest.

I confirm that I am authorized to approve research in this setting and that this plan complies with the organization's policies.

I understand that the data collected will remain entirely confidential and may not be provided to anyone outside of the student's supervising faculty/staff without permission from the Walden University IRB.

Sincerely,

[redacted]

Dr. [redacted]
Principal, [redacted] High School
[redacted]