Perceptions of Middle School Math and Language Arts Teachers on High-Stakes Testing Cultures in Public Schools

Robert Daniello
danielro@shu.edu

Follow this and additional works at: https://scholarship.shu.edu/dissertations

Part of the Curriculum and Instruction Commons, Curriculum and Social Inquiry Commons, Educational Administration and Supervision Commons, and the Educational Leadership Commons

Recommended Citation
Daniello, Robert, "Perceptions of Middle School Math and Language Arts Teachers on High-Stakes Testing Cultures in Public Schools" (2018). Seton Hall University Dissertations and Theses (ETDs). 2493.
https://scholarship.shu.edu/dissertations/2493
Perceptions of Middle School Math and Language Arts Teachers on High-Stakes Testing Cultures in Public Schools

Robert Daniello

Dissertation Committee

Barbara Strobert, Ed.D., Mentor
Anthony Colella, Ph.D.
Sharon Amato, Ed.D.

Submitted in partial fulfillment of the requirements for the degree of Doctor of Education

Seton Hall University
2018
SETON HALL UNIVERSITY
COLLEGE OF EDUCATION AND HUMAN SERVICES
OFFICE OF GRADUATE STUDIES

APPROVAL FOR SUCCESSFUL DEFENSE

Robert Daniello, has successfully defended and made the required modifications to the text of the doctoral dissertation for the Ed.D. during this Spring Semester 2018.

DISSERTATION COMMITTEE
(please sign and date beside your name)

Mentor: Dr. Barbara Strobert

Committee Member: Dr. Anthony Colella

Committee Member: Dr. Sharon Amato

The mentor and any other committee members who wish to review revisions will sign and date this document only when revisions have been completed. Please return this form to the Office of Graduate Studies, where it will be placed in the candidate's file and submit a copy with your final dissertation to be bound as page number two.
Abstract

Instruction in public schools today is highly prescribed to meet state standards, which, in turn, prepare students for success on standardized assessments. Teachers in language arts and mathematics are being held accountable for standardized assessment results in their end-of-year, summative evaluations. The development of curriculum and delivery of instruction is being impacted and, most specifically, revised and paced according to skills required to demonstrate levels of proficiency on standardized assessments. No Child Left Behind (2001) changed the game for teachers and started the mandate of placing teacher accountability and evaluation on high-stakes tests. However, it was difficult to place a federal mandate on test scores when states were taking different assessments and working with different standards for proficiency. Common Core Curriculum Standards (2010) attempted to formalize national standards and presented PARCC as the national standard for assessment. Race to the Top provided states monetary rewards for adopting Common Core Curriculum Standards and for entering into the PARCC Consortium.

Grades 3-8 language arts and mathematics teachers in New Jersey currently teach under a mandate that ties 30% of their end-of-year, summative evaluations to standardized assessment results. This is known in New Jersey as a Student Growth Percentile (SGP). This mandate holds teachers directly accountable for high-stakes testing outcomes. This does not take into consideration many of the proven factors that dictate academic performance and proficiency, such as socioeconomic status, limited English proficiency, and transience of students. Teachers are under pressure to make sure their students know the skills required for PARCC and to analyze growth and progress throughout the school year. This can get in the way of creativity
in delivering lessons and has the potential to narrow the curriculum to only skills required for standardized assessment proficiency.

PARCC also changed the way in which standardized assessments are administered to students. It delivered the assessments to students through computer software for the first time. This placed a burden on teachers to make sure their students were proficient in typing skills and other technology skills necessary to navigate through a timed, computerized assessment. This study examined whether or not instructional time is being taken away from language arts and mathematics instruction to ensure technology proficiency required to navigate through a computer-based standardized assessment.

It is important to examine perceptions of language arts and math teachers who are objectively held accountable for standardized assessment results. Specifically, the researcher examined any effects on the development of curriculum and delivery of instruction in a middle school setting. This study also examined any time spent away from teaching language arts and mathematics skills in order to ensure technology tools necessary to take a computerized assessment are learned.
Acknowledgments

I would like to acknowledge all of my colleagues and friends who have supported my efforts towards this accomplishment. First and foremost, I sincerely thank and acknowledge my mentor, Dr. Barbara Strobert, who worked with me patiently and tirelessly throughout this journey. You never lost touch and continued, over the years, to ensure that I was staying focused and on task. Your mentorship has developed into a friendship that has had a major influence on my professional career. Thank you for never giving up on me. To all of the colleagues and community members I have worked with for over 20 years in the Secaucus and Fort Lee school systems, I thank you greatly for allowing me to be part of your success and for giving me the opportunity to work with your children in and out of the classrooms. Thank you to my dissertation committee members, Dr. Sharon Amato and Dr. Anthony Colella, for the time and support you provided to me throughout the IRB and defense processes. To Dr. Amato, thank you for trusting my worth to schools and for giving me the opportunity to expand my career at Lewis F. Cole Middle School in Fort Lee. Your persistence on questioning my dissertation status kept me tuned to the importance of attaining goals. To my close ally and friend, Mrs. Gina Ruesga, thank you for all of your support throughout this process and for your unwavering devotion to making our school a great place for students to learn. I will use the knowledge gained from this research to help support all of my students grow socially and emotionally throughout their lives and to not place value of success or to determine career paths based on standardized test results.
Dedication

I dedicate this dissertation to my immediate and extended family members, who have shown me love and compassion throughout my life. To my wife Tia and children Alex and Ryan, I thank you for showing me the meaning of wealth and providing me the joys of making our home. This accomplishment would not have come to fruition without your support and acceptance of time spent away from family. To my parents, thank you for always encouraging the value of education and the importance of social and emotional growth. The tools you provided to me to continue moving forward both personally and professionally have made me vastly content with the choices I have made. To all of the students who have crossed my path as a teacher and school administrator, you have taught me so much more than I was ever able to teach you. You are the inspiration for change in the world and your power is never to be underestimated. Tenacity and persistence are the keys to success!
Table of Contents

CHAPTER 1. INTRODUCTION............................................................................................1
Statement of the Problem..................................................................................................6
Purpose of the Study...........................................................................................................8
Conceptual Framework.......................................................................................................12
Research Questions...........................................................................................................13
Design and Methodology.................................................................................................13
Significance of the Study...................................................................................................15
Delimitations and Limitations..........................................................................................16
Chapter Summary.............................................................................................................17

CHAPTER 2. LITERATURE REVIEW...........................................................................19
Historical Perspective........................................................................................................19
The Eight-Year Study........................................................................................................19
A Nation At Risk...............................................................................................................20
No Child Left Behind.......................................................................................................21
Race To The Top...............................................................................................................22
Existing Literature Regarding High-Stakes Testing......................................................23
Six Effects of High-Stakes Testing...................................................................................24
Curriculum and Teacher Effectiveness............................................................................24
Inability to Prove Validity of Standardized Tests............................................................26
School Funding and Costs of Standardized Testing.......................................................29
Political Agendas of Standardized Testing.....................................................................30
School Culture as a Result of Standardized Testing.......................................................34
Correlation of Standardized Testing and Dropout Rates.................................................36
Positive Culture of High-Stakes Testing...........................................................................38
Review Methods...............................................................................................................39
Limitations of the Review................................................................................................41
Criteria for Inclusion and Exclusion...............................................................................42
Corruption and Cheating Associated with Standardized Testing.................................42

CHAPTER 3 – METHODOLOGY.................................................................................44
Research Design...............................................................................................................47
Sampling...........................................................................................................................47
Profiles of the Schools Selected.......................................................................................48
Participant Data...............................................................................................................50
Data Collection................................................................................................................50
Data Management and Analysis Plan.............................................................................52
Field Notes and Reflective Memos................................................................................53
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coding Scheme</td>
<td>53</td>
</tr>
<tr>
<td>Validity and Reliability</td>
<td>53</td>
</tr>
<tr>
<td>Role of the Researcher</td>
<td>54</td>
</tr>
<tr>
<td>CHAPTER 4 – FINDINGS</td>
<td>56</td>
</tr>
<tr>
<td>Research Question 1 Themes &amp; Findings</td>
<td>57</td>
</tr>
<tr>
<td>Interview Results for Research Question 1</td>
<td>58</td>
</tr>
<tr>
<td>Research Question 1 Summary</td>
<td>63</td>
</tr>
<tr>
<td>Research Question 2 Themes &amp; Findings</td>
<td>64</td>
</tr>
<tr>
<td>Interview Results for Research Question 2</td>
<td>65</td>
</tr>
<tr>
<td>Research Question 2 Summary</td>
<td>71</td>
</tr>
<tr>
<td>Research Question 3 Themes &amp; Findings</td>
<td>72</td>
</tr>
<tr>
<td>Interview Results for Research Question 3</td>
<td>73</td>
</tr>
<tr>
<td>Research Question 3 Summary</td>
<td>78</td>
</tr>
<tr>
<td>Final Thoughts</td>
<td>79</td>
</tr>
<tr>
<td>Chapter 4 Summary</td>
<td>80</td>
</tr>
<tr>
<td>CHAPTER 5 – CONCLUSIONS AND RECOMMENDATIONS</td>
<td>81</td>
</tr>
<tr>
<td>Introduction</td>
<td>81</td>
</tr>
<tr>
<td>Summary of Major Findings</td>
<td>82</td>
</tr>
<tr>
<td>Summary of Findings for Research Question 1</td>
<td>87</td>
</tr>
<tr>
<td>Summary of Findings for Research Question 2</td>
<td>89</td>
</tr>
<tr>
<td>Summary of Findings for Research Question 3</td>
<td>93</td>
</tr>
<tr>
<td>Limitations of the Study</td>
<td>95</td>
</tr>
<tr>
<td>Findings Related to Theoretical Framework</td>
<td>95</td>
</tr>
<tr>
<td>Recommendations for Practice, Policy, and Future Research</td>
<td>96</td>
</tr>
<tr>
<td>Recommendations for Practice</td>
<td>96</td>
</tr>
<tr>
<td>Recommendations for Policy</td>
<td>97</td>
</tr>
<tr>
<td>Recommendations for Future Research</td>
<td>98</td>
</tr>
<tr>
<td>Conclusions</td>
<td>99</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>101</td>
</tr>
</tbody>
</table>
List of Tables

Table 1. Connecting Research Questions with Respective Interview Questions........51
Table 2. Teachers’ Responses Related to Factors..........................................................69

List of Figures

Figure 1. Creswell Analysis Process............................................................55
CHAPTER 1
INTRODUCTION

The discussions of teaching have often focused on classroom instruction as both an art and a science. Standards, curriculum, child development, and early interventions are examples of the science aspect of teaching. These examples are prescribed focuses that provide structure and understanding to the subject matter taught and expectations at each grade level. The art of teaching delves into how the subject and curriculum are delivered by the teacher. It also speaks to the manner in which a teacher understands and recognizes the skill differences within children and differentiates instruction. However, the age of high-stakes testing is upon us, and teachers in states like New Jersey are being evaluated based on annual student growth in standardized assessment results. The science is there; we look at the numbers and determine whether or not a pre-determined growth percentage has been met. Curriculum and instruction are now driven by standardized test scores, and instructional time is being spent on standards where students perform below par (Resnick & Resnick, 1992). Weaknesses in reading, writing, and math become identified by analyzing standardized test scores of individual students and in schools as a whole.

The questions then become apparent. Is the age of standardized testing creating an environment of success for our students by creating prescriptive instruction? Or does it hamper the teacher’s ability to deliver instruction in a creative, art-like, subjective manner? This research study focused on the perceptions of teachers affected by the new testing mandates that weigh student growth on standardized assessments as part of their summative evaluation. The primary focus was on middle school teachers in Grades 7 and 8 who teach language arts and mathematics in New Jersey. The overarching goal was to determine whether or not teachers who
accept the invitation for this research study feel creatively hampered by the culture of standardized assessments or if they notice student progress and advancement of skills in reading, writing, and math.

In 2001, President George W. Bush signed legislation titled No Child Left Behind, which placed an unprecedented amount of accountability on standardized testing data. This was the beginning of using standardized measurements of math and language arts proficiency to determine a school’s worth and the effectiveness of classroom teachers. The policy of using data in the classroom has strengthened as the 21st century has progressed. For example, in 2016, New Jersey Governor Chris Christie directed public schools to adopt a practice of weighing 30% of a teacher’s summative evaluation to PARCC standardized test score growth in language arts and mathematics. With that said, arguments have been made with claims that schools have shifted curriculum away from creative components of education and towards rigorous test preparation (Blazer, 2011). Proponents of accountability argue that standardized measurement is an objective means of gauging student achievement and teacher effectiveness (Kuncel & Hezlett, 2007).

The driving force behind standardized testing is accountability (Sahlberg, 2010). Accountability can be measured by examining critical attributes of school districts such as graduation rates of students, attendance rates of students, and gradual growth of language arts, math, and science scores on annual state standardized tests. Under NCLB (2001), schools started receiving report cards and scores indicating how students performed within these categories. Defined consequences were then laid out for schools deemed to be failing based solely on raw data. Consequences included allowing parents to have a choice in their child’s home school if the current school is deemed a failure. All of this placed a heavy burden on schools to meet pre-
determined proficiency levels in math and language arts, which created an unrealistic goal and held teachers accountable for students with low proficiency levels regardless of time in country, special needs, socioeconomic status, or other inhibiting factors. Under NCLB (2001), annual growth did not matter unless that growth reached the respective state’s pre-determined proficiency level.

Race to the Top was introduced in 2009 as part of the American Recovery and Reinvestment Act (U.S. Department of Education, 2009). This law generated $4.35 billion in grants in which states were forced to compete on the amount of progress made on improving student outcomes and closing achievement gaps, which were two distinct factors for success in NCLB. Race to the Top introduced accountability measures such as merit pay for teachers and administrators if test scores reflected student growth. In order to qualify for a piece of this funding, states needed to eliminate any barriers or laws that prevented standardized test scores from affecting a teacher’s or administrator’s evaluation (Brill, 2011). The intent was to improve teacher and administrator effectiveness and tie any retention of those teachers or administrators to high-stakes test scores (Baker, Green & Oluwole, 2013).

Common Core Curriculum Standards in language arts and mathematics were introduced in New Jersey in June of 2010. The goal of CCCS was to “make it easier for states to pool information and resources to develop a shared set of high-quality tests to better evaluate student progress” (NJDOE, 2010). The idea behind this is that we can measure all students across the nation on the same standards and accountability measurement tool; by 2011, two sets of high-stakes tests were being developed to measure students on one scale (Sparks, 2011). However, Nelson and Eddy (2008) made the argument that one test for all students is incapable of producing any measurable outcomes to be used for student growth models.
Opponents of a high-stakes testing culture will argue that standardized test preparation and the anxiety that comes along with placing data-driven, evaluative measurements on teachers has an adverse effect on a school’s culture and climate. Costigan and Crocco (2007) made the assertion that test-driven, data-filled environments disengage students from discovery and lead them towards an environment of always requiring the right answer immediately. The argument is that this can be extremely stressful to students and cause a negative attitude towards the classroom and school. Maylone (2002) argued that standardized test data do not account for major pre-determining factors such as socioeconomics, language acquisition, attendance rates, or parental involvement.

In 2008, Jones conducted research to find that a high school’s overall High School Proficiency Assessment (HSPA) achievement could be predicted based on that school’s demographic data as published in the annual New Jersey Report Card. This was highlighted in Turnamian’s 2012 research that looked at the value of demographic data in explaining Grade 3 NJASK scores. Turnamian (2012) argued that testing data could be predicted based on out-of-school variables such as socioeconomic status. Turnamian progressed with his argument and made the point that the data being produced from standardized testing are in no way a clear indication of measuring the successes or failures of a particular school’s programs, teachers, or administrators.

The culture of testing students for the purpose of driving instruction has become the norm in public schools across the nation. A heated debate has emerged on the validity of test data and the necessity of creating benchmarks for students to meet. Church and Elliot (2001) noted that the testing culture has limited the flexibility of teachers to respond to the realistic needs of their students. Santman (2002) argued that instructional approaches are lost when
teachers have to spend most of their time preparing students for testing. Kuncel and Hezlett (2007), however, argued that standardized tests are good indicators of success in higher education forums and “provide useful information for predicting subsequent student performance across many disciplines” (p. 2).

Cheating has also become a publicly raised issue, which opponents of standardized testing claim is a result of the pressures of high-stakes testing. A New York Times article written by Richard Fausset and published on September 29, 2014, reported on a criminal trial for Atlanta’s public school employees. According to the New York Times article, “A state investigation in 2011 found that 178 principals and teachers in the city school district were involved in cheating on standardized tests. Dozens of former employees of the school district were fired or resigned, and 21 educators pleaded guilty to crimes like obstruction and making false statements.” It was also explained that 11 of these Atlanta school teachers and administrators were convicted on April 2, 2015, on charges ranging from racketeering to theft and making false statements in the conclusion of this circumstance of standardized test cheating.

Arguments are made that pressures to have students perform up to growth standards as measured in NCLB, Race to the Top, and ESSA create a culture where teachers and administrators become desperate to be successfully evaluated and labeled as effective. Bohte and Meier (2002), however, made the claim that leaders in any organization gain incentives for increasing the quality of organizational outputs if performance is based on numerical data. Policy makers consistently make the argument that there needs to be an objective measurement of student growth in order to effectively gauge the performance of teachers in the classrooms.

Another argument that is raised consistently with standardized testing is that a testing culture generates large amounts of funding that get shifted away from the classroom. Opponents
of testing claim that funding is being shifted towards materials and resources used to prepare students for the tests. They also claim that decisions driving the high-stakes culture movement are being made by bureaucrats who designed legislation such as No Child Left Behind and its successors Race to the Top and Every Student Succeeds Act, with the intention that these initiatives will drive reform and shift focus onto accountability if students fail. In response to this argument, Race to the Top offered states financial rewards for following the rules and adopting initiatives considered reform, such as data-driven evaluation processes for teachers and the adoption of a standardized, nationally recognized set of language arts and math standards. This provided winning states incentives for creating a more objective, numbers-driven system of accountability. Legislators argued that funding for implementing this objective was being offset by federal and state monies and that accountability needed to be raised for teachers and administrators to reform schools that fail.

**Statement of the Problem**

Testing cultures have been created in America that hold teachers and administrators accountable for student growth in reading, writing, and mathematics. The theory accepted by opponents is that every school and every classroom in America cannot be evaluated equally using the same skill measurement due to pre-determining factors such as socioeconomic status and parental education levels (Kuncel & Hezlett, 2007). They argue that a teacher’s role has shifted from one that fosters growth towards a prepared, content life into one that fosters growth in test scores. Proponents of higher accountability argue that there needs to be an objective, numbers-driven system to measure growth and the effectiveness of the instruction being delivered. They argue that a teacher needs to teach towards the skills outlined in the Common
Core Curriculum Standards in order to effectively measure progress and growth in their academic setting.

In 2002, Maylone conducted a study that focused on the Michigan Educational Assessment Program’s (MEAP) data. Maylone (2002) concluded that standardized test scores could be predicted simply by the socioeconomic status of students. This was done through a multiple regression analysis in high school settings. Out-of-school variables were closely examined and weighed against in-school variables such as time on task for testing preparation. Maylone argued that we do not take into consideration the out-of-school variables such as socioeconomic status and English proficiency for English Language Learners. This causes us to be unfair to teachers or administrators when we weigh their evaluations based on student test scores. His conclusions included the opinion that tying evaluative measures to test scores created unnecessary anxiety to have students perform well on those tests regardless of the mastery of basic skills. Turnamian (2012) argued that high-stakes testing data are not an accurate measurement of the success of schools nor the academic growth of students. Ball (2001) and Niesz (2010) both noted a focus for departments of education to create an image of accountability as opposed to making real efforts for authentic school reform. They argued that reform in schools should come by identifying local concerns and addressing them as learning communities with educators leading the discussions.

To counter this argument, Phelps (2005) wrote a book titled *Defending Standardized Testing*. In his book, he makes the claim that standardized testing systems have become more advanced and relevant to measuring specific skills in reading, writing, and mathematics. He also states that tests have become more dependable and reliable in tracking student growth and that we have never before had the wealth of data we currently have in order to evaluate and assess a
student’s growth objectively without annotation. Phelps (2005) also weighs the value of accurate testing data versus the subjective assessment by teachers of their students.

Sergiovanni (2000), however, argued that high-stakes testing and an emphasis on accountability contradict the core values of education. He stated that a “one-size-fits-all” approach to national standards and high-stakes tests does not take into consideration key factors affecting the quality of a child’s education prior to even walking through the school’s doors. His leading argument is that high-stakes testing places undue pressure on teachers and administrators to meet growth objectives. His claim is that a pressure-filled approach greatly affects the quality of instruction in the classroom and places all students into the same growth model.

Under AchieveNJ (2013), New Jersey’s public school teachers must demonstrate growth in math and language arts standardized assessment scores, known as a Student Growth Percentile or SGP. Since 2013, teachers in Grades 3-8 who are responsible for delivering math and language arts instruction in New Jersey’s public schools have had a percentage of their year-end, summative evaluations tied to standardized assessment growth. Currently, in the 2017-2018 school year, that growth counts for 30% of the year-end, summative assessment. New Jersey uses PARCC as the assessment standard, which moved the traditional paper and pencil method of administering standardized assessment to a computer-based approach. It is also important to determine whether or not instructional time in language arts and math are being compromised due to the notion that technical skills such as timed typing and use of drop-down calculators are needed prior to taking a computer-based assessment.

**Purpose of the Study**

Testing has become an annual, high-stakes rite of passage in public schools. Across public middle schools, language arts and mathematics teachers are being evaluated and labeled
as effective or ineffective partially based on high-stakes test results. The culture of high-stakes testing weighs this level of effectiveness on what amounts to approximately three days each of language arts and mathematics testing in a lengthy school year. The testing that currently takes place in New Jersey’s public schools measures growth of students each school year in the areas of reading, writing, and mathematics. A recent shift in new evaluation systems of teachers based on test results has taken effect in New Jersey’s schools based on Department of Education mandates outlined in AchieveNJ (2013). The purpose of this study was to explore the perceptions of middle school teachers in New Jersey on the mandate that 30% of their summative evaluation is weighted by standardized assessment growth in language arts and mathematics. I explored perceptions of middle school language arts and mathematics teachers on the high-stakes testing environment as they design lessons and deliver instruction to students in the classroom.

Middle school is a major transitioning time for students. Physical and emotional maturity are key factors that drive their motivations. It has been proven that middle schools thrive in instructional environments that support project-based approaches, cooperative learning structures, and technology-driven lessons (Rothenberg, 1993; Willis, 2007). Questions have arisen about standardized testing preparation and whether or not there is a major effect on instructional approaches. Opponents argue that standardized testing preparation does not allow for much creativity or discovery beyond understanding of what the test will present to the students. Proponents argue that standardized testing preparation ensures that basic skills are being taught and that students should be capable of producing measurable outputs that gauge their levels of proficiency in language arts and mathematics.
Phelps (2005) made the claim that standardized tests are an effective means of measuring basic skills growth; however, he also argues that standardized tests need to be evaluated annually to create better, more accurate, systems of measurement. This is currently taking place in New Jersey’s public schools where the former standardized test titled New Jersey Assessment of Skills and Knowledge (NJASK) was replaced in the 2014-2015 school year with a standardized test titled Partnership for Assessment of Readiness for College and Career (PARCC), which claims to be directly tied to the Common Core Curriculum Standards (NJDOE, 2014). Language arts and mathematics teachers in Grades 3-8 currently in the 2017-2018 school year have 30% of their annual evaluation tied to test data. This provides each of these teachers with a Student Growth Percentile (SGP) and holds them accountable for student growth on language arts and mathematics scores based on the previous year’s results (NJDOE, 2015).

Another issue facing New Jersey public schools is the change from a paper and pencil testing delivery to a computer-based delivery. Opponents of testing argue that it is unfair to evaluate teachers on test results when computer literacy is now a factor in students proceeding properly and thoroughly through a standardized test. Proponents argue that a computer-based system brings us up to speed with modern technologies and allows for faster, more accurate, data to be produced. Throughout this research, I also attempted to address perceptions of middle school language arts and mathematics teachers on the new testing approach and format and whether or not instructional time is being spent on ensuring technology literacy.

The computer-based model of testing also generates other concerns with testing opponents. One argument being made and spoken of briefly in the previous paragraph is that testing preparation will now include assessing a child’s technology literacy in order to determine whether or not that child will be successful on the test. This is a factor completely outside the
scope of assessing for true skill-based proficiency levels, thus having the potential to impact teachers’ perceptions of standardized testing. Opponents of testing also argue that this forces schools to purchase technology devices and bandwidth infrastructure to support hundreds of students using web-based software at the same time. They argue that this adds one more pre-determining factor to the equation that was never introduced, which aligns significantly with socioeconomic status and a child’s ability to be privileged with technology devices in the home.

There is an overall claim by testing opponents that teachers are becoming more and more overwhelmed with the culture of standardized testing. Data from these tests are driving major decisions on curriculum and effectiveness of teachers. The argument from testing opponents is that it is impossible to garner a desired percentage of effectiveness with the overall goal of holding educators accountable for determining success or failure in instructional practices. Proponents of testing counter that argument by claiming that we need an objective, measurable outcome to prove whether or not a child is making academic gains in basic skill acquisition (Phelps, 2005).

This study is important because educators are getting little say in what determines the effectiveness or ineffectiveness of teachers and schools (NCLB, 2001). Legislators are making determinations and setting mandates for public schools without input from those who work with children on a daily basis. There is also a strong lack of research from the perspective of reading, writing, and math teachers as to the influence of standardized assessment data on classroom instruction. This research set out to determine whether or not educators in New Jersey are content with the mandate of using growth percentages in standardized assessment scores to partially weigh into their summative evaluations and if this process assists with instructional goals and output. It further expands on the notion of utilizing a computer-based approach to test
students and whether or not instructional time is being taken away from reading, writing, and mathematics skills instruction to ensure technological skills are proficient for the testing format.

**Conceptual Framework**

Political initiatives involving educational reform such as NCLB, Race to the Top, Common Core, and ESSA intend to add layers of accountability to all educators. The attempt at increasing accountability is connected to standardized test growth for students in language arts and mathematics. In all respects, a change theory is being implemented to alter the methods by which teachers and administrators are evaluated. A change theory tells us that we will look at the end result of a particular organizational initiative and work our way backwards to develop strategies to improve that end result (Creswell, 2014). In regards to this research, a change theory is being implemented by placing value on standardized assessment results as related to teacher effectiveness. Outcomes of students will provide us clear evidence of whether or not a teacher is effective in the classroom. Working backwards in a change theory, standardized assessment skills are identified as vital towards student preparedness and success and focused on in classroom instruction.

A conceptual change theory suggests that an organizational structure will find change through the attitudes, beliefs, and behaviors of its human subjects (Rea-Ramirez & Ramirez, 2017). Rea-Ramirez (1998) further describes a conceptual change theory as the restructuring of an organization’s existing concepts. Within the structure of public schools today, effectiveness is being tied to standardized assessment results in reading, writing, and math; and major decisions related to teacher retention, purchasing of resources, and classifying students’ abilities are being structured towards testing results. Lundholm and Davies (2013) tell us that change
theories in social sciences are under-utilized in research and demonstrate fragmented results in data.

**Research Questions**

1. What effect, if any, has standardized testing had on scripted curriculum and the pacing of it towards standardized assessment preparation?

2. What effect, if any, has the high stakes testing culture had on decisions of content delivery?

3. What effect, if any, has the shift towards computerized assessments had on the preparation of students for standardized assessments?

**Design and Methodology**

The framework for this research will be qualitative in order to gain insight into the perceptions of middle school language arts and mathematics teachers working in New Jersey’s public schools on the impact of a high stakes testing culture. Middle School language arts and mathematics teachers were specifically targeted because their classroom performance is directly being correlated to student test scores as dictated in Achieve N.J. (NJDOE, 2013). In the 2017-2018 school year, 30% of New Jersey middle school language arts and mathematics teachers’ year-end, summative evaluations are directly correlated to standardized testing proficiency levels of all students with whom they come in contact for at least six months of the school year (NJDOE, 2013).

The research questions were developed from common themes explored in the literature review. In order to properly address the research questions, 12 middle school language arts and mathematics teachers were interviewed. The interviewed teachers were chosen from three New Jersey public middle schools where the mandate exists to link 30% of language arts and
mathematics teachers’ year-end, summative evaluations to standardized test growth. The three schools use the PARCC assessment and utilize a curriculum based on New Jersey’s Student Learning Standards (NJSLS), which are a direct correlation to the Common Core Curriculum Standards (NJDOE, 2016). Utilizing a convenience, nonrandom sampling technique (Creswell, 2014), I gained qualitative data on the influence of the high-stakes testing culture on the classroom practice and preparation of middle school math and language arts teachers. Through this design, I garnered the perceptions of middle school language arts and mathematics teachers on the notions of tying 30% of their summative evaluations to test scores and the effects this has on the development and delivery of curriculum.

Prior to the search for reliable participating subjects, I contacted three school superintendents from school districts fitting into the aforementioned criteria. After receiving approval to conduct the research in their school systems, I contacted the school principals to garner contact information for teachers who meet the criteria and are currently teaching language arts and/or mathematics to students in Grades 7 and/or 8. An initial invitation to participate was sent to all of the teachers identified via professional e-mail. Those who replied and accepted the invitation were selected based on seniority of experience. There was an intention of creating an equal amount of math and language arts teachers to be interviewed. This did occur and six mathematics and six language arts teachers from the three public middle schools accepted the researcher’s invitations. To create more reliability in the research, two language arts teachers and two mathematics teachers were selected from each of the three middle schools. Their qualifications certified them as teachers who have worked under two systems of standardized assessment, which are identified as paper and pencil assessments and computerized assessments. Teachers identified and selected received an official letter of invitation for this study via their
school issued e-mail addresses. After the participants officially accepted the invitation, dates, times, and locations for the interviews were selected at the participants’ choosing and convenience.

Open-ended interviews were used to collect all qualitative data from the teachers. The questions chosen were intended to have the teachers speak openly on the effects of testing in relation to curriculum, instructional preparation, and the computerized format. Follow-up questions were used within the structure of the determined questions to create further clarity and understanding of responses. Three experienced middle school math and language arts teachers from a fourth chosen middle school with a minimum of seven years in a middle school language arts or mathematics classroom were asked to form a jury of professional educators to review the interview questions for validity and appropriateness in relation to the overriding topic of testing cultures in our schools. They were provided mock interview questions and asked to comment on the validity of the questions and to provide recommendations for increasing any validity.

Once the interviews were completed, I used a coding system to deliver qualitative data based on theory, research, and teacher perceptions (Boyatzis, 1998). I then created a summary to report common themes and compared them with all literature and theories “to confirm past information or diverge from it” (Creswell, 2003, p. 195).

**Significance of the Study**

Research for this divisive topic in education is important for several reasons. Primarily, language arts and mathematics teachers in New Jersey are now being held accountable in their annual evaluations for student growth on standardized tests to the value of 30%. Primarily, this research explored whether or not teachers consider this to be a valid and fair means of assessing professional performance and what, if any, effect this has on curriculum development and the
delivery of instruction. Second, this research explores perceptions of mathematics and language arts teachers on standardized assessments being administered to students through computer-based software. Determinations were made on whether or not resources such as instructional time and instructional preparation are being used to teach towards technology literacy in order to assist students in demonstrating annual growth on standardized assessments.

Data are required to create a structured approach towards addressing any changes that a culture of testing may create in a middle school math and language arts classroom. I have explored whether or not there is any empirical evidence to indicate that a culture of testing creates more motivated and better prepared teachers. This study distinguishes itself from others by focusing on the perceptions of middle school language arts and mathematics teachers on high-stakes assessments affecting their summative evaluations by 30%. Politicians and education administrators who provide the task of continuing this level of standardized testing rigor have the potential to look at this study and determine whether or not a culture of high-stakes testing provides us with the necessary data to prove the effectiveness of teachers in the classroom. This study also explores whether or not the preparation, instruction, and purchasing of instructional resources are a means to teach to the curriculum standards or to the standardized test format directly.

**Delimitations and Limitations**

Qualitative data were garnered from three northern New Jersey public school districts that are similar in size and demographics. Each middle school chosen currently utilizes New Jersey’s Student Learning Standards and administers PARCC to assess students annually in language arts and math in Grades 7 and 8. This study looks only at the perceptions of 12 teachers affected by changes to the New Jersey law regarding teacher evaluation (NJDOE, 2013)
and the significance of student performance on standardized tests being tied to those evaluations. Middle school language arts and mathematics teachers are the only chosen subjects since they are the only two content areas in Grades 3-8 tying high-stakes test scores to year-end, summative evaluations. The findings for this study cannot be assumed for school districts outside of New Jersey. Similarly, no findings can be related to any grade levels other than 7 or 8.

The results of this research apply only to similar themes arising from the teacher interviews. No specific PARCC data were garnered to correlate with or support the problem. The data gathered for this research are also a “snapshot” of time, documenting the transformation of evaluative systems based on student test scores in the 2017-2018 school year. It is expected that all participants in this study responded honestly about their perceptions of standardized testing effects.

Chapter Summary

NCLB (2001) placed extensively heavy emphasis on test data. The intent of the data garnered was used to determine effectiveness of educators without taking into consideration some of the pre-determining factors of a child’s academic realities such as socioeconomic status and English proficiency of English Language Learners. The culture of testing has progressed rapidly with the introduction of Race to the Top (2009), Common Core (2010), and ESSA (2012) with more accountability of teachers being placed on standardized test results. Teachers are now being evaluated based on testing data, and school-based decisions are being influenced by test results in regard to teacher effectiveness and development of curriculum.

Maylone (2002) indicated through her research that 56% of all high school standardized test scores in Michigan were pre-determined by elements such as race, socioeconomic status, and English language proficiency. Turnamian (2012) referenced Maylone’s 2002 study to make the
determination that demographic data had tremendous value in determining the results from standardized assessment data in New Jersey’s public schools. Policy makers, attempting to create accountability based on a child’s academic growth, argue that there needs to be an objective, numerical system in place to measure skill growth. Standardized assessments support this objective measurement and provide data to further drill down into strengths and weaknesses of students.
CHAPTER 2
LITERATURE REVIEW

Historical Perspective

High-stakes testing has become an integral component of assessing the efficacy of K-12 education in America. High-stakes testing denotes an annual standardized assessment that carries serious consequences for students, educators, and school districts (Urrieta, 2004). It is argued by advocates that high-stakes testing needs to be a part of diagnosing student strengths and weaknesses and a strong, objective manner in which we can measure progress in a child’s education. No Child Left Behind (2001) made the assertion that testing accountability placed on schools, teachers, and students, leads to significant improvements in the skills required to deem a child proficient in reading, writing, and mathematics. It is argued by testing opponents that schools are now spending too much time and too many resources preparing students for standardized tests while ignoring other important components of education that make a school and its students successful (Blazer, 2011).

The Eight-Year Study

The historical timeline of where we are today with high-stakes testing can begin with The Eight Year Study, ironically conducted over 12 years from 1930 to 1942 by the Progressive Education Association. The Eight Year Study found that students who attended the “progressive schools” performed comparably well academically and were more heavily involved in cultural activities. The Progressive Education Association determined that graduates from the progressive schools did not experience any deficiencies while attending college and that these students performed, at times, better than their traditionally prepared peers. A common curriculum was developed by 30 high schools in order to bring together academic skills and
content in science, mathematics, social studies, arts, and language. The only way for this to happen would be to ensure that all students were measured on the same scale. The staff placed in charge of developing the curriculum also had to decide what areas of the curriculum they needed to evaluate. They also sought to gather evidence and use the evidence to track progress or lack thereof. This can certainly be viewed as an essential piece of the high-stakes testing puzzle. Teachers, within this study, focused on gathering data and used data to make decisions regarding a college preparatory curriculum (Aikin, 1942).

**A Nation at Risk**

In April 1983, *A Nation at Risk: The Imperative for Educational Reform* was published by The National Commission on Excellence in Education (U.S. Department of Education, 1983). The intent of this report was to assess the current state of the American education system and to provide solutions to the problems facing American education. Six main charges of primary focus were the following:

1. Assessing the quality of teaching and learning in our schools
2. Comparing American schools with school systems of other advanced nations
3. Comparing the relationship between college admission requirements and college success
4. Identifying K-12 programs that assist students with college
5. Assessing how social and educational changes have affected student outcomes
6. Defining any problems and offering solutions to those problems in American schools

Much of the information being sought requires a measurement of data and data analysis. Student growth and progress can objectively be measured through standardized test scores or classroom grades. *A Nation at Risk* used data as its key premise in making the determination that
American schools are lagging compared to other industrialized nations. Under the section of the report titled “Indicators of the Risk,” all data cited to promote schools at risk are standardized test scores such as the SAT (U.S. Department of Education, 1983). Objective data were essentially being used to prove a point and to measure achievement.

No Child Left Behind

In 2002, the No Child Left Behind Act was signed into law by President George W. Bush. This law was a direct reauthorization of the Elementary and Secondary Education Act, which was signed into legislation in 1965 by President Lyndon B. Johnson (Elementary and Secondary Education Act, 1965). No Child Left Behind directly targeted the disadvantaged students and created annual accountability standards to hold schools more accountable for failing test scores (No Child Left Behind, 2001). This was a direct influence of the Elementary and Secondary Education Act of 1965, which developed such programs as Title 1 to target students in poverty (ESEA, 1965).

According to the No Child Left Behind Act, the following standards were used as measuring sticks to determine effective or failing schools:

- By 2006, states were required to test students from Grades 3-8 annually in reading and mathematics.
- By 2008, all students had to be tested once in elementary, middle, and high school on a standardized science assessment.
- All tests had to be aligned to state curriculum standards.
- By 2014, 100% of students had to be determined “proficient” on reading and mathematics standardized tests.
No Child Left Behind attempted to create accountability within schools through the use of standardized test score comparisons. Specifically within New Jersey, District Factor Groups were developed in 1975 and used to determine if a particular school was on par with other schools possessing similar socio-economic and demographic communities (NJDOE, 1975).

The outreach of No Child Left Behind, however, left curriculum and testing standards solely in the hands of state and local education agencies. High-stakes testing comparisons from state to state, prior to Common Core Curriculum Standards (2010), were largely skewed with the use of different standardized test formats and different levels of rigor within the curricula. Only tests such as the SAT and ACT were able to create a standardized playing field for students across the nation. However, these tests are not usually taken until the latter part of a student’s high school career. This idea of standardization was sent down the grade levels in order to create a fairer comparison table for schools across states.

Race to the Top

Race to the Top came along in 2009 to provide a more farsighted framework for curriculum and testing across all schools in America. States were provided the opportunity to compete for federal funds if they complied with federal education initiatives such as adopting national K-12 curriculum standards (U.S. Department of Education, 2009). These standards later became known as the Common Core Curriculum Standards. The Race to the Top Executive Summary also called for more standardized methods of evaluating teachers and administrators. Evidence-based approaches were adopted, and this marked a new path for teachers and administrators to be evaluated based on student growth on standardized tests. The intent was to close achievement gaps through objective, measurable methods and to use standardized testing data to make decisions about curriculum and programs (U.S. Department of Education, 2009).
The Eight-Year Study, A Nation at Risk, No Child Left Behind, Race to the Top, and the Every Student Succeeds Act create a historical perspective into the role and state of standardized testing in American schools today. Measurable accountability for teachers and administrators is the common theme that surfaces in each agenda. As the nation progressed after World War II, standardized testing rose as the most objective method of demonstrating growth and making comparisons amongst similar school communities. Accountability also grew for teachers, who are now evaluated based on their students’ growth on reading and mathematics tests. High-stakes testing is now more prevalent than ever and is used more often to make decisions for curriculum and programs in schools (Casbarro, 2005).

**Existing Literature Regarding High-Stakes Testing**

After reviewing the literature regarding high-stakes testing and its effects on students and teachers, it is apparent that standardized testing from Grades 3 to 8 has intensified after the implementation of No Child Left Behind. Much of the research read depicts the over-reliance of high-stakes testing as detrimental to schools in America. It is hard to argue that many school agendas are focused on test preparation. High-stakes testing and the act of attaching important consequences to standardized test results is what drove the No Child Left behind Act (Nichols & Berliner, 2008). Essentially, under No Child Left Behind, the threat of punishment would cause the teachers to work more effectively, thus creating more motivated students (Nichols & Berliner, 2008). Under No Child Left Behind, accountability became more objective and brought more evidence-based methods of evaluating teachers to schools (Danielson, 2011). This became a simple method of rooting out the “bad teachers” and applying a business model of productivity to ineffective schools (Nichols & Berliner, 2008).
Six Effects of High-Stakes Testing

Six major effects of high-stakes testing on K-12 schools and their systems consistently arise when searching for literature on the topic. The six effects all claim to have a negative influence on the direction towards measuring student achievement and the outcomes of school agendas. The negative effects of standardized testing in schools are as follows:

1. Curriculum and teacher effectiveness suffer
2. The inability to prove the validity of standardized tests
3. School funding and the costs of standardized testing are exorbitant
4. The surrounding political agenda pushes testing, not educational reform
5. The correlation of standardized testing and high school drop-out rates
6. School culture suffers as a result of testing

Curriculum and Teacher Effectiveness

One claim states that research-based instructional approaches are lost when educators have to spend time preparing for and worrying about high-stakes tests (Santman, 2002). The intent of high-stakes testing is to create a “top-down” method of driving instruction through standardized test data and, ultimately, to use these data for curriculum reform. Smith (1991) expressed her predictions for the negative impact of large-scale, high-stakes testing on schools and curriculum. She stated that a high-stakes testing culture limits curriculum to only the tasks to be mastered on the test. Curriculum decisions in design align state standards with the required standardized assessment (Bhola, Impara, & Buckendahl, 2003). Major decisions are being made based on one-time, high-stakes tests that affect the path of educational standards and curriculum for teachers and students. High-stakes testing now drives curriculum as teachers are reorganizing their lesson plans to reflect only those skills required on the standardized tests.
In 2011, Christie Blazer, supervisor in the Office of Assessment, Research, and Data Analysis for the Miami-Dade County Public Schools, wrote an article titled “Unintended Consequences of High-Stakes Testing.” In that article, she explained that high-stakes assessments narrow the curriculum and change what is taught. This, according to Blazer, negatively affects the quality of the classroom instruction. She argued that the greater the stakes, the narrower the curriculum becomes. Within this article, there were four overall categories that she organized and better described the narrowing of the curriculum in the Miami-Dade Public Schools. The first argument she made was that high-stakes testing excludes accountability for non-tested subject area teachers. Nichols and Berliner (2008) explain that subjects such as art, world languages, music, and physical education are being replaced by more practice in language arts, math, and science.

The second argument made by Blazer (2011) is that high-stakes testing creates an exclusion of non-tested topics within subject areas. What she found was that we are teaching to the low-level skills required to move up grade levels while not focusing on the more challenging aspects a curriculum has to offer. Barnes (2005) argues that high-stakes testing lends curriculum to focus only on what will appear on the test and does not give any regard to other vital bits of information that students should be learning. Blazer (2011) argues that this is not an appropriate pace of curriculum sequence for many of our students.

The third effect of standardized assessments on curriculum as explained by Blazer (2011) focuses on teachers having to adapt their teaching styles to the formats of the tests. She contends that teachers are practicing the art of repetition for only one or two isolated pieces of information that students will see on the standardized tests. She also makes the statement that instructional
practices such as cooperative learning, creative projects, and field trips are being abandoned for more traditional, lecture-style approaches to delivering information. The purpose, according to Blazer (2011), is solely to prepare students for high-stakes tests.

The fourth effect of testing on curriculum, according to Blazer (2011), is the notion that we are spending an extremely large amount of classroom time preparing students for a test. She argues that this creates less time for students to truly learn and grasp new material that will help them build on knowledge and skills. Most of the classroom time for students, according to Blazer (2011), is being spent in isolation reviewing sample test questions and understanding test formats. All of this, Blazer (2011) argues, leads to student boredom and burnout.

So, what to do?

For school administrators, teachers, and those given the charge of designing and implementing curriculum, this is a major directional issue for classroom teaching. Teachers and administrators must now decide what will be taught in the classrooms and make those determinations with the prospect of preparing students for a standardized, high-stakes test. Teachers must now focus on the repetition of instruction and resources designed to prepare students for these tests, which will lead to boredom and burnout.

**Inability to Prove the Validity of Standardized Tests**

A second supporting argument for the negative consequences discrepancy between individual state standards and the standards found on standardized tests. According to No Child Left Behind, the instruments in which states used to measure academic standards were to be set within the individual states (U.S. Department of Education, 2009). The states’ curriculum standards varied, but the tests being administered were similar in mastery skills. Hillard (2000) also found that standardized tests were narrowly focused and only covered lower levels of
thinking. Tests were often multiple-choice throughout with one or two narrative writing prompts. Some tests, such as the SAT, were multiple-choice only. Kohn (2000) stated that standardized tests are not a reliable method of measuring the potential of a student because instruction becomes limited and teaching methods become routine and structured.

The intent of high-stakes standardized assessments today is to determine the ability levels of students at a particular grade level. As much of the literature in this review suggests, the results of the test are often used to judge whether or not students are capable of advancing to the next grade or skill level and whether or not schools and teachers are doing their jobs properly in order to prepare students for proficiency in reading, writing, and mathematics. However, the validity of these tests and the performances of students and teachers come into controversy when we look closely and realize that, oftentimes, the tests being used do not align with the curricular outcomes prescribed by local school districts (Hornof, 2008). The United States Department of Education (2009), under No Child Left Behind, created a national focus of student achievement for language arts, mathematics, and science; however, each state was left to determine its own curriculum standards. The validity of creating a unified accountability system with different methods of achieving that accountability is suggested as being questionable (Amrein-Beardsley, 2009).

The issue of validity is also called into question when research-based instructional decisions are solely being made based on standardized test scores (Santman, 2002). Santman (2002) argued that the proficiency levels prescribed by state departments of education do not match realistic proficiency levels required for real-world application. Sloane and Kelly (2003) claim that more authentic tests which contain more open-ended questions and a grade-appropriate scoring rubric might give us more of an indicator of a child’s future success.
However, they also make the claim that these tests are more costly, time consuming, and take much more time to develop and design.

With this taken into consideration, are standardized tests still a valid means of measuring a student’s progress and predicting a level of success in the future? Nelson and Eddy (2008) argue that no single test is capable of giving us any of this information. They go on to say that high-stakes tests cannot give us guidance for curriculum or student deficiencies since they are only a small portion of what is learned throughout an entire school year. Smith (2009) tells us that high-stakes testing leads to “disjointed” learning in the classrooms and that the validity and reliability of high-stakes test results leave teachers and students questioning the amount of time spent in preparation. If the intent of high-stakes testing is to put pressure on schools in the form of consequences for the lack of progress or advancement of progress in student performance (NCLB, 2001), we are responsible for ensuring that those tests are authentic and valid (Kantor & Lowe, 2006).

In 1991, Paris, Lawton, Turner, and Roth concluded through survey results that high school students did not feel there was a realistic correlation between proficient test scores and academic ability. In 2002, Hughes and Bailey conducted a study in Indiana that indicated high school students did not show a high level of concern for standardized tests and did not display large anxieties over passing or failing a high school exit exam. Many of the students in the Hughes and Bailey (2002) survey felt it was unfair to correlate a passing test score with the ability to graduate on time. The surveys reflected students’ perceptions that there was not a lot of motivation to pass the test and that many students did not take the test seriously, often guessing at answers and not being on task for the entire time allotted to take the test.

So, what to do?
If teachers and students are to buy into to the culture of standardized testing, there must be a valid means of ensuring the worth of predicting outcomes for students and accountability of teachers. If the outcomes provide any level of invalidity, attitudes towards the initiative will not be positive.

**School Funding and the Costs of Standardized Testing**

In 2004, Lawrence Baines and Gregory Kent Stanley wrote an article titled “High-Stakes Hustle: Public Schools and the New Billion Dollar Accountability.” In that article, they argued that high-stakes testing costs American schools up to $50 billion per year. They make the claim that high-stakes testing costs in our nation rival the gross national products of small countries. They go so far as to calculate per pupil costs for high-stakes tests by making the claim the 5.5% to 14% of every dollar spent for public schools goes toward the testing initiative. Baines and Stanley (2004) also raised the argument that high-stakes testing now labels schools “in need of improvement.” Their focus for this statement was centered on New York City’s public schools. The financial task at hand in order to raise schools to an acceptable level of achievement based on test scores is not included in the $50 billion claim. They stated that in New York City alone, this task of bringing schools to an acceptable level of proficiency as dictated by a governmental bureaucracy will cost public school budgets an additional $25 million.

The claim of costs being exorbitant is supported by their assertion that the money spent on high-stakes testing can be put to better use by addressing issues such as crumbling infrastructure of schools and resources to accompany the current curriculum standards in subject areas other than language arts and math. Nelson, McGhee, Meno, and Slater (2007) made the argument that schools are spending almost all of their curriculum budgets on test preparation materials. According to Gentry (2006), programs supporting gifted and talented students, as well
as programs in the arts and sciences, are being eliminated due to the monetary shift towards preparing students to perform well on high-stakes tests.

So, what to do?

Most of the funding for America’s public schools comes from local taxpayer money. Taxpayers are not too happy when they find that their dollars are not being spent on proven methods of success. Also, school administrators must decide where to allocate their budgetary funds designated for curriculum and instruction. Initiatives such as technology in the classroom and project-based learning approaches will be pushed aside if funds are not available to accomplish the goals. These goals must heavily include teacher preparation and follow-up instruction for teachers in order for the proper outcomes to be gained. Training for teachers is vital for any program to succeed; however, teacher training is often costly and demands a large investment of time.

**Political Agendas of Standardized Testing**

There is also a strong claim within the literature of a political agenda surrounding standardized testing. This is by far the largest controversy surrounding high-stakes testing. The two major questions on both sides of the issue are as follows:

1. Con: “Do politicians and school leaders use standardized testing to push their political agendas in schools?”
2. Pro: “How else can we as leaders hold schools, teachers, and school leaders accountable for continuing to work towards raising student achievement levels?”

The most important effect of high-stakes testing has been on teacher accountability and evaluation. Grant Wiggins (1989) recognized a valid use of standardized tests; however, he reminded us that test developers warned against the use of one test to make major decisions...
regarding student performance assessments. Wiggins also saw standardized tests as a means for politicians to involve their agendas into public school decisions. A solid, objective score was a means of providing evidence of a school’s success or failure. Politicians could now use this objective number to support their own campaign or to use it against their opponents. The culture of high-stakes, standardized testing was now embedded into politics.

In 2004, Luis Urrieta, Jr. wrote an article that essentially labeled the high-stakes testing culture as Assistencialism. In this article, he defines assistencialism as the mindset that people need assistance, and it is our responsibility as a society to assist individuals in need. The article asserts that we assume to know who needs help and who doesn’t by simple observations and not by delving deeply into the problem or concern. He places high-stakes testing into this category by claiming that there is a political perception that high-stakes tests are the equalizers that will create a system of equality and give everyone the same opportunities to be educated in an effective school. According to Urrieta (2004), the politics comes into play with the implementation of serious consequences for schools that are not meeting proficiency standards as set by non-educators. Politicians were responsible for setting those standards with no evidence of how they would affect schools not meeting compliance. Urrieta (2004) states that some of those repercussions involved state department of education takeovers and conversions of public schools to charter schools, which is a major political issue in our nation’s urban, inner-city schools. For teachers, according to Urrieta (2004) in the same article, low test scores can mean losing their jobs. For students, he argues, low test scores can mean placement into special education programs or placement into lower ability tracks.

One of the larger political drawbacks, according to Urrieta (2004), was the burden placed on English Language Learners. Under NCLB, this subgroup of students was expected to be on
par with American-born students in the areas of language arts and mathematics regardless of their English acquisition level. Urrieta (2004) tells us that this created a major political battle as sensitivities for English Language Learners were challenged. The great language debate in America became evident in public schools, and teachers, students, and school leaders were being held accountable for English Language Learners whose English acquisition was not sufficiently developed.

In 2005, Joseph Casbarro wrote an article titled “The Politics of High Stakes Testing.” The opening paragraph of that article reads as follows:

A troubling reality in today’s political climate is that many political leaders actually believe that the best way to change schools is through an “end of a gun barrel” approach, rather than by building consensus.

Within the article, Casbarro (2005) goes on to explain that high-stakes testing is a means of creating accountability as “prescribed by No Child Left Behind.” His argument is that the intent of high-stakes testing is designed to measure student progress. However, the accountability issue only monitors the overall state of a school and not student progress. His claim is that we are missing the mark on the intended outcomes.

In the same article, Casbarro (2005) raises three major points regarding the state of testing in America and the politics it involves. The first point he makes is that high-stakes testing is a means of political coercion. He defines this coercion as coming in the form of both positive and negative coercion. The positive effects of the coercion consist of financial incentives (merit pay for teachers) and public recognition (school report cards and wide-range media reports). The negative coercion comes in the forms of threats regarding grade retention, loss of funding, and poor teacher evaluations. He argues that high-stakes tests are a “one-size-
fits-all remedy” that does not take into consideration key elements such as English proficiency of the students taking the test.

The second form of coercion described by Casbarro (2005) comes in the form of the politics of performance. He points to the fact that state departments of education have now developed specific bureaucratic agencies to develop tests and to analyze test results. They are put in charge of establishing benchmarks for student performance and judge schools publicly for the lack of progress in student growth. This brings about the point made earlier by Urrieta (2004) that allows non-educators to set a target score that determines the proficient skill level at each grade level for the students and teachers. This is described by Casbarro (2005) as completely flawed when a state department of education can raise or lower expectations based on how the students fared on the test in previous years. In other words, if the students are consistently failing, we will lower the proficiency level.

The third point Casbarro (2005) makes regarding politics in high-stakes testing is that of public perception. He states that many political campaigns focus on the theme of creating better performing schools. However, when politicians make this campaign promise, they are on the topic of continuing to test our students as the only viable means of determining growth. Casbarro (2005) claims that standardized tests are perceived by the public as a viable means of measuring the ability of our students from year to year. He argues that communities compare test scores of their respective students to the worth of comparable school systems.

When bringing all of this together, Casbarro (2005) argues that standardized testing has its place in schools; however, it should not be the only means of evaluating schools, teachers, and school leaders. He states that we only look at the final test score and do not evaluate what that score really means. This is a question that involves so many key factors, such as the socio-
economic status of the community, the rigor of the curriculum, the difficulty of the test, and how much preparation was put into preparing for the test. Casbarro (2005) makes the point that tests do not determine the overall skill level of a child if that child spent months practicing and preparing for that specific test. The results will only show that the student became a good test taker, not that he or she has the skills necessary for proficiency in reading, writing, and mathematics.

So, what to do?

If there is significant research proving the negative consequences of high-stakes testing, why are we continuing to be so over-reliant on the outcomes measured? The politics of school performance often embeds itself into political campaigns and promises of “fixing schools that are broken.” No Child Left Behind (2001) and Race to the Top (2009) are prime examples of testing standards being placed into political agendas. The costs of operating public schools is also observed by many as being overpriced and full of unnecessary spending. Politicians usually have an agenda of saving taxpayers money and cutting property taxes at the local level. What better way of getting elected to public office than by making a campaign promise to lower taxes and raise property values with better schools?

**School Culture as a Result of Standardized Testing**

Negative effects on school culture have also been a result of high-stakes testing. High-stakes tests have raised the level of accountability amongst teachers and school administrators under No Child Left Behind. No Child Left Behind and Race to the Top attached merit pay to teacher successes with raising test scores. High-stakes testing created pressure for schools to perform well on their respective standardized tests with an approach of attaching negative
consequences to low test performance (Amrein-Beardsley, 2009). As a result, accountability increased when high-stakes test results started getting tied to teachers’ evaluations.

In 2007, Arthur T. Costigan and Margaret Smith Crocco wrote about our ability to retain young teachers, particularly in urban areas. They explain that in New York City’s schools, specifically, there is now a culture of high-stakes teaching to a standardized test. This comes in the form, according to Costigan and Crocco (2007), of scripted lessons and strictly mandated curriculum. The essay provides several interviews with novice New York City teachers in English and social studies to determine their levels of frustration and anxiety over the idea that their evaluations and classroom performances will be judged solely on a snapshot of time. The essay concluded with the presumption that good, young teachers are being dissuaded from continuing with their jobs in urban, public schools due to the enormous pressure of high-stakes tests and the inability to have any instructional freedom. The essay makes the assumption that this is a common trend amongst all urban, public schools in America.

Advocates of high-stakes testing argue that testing places necessary pressure on schools to encourage students to work harder and for teachers to place greater emphasis on student growth and performance (Blazer, 2011). Critics argue that high-stakes testing creates a culture in schools that places the primary purpose of learning on the preparedness of the chosen test (Blazer, 2011). Researchers have also found that a culture of testing, overall, creates students who are less motivated in the classroom, less motivated to perform well on standardized tests, and have negative perceptions of curriculum that has a primary focus of preparing for a test (Jerald, 2006).

So, what to do?
School leaders believe that the culture and climate of a school has a direct correlation with student and teacher outcomes. If people feel positively about the building they enter each day, they will put more of an effort into what is being taught in the classrooms. If people feel negatively about the building they enter each day, they will not be motivated to perform to appropriate and necessary outcomes. This centers on basic tenets of human attitudes. If students and teachers are not given any level of academic freedom and consistently repeat the same material over the course of a school year, teachers and students will not be motivated to grow professionally and/or academically.

**Correlation of Standardized Testing and Dropout Rates**

Data have proven an unintended consequence of high-stakes testing, which details the phenomena of students dropping out due to the pressures and anxieties of being properly prepared for testing proficiency. The idea of high-stakes testing being correlated to rising dropout rates makes an extremely strong connection since the inception of No Child Left Behind. Under No Child Left Behind (2001), a school must meet Annual Yearly Progress (AYP) or it will be placed in a category of a “school in need of improvement.” If a school fails to meet Annual Yearly Progress for two consecutive years, parents have the right to school choice, which allows them to move their child to a school that meets Annual Yearly Progress. If that school continues to not meet Annual Yearly Progress, according to No Child Left Behind (2001), the school can be subject to restructuring and corrective action. That action can include the termination of school leaders and a poor evaluation for a teacher. If Annual Yearly Progress is a districtwide concern, a state department of education can essentially take over the school district and its daily operations.
Out of common sense, a school that is not performing to proficiency levels as deemed appropriate by department of education officials will target those students whose scores reflect such levels. Amrein and Berliner (2003) argue that strategies schools use are suspension, expulsion, and reclassification in order to have these students avoid taking the test. Bushweller (2004) claimed that students are being encouraged to drop out or transfer to different schools if their test scores reflect poorly upon the school and are a cause for not meeting Annual Yearly Progress. There are claims that schools use the tactics of “pushing students out” (Viadero, 2004) and ask low-performing students to stay home on the days the exams are administered (Jones, Jones, & Hargrove, 2003).

Anxiety levels of students and teachers regarding high-stakes testing are also a major factor in rising dropout rates (Jones, Jones, & Hargrove, 2003). According to Viadero (2005), there is a distinct connection between states requiring high school exit exams and the drop-out rates in those states. Viadero (2005) argues that states requiring high school exit exams also have lower SAT scores and lower graduation rates. Amrein and Berliner (2003) make the claim that 88% of states requiring high school exit exams report higher drop-out rates amongst high school students as compared to those states not requiring high school exit exams. Amrein and Berliner (2003) also tell us that students under the age of 20 taking the GED, an alternate program to achieving high school graduation, rose 73% between the years of 1986 and 1999. Stress, anxiety, boredom, and fear are all listed factors that prevent students from completing high school (Amrein & Berliner, 2003).

We can speculate on many causes of drop-out rates rising; however, one theory suggests that the high-stakes testing culture has forced schools to adopt programs catering only to the college preparatory track, which essentially eliminates any vocational training and career
education for high school students (Waits, Setzer, & Lewis, 2005). According to Monty Neill (2006), studies support the correlation between high-stakes testing and dropout rates, pushing approximately 40,000 high school students out the door of American public schools each year. As mentioned in No Child Left Behind (2001), student retention is a defined consequence for students not performing to levels of proficiency on high-stakes tests. Two grade retentions in a child’s K-12 school career create a 90% possibility that a child will drop out of school (Amrein and Berliner, 2003). If a child is retained in Grade 9, the chances of that child graduating from high school are extremely low (Viadero, 2005). Viadero (2005) takes this a step further and tells us that minority students and students with learning disabilities, limited English proficiency, and low socioeconomic status are more likely to receive low test scores and likely to drop out of high school.

So, what to do?

The purpose of a K-12 educational system should be to move students through the entire system and have them graduate in Grade 12 prepared for the next level of college or the work force. If, as the claim was made by Neill (2006), 40,000 students drop out each year in America, we are failing those students and creating a social system of individuals reliant on the government for support. K-12 administrators must find a means of supporting these students while existing in a culture of high-stakes, high-anxiety testing.

**The Positive Culture of High-Stakes Testing**

Several resources were found regarding the positive influences of high-stakes testing. However, the research is scant and limited to opinion articles, surveys, and polls. One such poll was taken in California in 2013 and the results were explained in an article written by Sharon Noguchi on September 4, 2013, in the *Contra Costa Times*. The poll showed that California
voters strongly support more state standardized testing as a way to measure the academic progress of children in reading, writing, and mathematics. The poll was taken by the Rossier School at the University of Southern California with Assistant Professor Morgan Polikoff as the lead for this project. The poll sampled 1,001 voters in English and Spanish online between August 27 and August 30, 2013. The results of the poll were as follows:

- 43% said that teachers should be evaluated on students’ standardized test performance.
- Two-thirds of those polled said that California should test students in every grade level.
- 45.3% of respondents and 48.4% of parents of school children gave the California school system a grade of C or below.
- 50% said that California schools have gotten worse over the last 10 years.

Although this polling data shows a positive light to high-stakes testing advocates, one statistic was raised within this poll. Forty-three percent of respondents said that removing teachers from the classroom is the best remedy to making schools better. 48% blamed the teachers for non-proficient students, while only 28% blamed the parents of those students.

This poll was conducted by MFour Research and Tulchin Research for University of Southern California as a means of providing supporting data to Policy Analysis for California Education, an independent, nonpartisan research center. The poll’s margin of error had a 3.5 plus/minus margin of error.

**Review Methods**

The literature was searched through the Seton Hall University library database system. More specifically, the ProQuest and ERIC databases were utilized for most of the searching.
Journal articles were found through a general “Education” search within the library’s webpage. When accessing the library’s web page, titles for disciplines are listed. They consist of such search options as Mathematics, Science, Education, Medicine, Business, and so forth. The Education tab was clicked leading to a wide array of educational journals. The journals accessed were ARRA, Educational Researcher, and Education Weekly.

ProQuest and ERIC databases provided much of the literature used in this review. A heavy reliance was placed on dissertations and case studies found throughout these search engines.

Government websites were also visited to gain a better understanding of the historical aspects of accountability and high-stakes testing. For example, The United States Department of Education’s website offered the Race to the Top agenda and proposal.

The search words used when accessing the literature were as follows:

High-Stakes Testing Politics
High-Stakes Testing School Culture
High-Stakes Testing Drop Out
High-Stakes Testing Costs
High-Stakes Testing Funding
High-Stakes Testing Efficacy
High-Stakes Testing Teacher Evaluations
High-Stakes Testing Rewards
High-Stakes Testing Pitfalls
High-Stakes Testing Validity
High-Stakes Testing Limitations
High-Stakes Testing Attitudes
No Child Left Behind
Eight-Year Study
Race to the Top
A Nation at Risk
High-Stakes Testing Accountability
High-Stakes Testing Corruption
High-Stakes Testing Cheating
High-Stakes Testing
High-Stakes Testing Quantitative Data
High-Stakes Testing Curriculum
High-Stakes Testing Skills

Limitations of the Review

The review of the literature did not reveal any significant data or research on the positive aspects of high-stakes testing. There was an overwhelming amount of research, however, found on the negative aspects of high-stakes testing. This forced me to think critically about the positive argument for the issue and whether or not major decisions regarding high-stakes tests were being made without proper research to support the initiative. It is difficult to weigh data on two sides of an issue when the literature is heavily in favor of one position.

Another limitation of this review is that there was an over-reliance on research provided through ProQuest and ERIC. This created limitations for the types of resources found. For example, a large array of dissertations and case studies were found with qualitative data.
However, a much smaller amount of quantitative data was researched. This could have created a more significant argument for the perils of high-stakes testing.

**Criteria for Inclusion and Exclusion of Literature**

The intent of this literature review was to examine studies that significantly outlined qualitative data. Quantitative data for the correlation of high-stakes testing and drop-out rates were also examined, but the quantitative data were limited to that topic. The qualitative data were taken from ProQuest and ERIC databases, which provided a wealth of information pertaining to the negative effects and consequences of high-stakes testing. The studies of focus pertained mostly to attitudes and perceptions of students, teachers, parents, and school leaders on the effects of high-stakes tests. The data received within the chosen literature were garnered, in large part, through interviews and experiences of teachers and students preparing for high-stakes tests.

**Corruption and Cheating Associated with High-Stakes Testing**

The focus of this literature review was to look closely at empirical research where cheating and corruption were the direct effect of the pressures placed on schools, teachers, and administrators to meet levels of Annual Yearly Progress as defined by No Child Left Behind (2001). In 1975, Campbell made the bold statement, “The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor” (p. 35). According to No Child Left Behind (2001), a social culture of placing pressure to perform on high-stakes tests was placed heavily on schools, teachers, and students alike. Negative consequences under No Child Left Behind (2001) ranged from student retention to the loss of employment for school staff. The positive consequences under No Child Left Behind (2001)
ranged from merit pay for school personnel to positive media reports regarding the state of the respective school. With negative and positive consequences attached to employment status, cheating and corruption became inevitable.
CHAPTER 3

METHODOLOGY

Testing has become an annual, high-stakes rite of passage in public schools. Across Grades 3-8, language arts and math teachers in New Jersey are being evaluated and labeled as effective or ineffective partially based on high-stakes test results. The culture of high-stakes testing weighs this level of effectiveness on what amounts to three days of language arts testing and three days of mathematics testing in the school year. The testing that currently takes place in New Jersey’s public schools measures growth of students annually in Grades 3-8 in the areas of language arts and mathematics. A call for new evaluation systems of teachers based on test results has taken effect in New Jersey’s schools based on Department of Education mandates outlined in AchieveNJ (2013). The purpose of this study was to explore the perceptions of middle school teachers in New Jersey on the mandate that 30% of their summative evaluation is weighted by standardized assessment growth in language arts and mathematics. I explored perceptions of middle school language arts and mathematics teachers on the high-stakes testing environment as they designed lessons and delivered instruction to students in the classroom.

The qualitative data from this study were collected through the utilization of three research questions containing a total of eight interview questions that are open-ended in nature. Due to the nature of the qualitative approach, the interviews served as data collection opportunities. The interviews allowed the teachers to speak openly and honestly in a setting of their choice in regard to the effects of a standardized testing culture on their input into curriculum design and their individual classroom instruction. Individual interviews prevented the teachers from being influenced by their colleagues’ responses. Utilizing this method of data collection, the subjects chosen provided data comparable to other means of data collection. The
three pre-determined questions were as follows: (1) What effect, if any, has standardized testing had on your scripted curriculum and the pacing of it towards standardized assessment preparation? (2) What effect, if any, has the high-stakes testing culture had on your decisions of content delivery? (3) What effect, if any, has the shift towards computerized assessments had on the preparation of students for standardized assessments?

The design of this study was conducted through personal interviews of twelve middle school math and language arts teachers in order to examine common themes and trends (Krueger & Casey, 2000) in middle school math and language arts classrooms. Twelve middle school math and language arts teachers were selected to create factual research in a qualitative method. Each of the participants recruited and selected had at least five years of experience in a middle school math or language arts classroom. The researcher believes that the topic of effects on classroom instruction and instructional attitudes deserves recognition and study based on the amount of instructional preparation, time, and resources being allotted towards standardized testing preparation. Since the researcher is a middle school building administrator, this topic serves as interest to the attitudes of teachers in the preparation, planning, and delivery of middle school math and language arts curricula.

In order to develop appropriate and valid research questions, the researcher gathered a jury of experts to review and amend language to better reflect experience and practice. Three middle school language arts teachers and three middle school math teachers, each with at least seven years of experience in a middle school language arts or mathematics classroom, were asked to be part of a jury of experts in order to determine themes of dialogue amongst teachers related to standardized assessments and any effects on curriculum and/or instruction. Each teacher invited volunteered to participate and is employed at a fourth middle school in the same
geographic region as the researcher. A survey was sent to these teachers via e-mail. After the survey was completed, a brief meeting with the six experts took place. It was agreed upon that the most affected areas of teaching due to the high-stakes testing culture are (1) the development and design of curriculum, (2) costs to prepare students, (3) annual adjustments and revisions to curriculum, (4) effects on the creativity in developing lessons, (5) facts of having 30% of a year-end, summative evaluation tied to standardized test growth, (6) effects on school schedules and instructional time, and (7) instructional time spent practicing technical skills needed to take a computer-based assessment. These themes were broken down into three classifications that created the framework for the research questions: (1) development of curriculum, (2) delivery of curriculum, and (3) teaching towards a computer-based assessment. Face validity was verified with three opening questions as follows: “What subject do you teach? What grade level(s) do you teach? How many years have you taught this subject area to that/those grade level(s)?

Teacher responses to eight open-ended questions provided valid and systematic research that allowed the researcher to code common themes and attitudes towards the effects of standardized testing preparation. The data were collected through the question and answer dialogue between the teachers and the researcher. Recordings were made of the interviews and the responses were transcribed solely by the researcher. A descriptive manner was used to measure the data since there was no data collection device to quantitatively measure the teachers’ interviews (Krueger & Casey, 2000). The interviews took place in the teachers’ respective schools in a location of their choice. Settings with no other individual present were chosen in order to avoid any threatened attitudes or a lack of willingness to answer the questions openly and honestly. The responses provided were directly related to the teachers’ experiences in preparing students for standardized testing in math and language arts at the middle school level.
Research Design

I used a case study approach by selecting 12 middle school math and language arts teachers who have shared the experience of preparing students in Grades 7 and 8 for high-stakes standardized assessments for at least five years. This study explored combinations of factors that contributed to the outcomes of findings, concepts, hypotheses, or theories (Ragin, 1997). Since data garnered for this research project derive from perspectives based on classroom practice, open-ended interview questions were posed to each of the case-study participants with opportunities for follow-up questions to emerge. The responses to all interview questions served as the source of all data so that research was developed into the insights of participants’ interpretations of effects of high stakes testing on classroom practice (Bogdan & Biklen, 2007).

Sampling

In order to investigate how a culture of standardized assessments affects the creativity and structure of middle school math and language arts curriculum and instruction, I used a convenience-based, criterion sampling approach. In order to gather strong evidence and support of all conclusions, middle school math and language arts teachers were chosen from three different New Jersey public middle schools in close proximity to the New York City tri-state area. Approvals from each school’s superintendent and principal were obtained. After approval to conduct the research at the respective middle school, all language arts and math teachers teaching Grades 7 and/or 8 language arts or math were sent an e-mail directly by the researcher inviting each of them to participate. The respective school principals were not involved in the invitation process. Two language arts and two mathematics teachers from each of the three middle schools accepted the invitations and interviews were scheduled. Since the researcher is a middle school principal, no subjects were invited or selected under his supervision within the
same middle school in order to avoid a conflict of interest and the possibility of coercion. No teachers under the direct supervision of the researcher were used in the process of developing the research questions.

The researcher utilized a professional network of school and district administrators in order to gain approval for research in their respective middle schools. Clear descriptions of standardized assessments, curriculum standards, and classroom practices were provided to all participants prior to any interview taking place. The three schools chosen share the experiences of preparing for and administering language arts and mathematics standardized assessments to students in Grades 7 and 8. Consent from each district’s superintendent of schools and each school’s principal was obtained prior to any contact with the math and language arts teachers. All interviews were scheduled to meet the demands of the teachers’ availability so that instructional time in the classroom was not lost. The three middle schools selected are similar in student population size, socioeconomic status, and demographics. The three schools are identified as School A, School B, and School C; each teacher selected is identified as Teacher 1 to Teacher 12 in order to keep all schools and teacher identities anonymous.

**Profiles of the Schools Selected**

The three middle schools selected are New Jersey public schools located in the New York/New Jersey metropolitan area and are mandated to have 30% of a math or language arts teacher’s summative, year-end evaluation based on standardized assessment growth in the 2017-2018 school year. All have diverse student populations and currently administer the PARCC assessment in math and language arts for Grades 7 and 8. The three schools are profiled as follows:
School A

School A is a K-8 public school that is housed in the only school building in its respective New Jersey district. It lies within the immediate New York City area and a majority of its students fit into White and Hispanic demographic groups. The three highest represented demographic groups, in order, are Hispanic, White, and African-American. Grades 7 and 8 are within the same building and the school building holds K-8 students. High school students in this district are sent to a surrounding community’s high school. There are approximately 300 students total in Grades K-8 and approximately 70 students solely in Grades 7 and 8. Although the student population is limited, four teachers meeting the criteria (two language arts and two mathematics) are employed at this school and accepted the invitation to be interviewed.

School B

School B lies within the immediate New York City area. Its district is comprised of one Pre-K/Early Learning Center, two K-5 elementary schools, one 6-8 middle school, and one 9-12 high school. The middle and high schools are concentrated in one building on the same campus; however, both operate under separate administrative teams. There are approximately 2,200 K-12 students in this school district. Grades 7 and 8 contain approximately 300 students. The three major demographic groups in this district in order from greatest representation to least are White, Hispanic, and Asian. Four teachers meeting the criteria (two language arts and two mathematics) are employed at this school and accepted the invitation to be interviewed.

School C

School C lies within the immediate New York City area. Its district is comprised of approximately 1,700 students in Grades K-12. Approximately 225 of those students comprise Grades 7 and 8. Grades 7 and 8 reside in a Junior/Senior High School. The Junior/Senior High
School operates under the same administrative team. The three major demographic groups in this district in order from greatest representation to least are Asian, Hispanic, and White. Four teachers meeting the criteria (two language arts and two mathematics) are employed at this school and accepted the invitation to be interviewed.

**Participant Data**

Potential participants in each school received an e-mail of inquiry to see if they were interested in participating in this research study. Each participant and school received an alpha/numeric identification code in order to protect confidentiality with this research. Information regarding total years of teaching experience, years of experience in their respective school, grade levels taught, and accountability to standardized assessments were identified prior to participants becoming solidified. In order to create face validity with the participants, three introductory questions were asked at the beginning of each interview to ensure criteria for participation was reached. Teachers interested but not selected received an e-mail of regret but they were thanked for their interest in this research.

**Data Collection**

The primary method of collecting data was conducted through direct, one-on-one interviews with each of the participants. Each interview was conducted privately in a school setting of the participant’s choice with no opportunity for outside, collegial influence on any response. An open-ended format was utilized and opportunity was provided for follow-up questions from the researcher. This allowed each of the participants to generate qualitative data through all responses. Pre-determined, open-ended questions provided a framework for responses related to practical implications or non-influence of standardized assessments on curriculum development and classroom instruction. Open-ended topics included the following:
(1) exploration of responsibilities related to a teacher’s input into curriculum development; (2) exploration of perceptions related to the influence of standards and standardized assessments on classroom instruction; (3) the selection of resources and materials used to deliver instruction towards standards and standardized assessments; and (4) perspectives on instructional time spent, if any, on teaching technical skills due to the move towards a computer-based assessment.

An open-ended interview took place with each participant. Each interview session was semi-structured with eight pre-determined questions and the opportunity for follow-up on each. A grounded theory approach allowed for responses to identify phenomena based on a particular situation (Creswell, 2003). Strauss and Corbin (1998) also stated that a grounded theory approach holds value in finding data from theory. Preliminary questions asked were as follows: What subject area do you teach? What grade level(s) do you teach? How many years have you taught this subject area to this/these grade level(s)?

Table 1

 Connecting the Research Questions with Respective Interview Questions

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Relative Interview Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>What effect, if any, has standardized testing had on your scripted curriculum and</td>
<td>1. How would you describe your input value to the curriculum content?</td>
</tr>
<tr>
<td>the pacing of it towards standardized assessment preparation?</td>
<td>2. What types of formal programs does your school employ to prepare students for standardized assessments?</td>
</tr>
<tr>
<td></td>
<td>3. In what ways, if any, does standardized assessment preparation alter the curriculum for students?</td>
</tr>
<tr>
<td>What effect, if any, has the high-stakes testing culture had on your decisions of</td>
<td>4. What effects has standardized assessment preparation had on time spent on other instructional units or content in your curriculum?</td>
</tr>
<tr>
<td>content delivery?</td>
<td></td>
</tr>
<tr>
<td>Question</td>
<td>Response</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>5. Has standardized assessment instruction affected your creativity in delivering lessons?</td>
<td></td>
</tr>
<tr>
<td>6. What is your perspective on tying 30% of your summative evaluation to standardized test growth of your students?</td>
<td></td>
</tr>
<tr>
<td>What effect, if any, has the shift towards computerized assessments had on the preparation of students for standardized assessments?</td>
<td>7. How does your school schedule and administer the computer-based standardized assessment?</td>
</tr>
<tr>
<td></td>
<td>8. In what ways, if any, does a computer-based assessment differ from the paper and pencil assessment?</td>
</tr>
</tbody>
</table>

**Data Management and Analysis Plan**

The data management plan allowed for collection, storage, access, and security back-up of all qualitative data collected. All interviews were audio-recorded on an iPhone 7, downloaded to a computer, and saved to a computer hard-drive. A back-up of all audio files were made on two separate USB thumb drives. The audio files were transcribed solely by the researcher and copies of all interview questions and responses were printed. The printed documents were checked for accuracy by reading the printed document as the interview’s audio was played back. The transcribed interview was then sent to the participants to check for accuracy and validity. Final copies of the printed transcriptions were used to record notes, re-examine the depth of responses, and seek common codes.

It is important to utilize time to break down the process in between collection and analysis to allow for perspective on responses (Bogdan & Biklen, 2007). A structured approach to data collection, analysis, and conclusions identified patterns by coding responsive themes. All data were read and analyzed on two separate, distanced occasions.
Field Notes and Reflective Memos

Field notes were taken as responses were heard and recorded. Notes included ideas for follow-up questions and thoughts. Notes were made during and after each interview. Notes also included notions of coding patterns and the researcher’s building theories.

Coding Scheme

Specific procedures were followed when developing coding patterns in order to garner qualitative data. All transcribed interviews were thoroughly read and familiar to the researcher. Throughout the reading, a bullet-point list was made respective to the potential coding categories that emerged. Responses were categorized into coding themes as chunks of information serving as raw data. Responses were compared among participants in order to identify patterns, assign specific codes, and filter all emergent themes. A formal line of communication was maintained between the researcher and participants in order to find more clarity and to better understand context if needed. This also helped the researcher to rely on follow-up feedback and not solely on data. Upon completion of the coding process, tangible conclusions were drawn on emerging themes and categories for interpretation. The data were fully interpreted and conclusions were developed within the scope of the case study. Connections relative to literature and theories on this subject were presented. The implications for this research may be relevant to educators responsible for curriculum development and lesson design.

Validity and Reliability

Validity of data is essential for research credibility. Consistent procedures and standards for methods were necessary for validity. Creswell and Miller (2000) believe that validity is drawn from the theories and conclusions rather than the data. Creswell and Miller (2000) developed a three-step procedure in order to ensure validity of conclusions. They are peer
debriefing, creating an audit trail, and reflexivity of the research. In order to gain insight and validity to my own thoughts and theories, I relied on peers to assess the same data and agree on theories. In order to audit all phases of the research, I used journals, research logs, and a strong chronology of data collection. All data collected and presented were reviewed by an outside party. In order to be reflexive with the research, the researcher included a section titled *Role of the Researcher* to reveal any pre-determined biases that may influence research. After interviews were conducted and transcripts were read twice, the researcher ensured that the proper context was portrayed and the message did not drift from the original codes and themes. The researcher also provided each of the participants with a copy of their respective transcript to validate its accuracy.

**Role of the Researcher**

Interest to the researcher on the effects of standardized assessments on classroom instructional perspectives comes from professional experiences as a language arts teacher, ESL teacher, and middle school principal. This career path has provided much stability as the only profession maintained since earning an undergraduate degree. With 22 years of educator experience, the researcher has observed and witnessed a shift towards measuring teacher worth in standardized assessment results. The researcher works closely with middle school language arts and mathematics teachers and was determined to understand the effects, if any, on development of curriculum and classroom instruction when stress is placed on standardized assessment growth.

Any bias predisposed to this research was related to the researcher’s experiences and observations of administering standardized assessments and using the data to drive instructional
practices. In order to minimize this bias, memos were written and collected for proper bias analysis. All memos and written notes were coded in order to determine if patterns emerged. The researcher followed the Creswell (2003) analysis process as illustrated below:

*Figure 1. Creswell Analysis Process.*
CHAPTER 4

FINDINGS

The purpose of this study was to examine the perceptions of New Jersey middle school language arts and math teachers on the current culture and stress placed on standardized assessment growth for students. This qualitative study was conducted during the spring of 2018 to garner perceptions of teachers who are affected by the 30% year-end, summative evaluation mandate. Specifically, teachers chosen work with students in Grades 7 and 8 in a language arts or mathematics classroom in a public middle school in New Jersey. Six language arts and six mathematics teachers were interviewed using the same interview questions in a structured interview format. The interview questions were the product of three research questions that gave direction to this study: (1) What effect, if any, has standardized testing had on scripted curriculum and its delivery towards standardized assessment preparation? (2) What effect, if any, has the high stakes testing culture had on decisions of content delivery? (3) What effect, if any, has the shift towards computerized assessments had on the preparation of students for standardized assessments?

This qualitative study’s researcher selected three public school districts in New Jersey from a sample of convenience. The three school districts are all in proximity of each other and contain similar demographics and student population numbers. Each public school district maintains Grades 7 and 8 and utilizes the PARCC assessment as mandated by the New Jersey Department of Education. Two middle school math teachers and two middle school language arts teachers who teach Grades 7 and 8 were interviewed for a total of six middle school math and six middle school language arts teachers.
Throughout the interviews, it was common for themes of conversation to cross between topics of curriculum and instruction. This was most apparent in the theme of pacing curriculum to meet the ability levels of students while needing to “spiral back” to ensure skill deficiencies are met prior to moving forward. Staying on pace with instruction was also a common theme that crossed paths amongst the research questions. In order to cover all required content prior to the standardized assessments, teachers faced pressures with staying on pace and meeting all curriculum requirements that ensure preparedness for testing content.

**Research Question 1 Themes and Findings**

**Research Question 1**

What effect, if any, has standardized testing had on scripted curriculum and the pacing of it towards standardized assessment preparation?

The goal of this question was to determine whether or not the development of curriculum was geared directly or indirectly towards the notion of preparing students for standardized assessments.

**Findings**

Standardized assessments are provided to students in language arts and mathematics in New Jersey’s public schools each school year in Grades 3-8 and once in high school. It is expected that students will sit for standardized assessment each year in Grades 3-8 and once in high school in order to determine strengths and weaknesses in reading, writing, and mathematics. The culture in schools today places great importance on standardized assessments, which is evidenced by Grades 3-8 teachers of language arts and mathematics having 30% of their year-end, summative evaluation tied to standardized assessment growth. With this evaluation standard for teacher effectiveness in place, teachers are currently motivated to ensure that
standardized assessment preparation is covered in instructional time. Language arts and mathematics instruction is influenced and prepared through a school’s scripted curriculum, which is designed primarily by the grade level standards.

A scripted curriculum, for language arts and mathematics, relies heavily on the New Jersey Student Learning Standards (NJSLS) in language arts and mathematics for all K-8 grade levels. NJSLS mirrors, in most respects, the standards set forth with the Common Core Curriculum Standards, which were designed to meet testing expectations on the PARCC assessment. The findings for Research Question 1 indicate that Grades 7 and 8 language arts and mathematics curricula in New Jersey are designed first by unpacking the NJSLS (formerly CCCS) standards and developing content that will support the skills expected at the end of each grade level. Content is then organized into a paced framework for the school year. Although standards are prescribed, the selected content is controlled locally; and teachers get the opportunity to provide input and develop a pacing schedule of the curriculum throughout the school year. There is a strong correlation between skills presented in the CCCS and NJSLS and skills required to be deemed proficient in PARCC.

**Interview Results for Research Question 1**

Three interview questions were asked of each teacher to address the topic of the influence of standardized assessments on a middle school’s scripted curriculum in language arts and mathematics. As the interview process evolved, the researcher added two questions involving data used to adjust curriculum and how content for curriculum standards is developed. Because these two topics of discussion came up frequently, it was relevant for them to be added. According to Creswell (2014), it is appropriate in qualitative research to add interview questions
during the research in order to reflect topics and themes that arise and become relevant to the research question.

1. How would you describe your input value to the curriculum content?

All 12 teachers interviewed stated that the curriculum adopted in their school was developed first by understanding and unpacking the standards required by the New Jersey State Department of Education. Standards for this research identify specific skill sets recommended for language arts and mathematics to be completed by the end of the school year. In the 2017-2018 school year, New Jersey’s public schools utilize the New Jersey Student Learning Standards (NJSLS), which are a direct derivative of the Common Core Curriculum Standards (CCCS).

Ten out of the 12 teachers interviewed participate in meetings at the beginning of each school year where curriculum is adjusted based on standardized assessment data from the previous school year. These beginning-of-the-year adjustments in all three schools come from the analysis of students’ strengths and weaknesses as identified primarily by standardized assessment data from PARCC. Teacher 1 stated, “The adjustments in curriculum give me the chance to place more time and emphasis on skills the students struggled with. We see where the gaps are and where we need to focus.” Again, these skills have a direct correlation to expectations in the standards. All three middle schools had teachers reporting that adjustments are made at the beginning of the school year in collaborative settings where teachers examine data together and come up with a common plan for emphasis. Two teachers from the same school identified this practice as a PLC, or Professional Learning Committee.

Curriculum adjustments also occur throughout the school year with a “spiraling back” approach. This terminology was used by three teachers, all teaching language arts. This
approach allows teachers to use formative and summative assessments to determine skill weaknesses of individual students. After the weaknesses are identified, students are brought back to skill concepts introduced earlier to ensure comprehension is gained before moving forward. This can be identified as a differentiated approach for each student.

Although teachers are permitted to adjust the curriculum to meet the skill levels of students, pacing remains an important component to curriculum and mandatory to their practice. They must ensure teaching of all standards expected for the skills required on the standardized assessments. Since most standardized assessments are administered in April, the entire school year is not available to ensure preparedness on the standardized assessments. Pressures of remaining on pace and covering all material prescribed in the curriculum prior to the standardized assessments were noted by three of the teachers interviewed. Teacher 4 stated, “We are able to modify the curriculum as the year progresses but as time permits, of course. We are in a time constraint of when the test is.”

Although the standards drive curriculum, which in turn prepares students for standardized assessments, 10 out of the 12 teachers interviewed stated they have a vast amount of input into the content that delivers the curriculum and their lessons. Two teachers interviewed stated they have no input at all. Freedom of choosing resources and materials is the key component to the input and development processes as described by 10 teachers. Local decisions are made collaboratively by teachers, supervisors, and administrators of each respective middle school in all three districts chosen. Freedom of content, according to eight of the teachers, includes choosing textbooks and other supplemental resources purchased, shared, or found online. Teacher 4 stated, “I’ve created my own Power Point presentations and use the internet to find resources, so I can put my own spin on the curriculum.”
2. What types of formal programs does your school employ to prepare students for standardized assessments?

Two of the three middle schools chosen use a formalized, computerized benchmarking program that is purchased. This program has students sit in front of a computer to take a diagnostic assessment in the beginning of the year only. No follow-up benchmarking assessments are provided throughout the rest of the school year in those schools. According to six of the teachers interviewed, those benchmarking programs are used to place students into ability level groupings and serve as a source for ability level planning.

3. In what ways, if any does standardized assessment preparation alter the curriculum for students?

During the teacher interviews, five teachers brought up the topic of pacing and how this creates pressure to present specific content in a short timeframe prior to PARCC testing. Teacher 12 stated, “We have the opportunity to go slower when students are struggling, but you have to get a certain amount done before that test. You have to get it completed because then they won’t be exposed to a large chunk of what they will see on the test.” Three of the teachers interviewed raised the topic of taking the time to “spiral back” to ensure understanding of a specific skill before moving forward in the curriculum. Teacher 7 indicated that some students begin middle school with skill proficiency equivalencies of Grades 4 and 5 standards. These students, according to Teacher 7, get placed in leveled classes that begin with the appropriate grade-level standards, such as Grades 4 and 5, prior to advancing to the Grades 7 and 8 standards. The opposite is true for students with higher levels of proficiency in language arts and mathematics. Teacher 10 stated, “We start the honors students later in the book because the first couple of chapters are review, which we now gloss over.” Teacher 10 also spoke of the
pressures of pacing by stating, “We are not getting a lot of time to review in September. A lot of our kids need that review, but we don’t have the time because we need the time in April to get ready for the test.”

Higher-order thinking skills is a theme that was mentioned by six out of the 12 teachers interviewed. All six spoke about the changes in PARRC and the fact that all questions asked look for more than just the final answer. The process of attaining the answer is also analyzed. This creates multi-level questioning, which comes with two or three components to open-ended questions in both language arts and mathematics. Teacher 8 stated, “The difference between PARCC and other standardized assessments is that it will ask you a question (there are a lot of two-part questions) and it’s all about providing evidence. So even if they know the answer, they are looking at two different skill sets.” Teacher 2 commented on the curriculum changes to address higher-order thinking questions. Teacher 2 stated, “The answer is the least important part of the question on the PARCC. It’s now about how they arrived at the answer. There’s only one point for getting the answer right now.”

Creativity in designing lessons emerged as a theme throughout the interviews. This topic of discussion focused on the materials chosen to deliver the curriculum, which is directly correlated to the standards. The focus of this question varied between the language arts and mathematics teachers. The language arts teachers are able to be more flexible in the materials and resources chosen to target the standards. Math teachers, primarily in all three middle schools selected, utilize a strict reliance on their respective textbooks, which are all geared towards Common Core Curriculum Standards and provide online resources geared towards PARCC practice. Teacher 1, a language arts teacher, stated that she is able to focus on the same skills across varied content. “If I think about teaching different text structures, we will dive into a text
and focus on different text structures within that text. We will review it, practice it, deliver a fresh read, and apply the same types of skills to different text content to see if they can apply it to a more formal assessment.” Teacher 12, a math teacher, spoke about “rearranging the content in the textbook” in order to meet the needs of proficiency levels of students throughout the school year. Teacher 11 spoke about losing creativity in teaching due to standardized test prep. “We should be doing more student-based, project-based work, which is more interesting for the students. I find myself taking the curriculum and trying to adjust it so that it follows the format for PARCC testing.”

A final theme that emerged was the adjustment of curriculum in order to provide time spent on direct standardized test preparation. This includes sample questions being posed to students in an informal testing setting and the analyzation of the expectations for a proficient response. Five of the teachers interviewed brought up the topic of going into “test prep mode” and gearing the curriculum solely towards this at least two weeks prior to the formal standardized assessment administration. Teacher 9 stated, “About two weeks before the assessment, we spend more time prepping for the test. But the prep work is not done for an extended amount of the school year.” Teacher 10 stated, “We need to take the time in April to get ready for the test.”

**Research Question 1 Summary**

Research Question 1 investigated the influence of standardized assessments on a middle school’s scripted curriculum in language arts and mathematics. This included the development of curriculum, which focuses on content chosen and the standards used for the framework of skills. The pace of a curriculum dictates how much time is spent on specific skills and content, and it is focused on meeting deadlines for standardized assessment administration. It was
important to investigate a teacher’s input into curriculum, benchmark programs outlined in the curriculum to ensure standardized assessment preparedness, and any possible relationships of standardized assessments to narrowed curriculum.

Research Question 1 findings indicate that teachers have no control over the skills required to be learned at the end of each grade level. However, curricular control comes with choosing content, resources, and materials used to teach towards those required skills. Teachers interviewed indicated that pacing of curriculum is the largest inhibiting factor of identifying student deficiencies and spiraling backwards to ensure students grasp concepts before moving forward.

Research Question 2 Themes and Findings

Research Question 2
What effect, if any has the high-stakes testing culture had on your decisions of content delivery?

The goal of this question was to determine whether or not teachers were spending significant amounts of time on standardized testing preparation and whether or not teaching to a computer-based assessment took time away from teaching skills required for language arts and/or math.

Findings

The PARCC standardized assessment closely mirrors the Common Core Curriculum Standards, which are now heavily similar to the newly adopted New Jersey Student Learning Standards. Teachers feel that the NJSLS are just a re-numbering and grouping of the Common Core Curriculum Standards. Teacher 12 stated, “The New Jersey Student Learning Standards are the same as the Common Core Standards; they just renamed them. There are a few minor words or phrases that changed. But I compared them and they are almost exactly the same.”
Curriculum delivered, even with a shift to NJSLS, remains geared towards standardized assessment preparation, targeting the skills required on the PARCC assessment. Delivery of instruction does change to prepare for standardized assessments approximately 2-3 weeks prior to the administration of the assessment. This is defined as practicing sample PARCC questions and doing this on a more consistent, daily basis. Throughout the rest of the school year, curriculum is delivered as scripted in the pacing of content and focused on the skills outlined in the NJSLS.

Instruction in the classroom and the design of lessons are directly correlated to a school’s curricula. Specifically, in middle school language arts and mathematics classrooms, standards are tied to high-stakes tests, which are tied to teacher effectiveness models. It is important to look at perceptions of teachers on whether or not standardized assessment preparation exists and has any effect on instructional time and the creativity in designing and delivery of lessons. Teacher 4 made the statement as related to the effects of standardized testing on instruction, “We are always trying to chase after results.”

Interview Results for Research Question 2

1. What effects, if any, has standardized assessment preparation had on time spent on other instructional units or content in your curriculum?

Content within the middle school math and language arts curricula, as utilized and approved at the middle schools targeted for interviews, is highly focused on the New Jersey Learning Standards, which are required by the New Jersey Department of Education. These standards reflect the NJSLS and are the direct descendant of the CCCS in the State of New Jersey. When delivering the curriculum, five out of the 12 teachers commented that the content taught is highly centered on PARCC topics and the standards related to PARCC. Teacher 6
stated, “We definitely vary how we deliver it. We lecture, deliver small group instruction, large group instruction, do a lot of hands-on activities and a lot of moving around the classroom. Students are able to do all of these activities and learn the skills in the curriculum. But the curriculum is focused on the standards.” Teacher 10 stated, “I don’t remove things from the past because I still find value in teaching them. I try to incorporate all topics as much as possible so that I’m not always teaching to the book as related to the standards. But, in all, the students are still learning towards the standards.” Teacher 2 stated, “Although most of the curriculum relates to the standards on PARCC and we focus on PARCC’s major content areas, we have the freedom to do things off of the PARCC topics. We do, however, mostly focus on PARCC’s major content areas.”

Five out of the 12 teachers interviewed made statements linking the pacing instruction to preparedness on PARCC. This also includes the removal of material that might have been taught prior to the inception of Common Core Standards and PARCC. Teachers 8 and 10, both math teachers in different middle schools, stated that they now start the school year in Chapter 3 in their math textbooks. This does not allow for the content review from the previous school year that used to take place in years prior to the administration of Common Core Standards. Three teachers spoke about lessening content. Teacher 8 stated, “A lot of our kids need the foundational blocks that should be reviewed but are missing it because you need that time in April to get yourself ready for that test. You’re definitely pulling pieces out.” Teacher 11 stated, “It starts getting into the back of my head that PARCC is coming up and I have to cover specific material that the kids will be tested on. I might want to get into a poetry unit or something to that effect, but now I have to think twice because I have to prep the students for PARCC.” Teacher 12 stated, “I definitely feel you have to rush things and don’t have time for the more interesting
aspects of it. You don’t have the time to get to everything on the test so instruction definitely has to be altered to some extent.”

The removal of content and increased pacing of instruction were followed by statements of reduced flexibility and a majority of instructional time focused on PARCC assessment content. Five out of the 12 teachers interviewed made statements related to PARCC content taking up a majority of instructional time. Teacher 4 stated, “I rarely deviate from what the textbook is doing. Our focus in the classroom has been on the standardized assessments.” Teacher 2 and Teacher 4 spoke about focusing instruction on student weaknesses on PARCC related skills and content. Both teachers made statements about instructional content being removed in order to make time for “spiraling” students to address skill deficiencies.

2. Has standardized assessment instruction affected your creativity in delivering lessons?

Nine out of the 12 teachers stated that there was no loss of creativity in creating lessons designed to teach to the skills outlined in the New Jersey Learning Standards. Teachers 3 and 4 both stated that it was crucial to make learning fun for their students, which lends itself to their creativity. Teacher 5 used experience with the standards as key to being creative since positives and negatives of former lessons can be drawn upon. Teacher 5 stated, “As you do these skills year after year, you try to get a little more creative with it. If you weren’t creative with your lessons, the kids would not be successful with learning the skills.” Teachers 1, 3, 6, 7, and 11, all language arts teachers, spoke about the freedom to choose content related to the skills. This included the choice of curriculum content such as stories, topics for writing and real-life scenarios related to students’ lives in order to target the standards. Teacher 5 stated, “I now have to get deeper into questions since the word problems are more multi-level. The students now are responsible for showing the work and I have to come up with real-life scenarios to the PARCC
skills. I’d like to think that I have to use more creativity to come up with these scenarios than five or ten years ago.” Teacher 3 stated, “When the lessons aren’t creative, the kids aren’t learning. I make sure I keep that creativity in there because that’s what makes them want to learn.” Teacher 10 stated, “They can change the standards a million times. Teaching Pre-Algebra is still Pre-Algebra. I can still use all the hands-on activities, all the fun things you can do. The art of teaching hasn’t been lost in the standardized testing culture.”

Although most of the teachers relayed a statement of no loss of creativity in designing lessons, three of the teachers made comments related to the focus still needing to be on the standards, which are directly related to PARCC. Two of the teachers made statements related to pressures of pacing affecting creativity. Teacher 11 stated, “Although I am able to be creative in how I deliver the lesson, I still have to pace it accordingly and move on.” Teachers 11 and 12 spoke about pacing pressures and not being able to delve deeper into project-based approaches because they take up too much instructional time. Teacher 8, a math teacher, stated, “Sometimes I cannot stop long enough to do projects. I do very few projects these days. However, you can spend time on things like statistical games or use manipulatives for the more real-life applications. We are definitely at a much faster pace.”

3. What is your perspective on tying 30% of your summative evaluation on standardized test growth of your students?

Grouping students together in a growth model and not taking ancillary factors into consideration with standardized tests was a strong theme of the 12 teacher interviews. Factors that teachers feel, are not taken into consideration are demographic information, socioeconomic status, apathy and indifference of students to testing, transience of students, limited English proficiency, ability levels, and options for test refusal. Ten of the teachers responded to these
factors as being the essential ingredient in determinations of testing growth and proficiency. The teachers’ responses related to factors are listed in Table 2 (teachers were able to list more than one factor as a major determinant)

Table 2

*Teachers’ Responses Related to Factors*

<table>
<thead>
<tr>
<th>Demographics as a key factor</th>
<th>Teachers 2 and 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Socio-Economic Status</td>
<td>Teachers 2, 7, 9 and 11</td>
</tr>
<tr>
<td>Apathy and Indifference towards Testing</td>
<td>Teachers 4, 5, 6, 8 and 12</td>
</tr>
<tr>
<td>Transience of Students</td>
<td>Teacher 9</td>
</tr>
<tr>
<td>Ability Levels of Students</td>
<td>Teachers 5, 6 and 10</td>
</tr>
<tr>
<td>Test Refusal</td>
<td>Teachers 6, 8, 10 and 12</td>
</tr>
<tr>
<td>Testing in Languages other than English</td>
<td>Teacher 10</td>
</tr>
</tbody>
</table>

As related to these findings, Teacher 4 stated, “It can all come down to how a student feels on the day of the test.” Teacher 11 stated, “Some kids can afford to hire tutors and some cannot.” Teacher 8 stated, “Politics can dictate a teacher’s evaluation. Parents can force their children to refuse the test and the accountability falls solely on the school. Teachers get stuck in the middle of the politics between legislators and parents,” adding, “Some of our brightest kids are refusing to take the test.” Teacher 12 stated, “I don’t feel that seventh or eighth graders put all of their effort into the test because it means nothing to them. They think that if it has nothing to do with their grade; they will just answer the questions randomly.”

The topic of fairness also was raised quite often in the teacher interviews. Nine out of the 12 teachers spoke about the lack of fairness in weighing 30% of their summative evaluation on
student growth in PARCC. Teacher 4 stated, “30% adds a lot of pressure to the teacher. We spend the entire school year teaching the content and it all comes down to three days . . . 30% in three days. That sounds a little ridiculous when you say it out loud.” Teacher 5 stated, “Prepping for the PARCC has a direct result on your evaluation. I don’t think it’s fair to teachers that there is part of your evaluation that you don’t have any control over.” Teacher 7 stated, “I’m not saying that we should not be held accountable because we do teach the most important skills necessary for success in life. But 30% of my evaluation is a lot of weight.” One teacher also spoke about the lack of fairness being related to colleagues teaching content areas other than math or language arts. Teacher 10 stated, “If a student is not strong in math, he might not also be strong in science. But that science teacher isn’t being scored and doesn’t have the pressure. I think it should be all subjects or none.”

There are teachers, however, who stated the 30% mandate did not alter their instructional attitudes or did not put pressure on them in the classroom. Teacher 1 stated, “It doesn’t really phase me that much because I know what I have to teach and what I have to cover. It doesn’t bother me so much.” Teacher 8 stated, “Every math and language arts teacher (in Grades 3-8) is in the same boat. As long as you’re on par with your peers you are going to be ok. I don’t get too worked up about it.” Teacher 2 stated, “We have to remember that the 30% is related to growth and not overall proficiency. If you can focus on the students’ weaknesses, you should be able to show growth.” Teacher 3 stated, “It doesn’t change the way I’m teaching. I am going to do the best I can whether they’re tying 30%, 0% or 100% into it. I teach the best way I know how.”

Two outliers for responses to the 30% mandate included Teacher 11’s concerns with teachers focusing only on weaknesses of students who are capable of demonstrating growth and
the same teacher’s concern with the lack of qualitative attention paid to social/emotional growth of middle school students. Teacher 11 quoted, “Why am I going to worry about the kids who struggle the most and won’t do well when I can focus on the students who can improve my growth [in test results].” Potentially, it is a numbers game that is being set up. You are creating a quantitative measure that you have no control over.” and “The standardized test growth doesn’t take into consideration the social and emotional well-being and growth of our students. This type of system rewards teachers who are all about the academics and not the social/emotional development.”

**Research Question 2 Summary**

Research Question 2 focused on any possible effects of standardized assessments on the delivery of instruction. This includes the pacing of instruction and directly preparing students for assessments by presenting sample PARCC questions to students for practice. It was important to explore impacts, if any, this might have on a teacher’s creativity in developing lessons and time spent on identifying and working on students’ skill weaknesses. It was also important to gain perceptions of teachers affected by the AchieveNJ (2013) mandate of tying 30% of their end-of-year, summative evaluation to PARCC results. Instructional pressures to prepare students for PARCC, if any, were explored.

Research Question 2 findings indicate that instructional time is being utilized to teach towards standardized assessment proficiency. Teachers have pressure placed upon them to teach to the test when 30% of their summative evaluation is weighted towards test results. Teachers do have creativity in creating and delivering lessons. They feel that the recent culture of standardized assessments has taken attention away from key components of education such as social and emotional growth. Teachers also indicated that standardized assessment results do not
Research Question 3 Themes and Findings

Research Question 3
What effect, if any, has the shift towards computerized assessments had on the preparation of students for standardized assessment performance?

The goal of this question was to determine whether or not a computer-based assessment affected a school’s instructional time more than when a paper and pencil test was administered. This is due to instructional time possibly being utilized to teach technology skills required to effectively take a computerized assessment.

Findings

Computer-based assessments have changed the way in which students prepare for and take standardized assessments. This is a fact in knowing that standardized assessments were delivered, in the past, with paper and pencil. When the PARCC assessments were first introduced, concerns arose such as, but not limited to, typing proficiency in a timed setting, typing out open-ended questions, proving work in mathematics and using a drop-down, digital calculator. However, students today are “digital natives” and extremely literate in the use of technology. They quickly adapted to a computerized assessment and the skills required to complete them.

Scheduling computer-based assessments remains a challenge for some school districts not adequately equipped with 1:1 device programs and small quantities of computer devices. Scheduling requires schools to have students taking the assessments within windows that are approximately four to six weeks in length. School C adopted a special schedule during testing
weeks in order to create two-hour blocks of time where students can rotate and be tested in the morning and afternoon hours. All three schools shut down classrooms for testing security mandates and to limit internet access to instruction due to bandwidth concerns.

Since Research Question 3 related to the school’s schedule during testing and the number of devices available to take the assessment, the responses were extremely similar respective to the school in which the teacher was employed. Therefore, findings for Research Question 3 are delivered by identifying school, the schedules they employ during testing, and the number of devices available to the students during testing.

**Interview Results for Research Question 3**

1. How does your school schedule and administer the computer-based standardized assessment?

**School A**

The four teachers employed at School A all reported that, in the past, there were not enough devices for all of the students to test at the same time. The laptops were rotated amongst students and a computer lab was occupied based on a pre-determined schedule that had students testing on different days and at different times. Student groups were rotated throughout the day with some taking the assessment in the morning hours before lunch periods while others took the assessment after lunch periods in the afternoon hours. While some students were being tested, instruction continued for the rest of the students in the school. Teachers were also rotated in and out as test administrators and proctors, which meant that substitute teachers had to be hired in order to cover their classes during testing administration. Classrooms and science labs were used for students who tested on the laptops. In total, School A, in the past, would test all students within a five to six week timeframe that spanned April into May. Teacher 2 reported, “In the
past, it did take over a month of instructional time away. We were missing class a lot.” Teacher 3 reported, “In the past, some kids tested in the computer lab, some in the science labs, and some in the classrooms with laptops.” Teacher 4 reported, “Testing, in the past, was scheduled based on the availability of equipment.”

During the 2017-2018 school year, School A purchased Google Chromebooks for all of their students under a 1:1 initiative. Students will test at the same time while remaining in their homeroom period for the morning hours only. In total, during the current school year (2017-2018), testing will be completed in six school days – three days for language arts and three days for math. Teacher 1 reported, “The middle school will essentially test in one school week since we now have that ability.” Teacher 2 reported, “The school will kind of shut down for one to two weeks, but that’s better than in the past when it was five to six weeks.” Teacher 3 reported, “This year, the Chromebooks are going to make a big difference.”

School B

All four teachers from School B reported that last school year (2016-2017), testing was conducted over a four-day period at the middle school with all students testing for language arts and mathematics in the morning and afternoon hours. In total for the school district, testing consumed five to six weeks. They rotated in and out of a computer lab and were distributed laptop computers in the classrooms. These laptop computers were the same devices utilized for students throughout the entire school district. This meant that the middle school, according to Teacher 5, had to “wait their turn” for the laptops to be delivered to the middle school. Teacher 6 reported, “There had to be a redistribution of devices to one building at a time.” This also meant that computers were unavailable for use until the entire district completed testing. As related to testing being completed in a four-day period, Teacher 5 stated, “That was too much.
They tested in the morning, ate lunch, and then tested again in the afternoon for four days straight. That was not the optimal condition for kids to be reading and writing for that long."

Teachers 10 and 11 also reported that the WiFi capabilities of the school were shut down during testing periods in the past. This was done, according to Teacher 11, because “there was not enough bandwidth for the testing software application to work if teachers were using the WiFi for instruction.” Teacher 11 also stated, “I couldn’t use any internet in the classroom during the five to six week (district-wide) testing period, which was extremely disruptive to how I teach.” Teacher 10 reported, “In the past, no computers or computer labs were available during the (district-wide) testing periods.”

This year, the school district in which School B is located purchased additional Google Chromebook carts, and testing at the middle school will take place in approximately one week in the morning hours only. Teacher 10 reported, “We are getting better at this with experience. They updated our WiFi and purchased more devices, which should give us the ability to test more students at the same time.” Teacher 11 reported, “Things should be much better this year.”

School C

School C teachers reported that their school utilizes a special schedule named for their school’s mascot. The schedule is a rotating, block-like schedule that allots two, two-hour blocks of time in the school day. One two-hour block occurs in the morning and one in the afternoon. Students are tested within these two-hour blocks of time allotted in the morning and afternoon hours. This does not mean that students test in both two-hour blocks of time within the same day, although this is a possibility. If a student is not testing during that two-hour block, they will sit in the scheduled instructional class for the two-hours. This is across the board for all subject areas. Outside of the two-hour blocks, all other periods of the day are 30 minutes in length.
Instructional time for each course over the timeframe of the standardized testing window adds up to an equal number of minutes when the school operates under a normal, full-day schedule respective to the same amount of school days. Overall, this special schedule is in effect for approximately two to three weeks, and approximately 20-30 students are tested at one time. Teacher 7 stated, “It alters the entire schedule. Even when you’re doing the special schedule for testing, the other classes are only 30 minutes. It affects every class—music, art, math, etc. Students not testing can also be sitting in art or social studies for two hours straight, which is a lot of time. You’re still trying to teach and getting the right overall amount of time, but it’s broken up and alters things.” Teacher 9 stated, “The students can be tested in the morning and afternoon in the same day, which is disruptive to their school day.”

Two of the four teachers interviewed at School C commented on the lack of computer devices to test the students and how this also affects instruction. Teacher 8 stated, “There are still not enough devices to test a larger number of students at once. No new devices were purchased over the last several years.” Teacher 12 stated, “No other classes can use computers during the testing window because they are all being used for PARCC. This can take over three weeks to complete.”

2. In what ways, if any, does a computer-based assessment differ from the paper and pencil assessment?

The technological tools required for navigating a computer-based assessment in language arts and mathematics were reported, overall, by the teachers as follows:

- Secondary pop-up screens to take notes
- Eliminating answers by crossing them out
- Enlarging text boxes to type more information on essays and short answers if needed
• Turning pages in text for reading components
• Typing skills to ensure speed is on par with timeframe of respective section of the test
• Using the cursor and tab keys
• Underlining text
• Highlighting text
• Accessing reference sheets
• Using drop-down features
• Entering numbers as fractions
• Utilizing the mathematics equation editor
• Dragging on the screen to create the lines for graphs
• Using a digital, drop-down calculator (replaces a hand-held calculator)
• Using a measurement tool (replaces a hand-held ruler)
• Dragging and sliding text and numbers into an answer box

Testing websites and private software are available for students to practice with PARCC’s computerized format. This includes, but is not limited to, the New Jersey Department of Education’s website, PARCC practice assessments offered through Pearson websites, textbook supplemental online components, and privately owned and operated software sold to school districts through numerous educational vendors. Teachers interviewed did report a continuation of paper and pencil practice in order to develop an understanding of the types of questions being asked, such as multi-step questions and open-ended questions. Overall, across the 12 teachers interviewed, all stated that the students are very tech savvy today and do not need much practice prior to the administration of the standardized assessment. An approximate average of two to three weeks of computerized practice prior to the assessment was reported.
Teacher 1 stated, “The tech skills are more related to taking notes or writing in the margins of the text on the screen. The kids have adapted very well to it because they are tech savvy. The computer format doesn’t interfere too much.” Teacher 2 reported, “We take the kids up to the lab at least a half dozen times to get them familiar with it.” Teacher 3 stated, “We spend some time teaching them how to use the text box and expand it if they need more space to write.” Teacher 4 stated, “We do reserve a little bit of time before the test to get on there and do a practice run. There are manipulatives on the test that you cannot simulate on paper. This is done about two weeks prior to the test.” Teacher 5 reported, “Middle school is not the first time they’ve seen this testing format. By the time they get to us, they are familiar with it.” Teacher 6 stated, “The only time we spend on the computer throughout the school year is when we are preparing for PARCC. We show them how to input their answers, make graphs, and plot with a ruler, which takes away from instructional time.” Teacher 7 stated, “We practice on the computer a few weeks before the test. The graphing tool is pretty unique and they have to know how to use that when they take the test.” Teacher 9 stated, “I go over the underlining and magnifying features with the students a couple of weeks before the test.” Teacher 10 stated, “It’s not brand new. A lot of the programs offered today use the same tools they need to use on PARCC.” Teacher 11 stated, “The kids are digital natives but they do need to have some practice. This puts us at a disadvantage with districts who have a 1:1 initiative.” Teacher 12 stated, “They are used to the test by the time they are in middle school. They’ve taken it before in elementary grades.”

**Research Question 3 Summary**

The goal of Research Question 3 was to determine effects, if any, a computerized assessment has on the delivery of language arts and math skills. Teachers identified 16 different
technology skills required to navigate through the computerized PARCC assessment. It is important to determine whether or not teachers are taking instructional time away from the skills listed in the standards to teach technology skills. It is also important to investigate whether or not a computerized assessment affects school schedules and use of instructional technology during testing windows. This has the ability to include the number of computers available and bandwidth available to properly have a pre-determined number of students taking the assessment at the same time.

Research Question 3 findings indicated that students today are “digital natives” and have adapted well to a computer-based standardized assessment. Software programs being used in the classroom assist students with skills needed to navigate through a standardized assessment. Sixteen technology skills were identified by the teachers interviewed as being necessary to complete this task. Instructional time to teach these technology skills does not take time away from math or language arts instruction.

Final Thoughts

All 12 teachers were asked the following question to close the interviews: Is there anything you would like to add that wasn’t covered in our interviews related to standardized assessments in schools today?

The overarching themes of this final question had seven of the 12 teachers mentioning the lack of fairness in creating a 30% influence of standardized assessment results on a summative, year-end evaluation. All of them understood the importance of accountability but did not feel that many of the student factors affecting kids today are taken into consideration when measuring student progress and growth. Three teachers stated that standardized assessments are not new and that they have been around for many years; however, the lack of fairness in tying results to
evaluations was mentioned by all three after the initial statement was made. Teachers 2 and 4 spoke about the development of standardized assessments in measuring appropriate skills such as higher-order thinking and problem-solving. Both of these teachers, along with Teacher 9, made statements that linked the skills required on standardized assessments as skills the students should know prior to entering high school. However, both are included in the statements related to the lack of fairness in tying results to teacher effectiveness. Ten out of the 12 teachers spoke about the pressures of the 30% mandate and how this is in their minds as they are teaching students.

Summary

Overall, teachers interviewed support the skills required in the standards. They also feel that the curriculum developed based on these skills does not overwhelmingly dictate the content nor the creativity in delivering lessons. The choosing of content, however, was different with language arts and math teachers. Language arts teachers have more freedom in choosing material to deliver the skills, such as novels, short stories, and writing prompts. Math teachers collectively teach from the textbook and follow the patterns of units presented in numerical chapters. This does not, however, take away their creativity in designing projects or having students do collaborative group work. The creativity in designing and delivering lessons is not affected. The notion of tying assessment results to accountability and effectiveness is the major detail that all 12 teachers focused on during the interviews. They believe there is a lack of fairness towards math and language arts teachers solely having a mandate directing 30% of their evaluation towards the quantitative measure of standardized assessment growth. The teachers also feel that standardized assessments do not reflect all of the factors that affect students and their performance on days the tests are administered.
CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

Introduction

The purpose of this study was to examine the perceptions of middle school language arts and mathematics teachers on the current culture of accountability being tied to student assessment results. Specifically, this study focused on public middle school language arts and mathematics teachers teaching Grades 7 and 8 in New Jersey. All teachers interviewed for this study currently have 30% of their summative, year-end evaluation directly influenced by PARCC scores; therefore, it was important to gain perspectives directly from the affected educators (Kruger & Casey, 2000). This study addresses any concerns related to this 30% mandate and any direct or indirect influences of its existence on curriculum and instruction.

This study also focused on the shift from paper and pencil assessments to computer-based assessments. There are technology skills required for students to navigate through the assessment in order to read instructions and answer questions. It is important to determine whether or not teachers are replacing skill practice in language arts and mathematics instruction with lessons dedicated to teaching technology skills needed to take the PARCC assessment. This study provides insight into the effects of standardized assessments on curriculum and instruction in public middle schools and any pressures caused by a 30% mandate affecting end-of-year summative evaluations for language arts and mathematics teachers.

This was a qualitative study designed appropriately in methodology (Creswell, 2003) to determine direct and indirect influences of standardized assessments on the practice of teaching. After ensuring anonymity in name and school, participants volunteered to share honest and valid opinions related to their experiences of teaching for proficiency on standardized assessments.
Impacts on development of curriculum, pacing of curriculum, development of lessons, and instructional delivery were explored. This study’s sample consisted of 12 middle school language arts and mathematics teachers who directly impact instruction for students in Grades 7 and 8. Six language arts and six mathematics teachers were selected from three public middle schools in New Jersey; two language arts and two mathematics teachers were selected from each school. Each teacher selected currently receives an end-of-year, summative evaluation that ties 30% of PARCC results to that evaluation, also known as a Student Growth Percentile (SGP). All of the participants have teaching experience of five years or more. The research was gathered by conducting individual interviews (Creswell, 2014) with each of the qualifying teachers in the spring semester of 2018.

The interview questions were reviewed by a jury of experts and designed to follow a structure that demonstrates understanding to any effects on the beginning-to-end process of educating children (Kruger & Casey, 2000). The researcher started with effects on curriculum development and then moved to the process of designing lessons and delivering instruction. The research questions selected were as follows: (1) What effect, if any, has standardized testing had on your scripted curriculum and the pacing of it towards standardized assessment preparation? (2) What effect, if any, has the high-stakes testing culture had on your decisions of content delivery? (3) What effect, if any, has the shift towards computerized assessments had on the preparation of students for standardized assessments?

**Summary of Major Findings**

Standardized assessment results have become weighted with accountability and determinations of effectiveness of teachers. Most affected by this is the fact that New Jersey teachers who teach language arts or mathematics to students in Grades 3-8 are impacted by being
mandated to tie 30% of their end-of-year, summative evaluation to PARCC results. This places unnecessary pressure on the teachers affected to ensure that all students are prepared to take the assessment and all students demonstrate growth regardless of circumstances. Teachers interviewed made statements related to the pressures of pacing and preparing students for standardized assessment deadlines. “Chasing after results” was mentioned by one teacher as the overall pressure being faced when tying 30% of an evaluation to those results.

Teachers interviewed understood the importance of teaching critical skills in reading, writing, and mathematics. Teachers of language arts and mathematics recognize the vital element to ensuring skill growth and getting students prepared for the next grade level. However, a cut and dried 30% mandate on their evaluations does not take a number of factors into consideration on testing days. The factors ignored in this mandate and listed in the Chapter 4 Findings include socioeconomics, transience of students, limited English proficiency, learning disabilities, and the fact that students are indifferent to test results or outright refuse to take the test. The teachers interviewed often associated this with a lack of fairness and a lack of understanding the realities that teachers are faced with in extremely diverse demographic school settings.

Standards provide the framework for curriculum development and the delivery of instruction (NJDOE, 2016a). In the state of New Jersey, three sets of standards have been implemented in the span of 10 years (NJDOE, 2016b). The major shift was moving from the New Jersey Core Curriculum Content Standards to the Common Core Curriculum Standards in 2010. This change, in language arts and mathematics specifically, altered the skills students are prescribed to learn by the end of each grade level. Common Core also narrowed the standards and removed what was deemed repetitive and non-essential (Common Core Standards, 2010). In
2016, New Jersey adopted the New Jersey Student Learning Standards, which are strikingly similar to the Common Core Curriculum Standards. One teacher commented on how NJSLS were identical in skill language to Common Core but with just an overall renaming and minor reshuffling of the skill indicators.

Common Core Curriculum Standards are directly tied to PARCC assessments. The reading, writing, and mathematics skills popping up on PARCC reflect the standards written into the local curriculum. Since NJSLS closely mirrors the Common Core Standards, proponents of standardized assessments will argue that PARCC is a good indicator of whether or not those standards are being met. It provides an objective measure of whether or not the standards prescribed in the curriculum are being demonstrated in the language arts and mathematics classrooms. However, it is not the only measure available to determine success and growth of the skills outlined in the standards, and there are more authentic methods of assessing students (Wiggins, 1989). Teachers interviewed made statements related to internal benchmarks and formative and summative assessments as better measures of real-time progress and having the ability to measure strengths and weaknesses as the school year progresses.

Teachers are professionals who are trained to differentiate instruction in order to meet the needs of every child. Standards are the framework, but they do not tell us what to do if a student struggles with specific skills, falls behind, and needs more time to master those specific skills before moving forward. Teachers interviewed spoke about “spiraling” their curriculum in order to go back and review skill components. They make determinations of skill struggles based on their own formative and summative assessments. This can also include using benchmark assessments to determine strengths and weaknesses. The weight of tying 30% of a teacher’s evaluation to one assessment that is a brief “snapshot in time” narrows curriculum and does not
take into consideration the daily academic growth or the social ability to work collaboratively and collectively with peers for success (Costigan & Crocco, 2007).

Pacing pressures of both curriculum and instruction due to the influence of standardized assessments was also discussed consistently with the teachers. In New Jersey, the 2017-2018 start date for Grades 3-8 standardized testing is April 16. Teachers must ensure all skills required for testing proficiency are taught and practiced in the classroom before this date. There are approximately two months left in the school year when the assessments begin, which gives teachers a little more than seven months of school to prepare students for an assessment that will weigh 30% of their effectiveness. The skills required in the New Jersey Student Learning Standards are framed to meet a 10-month school year, which is reflective of the developed curriculum for each local school district. Schools must decide how to pace their curriculum based on skill strengths and weaknesses with the knowledge that they must eventually move forward in order to meet deadlines for testing. When pacing falls behind and cannot meet all of the standards that might be presented on the tests, decisions of content focus, categorizing skills, and eliminating those not as frequent or prevalent on past assessments must be made. This does not indicate that these skills are not presented, taught, and practiced after the standardized assessments are taken. This does indicate that those skills are not taught prior to students being assessed.

When New Jersey moved to the PARCC assessments in 2014, all standardized assessments in the state moved to a computerized format. This was a major caveat that brought tremendous logistical concerns to those responsible for ensuring testing protocols and test administration. Logistical issues were mostly due to the lack of devices and the lack of adequate bandwidth for all students to take the test at the same time. This required schools to create
rotating schedules for all students to be tested on the same limited number of devices and forced teachers to shut down instructional technology during testing windows. Prior to 2014, New Jersey utilized the NJASK assessments, and all students took the test together with paper and pencil. The creative schedules developed have students in the same grade level taking assessments at different times of the day (morning and afternoon), different days of the week, and in different weeks altogether. This had the capability of affecting a school’s master schedule for up to three weeks in one instance. Two of the three schools have resolved a lot of the logistical scheduling issues by purchasing Google Chromebooks and, for the first year, are able to complete testing in approximately six school days.

Bandwidth infrastructure was another logistical issue presented with the implementation of the PARCC assessments. All three schools, in the past, made it a mandate to shut down any computer or internet use in the classrooms during PARCC testing administration. The PARCC software that delivers the assessments to students requires bandwidth without interruption. In order to be pro-active and accommodate for this concern, technology for purposes other than PARCC were not permitted. One teacher commented on how this greatly affects classroom instruction and the lessons delivered. One of the schools, as reported by two teachers, upgraded its internet bandwidth, which will allow for more students to test at the same time. One school is in the first year of a 1:1 Google Chromebook initiative, which will allow all students to test at the same time. This also required an upgrade in bandwidth available. All of this preparedness and infrastructure planning comes with costs towards testing that are not adequately offset by state and federal funding (Baines & Stanley, 2004).

With the implementation of a computerized assessment, students also needed to know the technology skills necessary to navigate through the assessment, which include turning pages
forward and backward, highlighting text, expanding answer boxes and eliminating answers. Teachers overwhelmingly reported that this was a concern when PARCC was first delivered. However, students have now been taking a computerized assessment for nearly four years at the time of this research. They are described as very tech savvy and able to navigate the PARCC format quite easily. By the time they get to the middle school level, they have experienced sitting for this test in the previous school years. Two teachers raised a concern related to typing speed in a timed setting. Students now use software for benchmarking and guided practice that utilizes the same technology skills the students see on the PARCC assessments.

Overall, teachers interviewed did not feel that standardized assessments take away the art of teaching and designing lessons for students. Regardless of standards or curriculum framework, teachers need to be creative in order to make lessons engaging, interesting, and fun for their students. Without this practice, instruction becomes a highly restrictive and repetitive practice. Freedom of choosing content and delivering that content are the primary sources of creativity and engaging designs of classroom lessons. This differs in language arts and mathematics classrooms. Math teachers are more reliant upon the structure of the textbook and the sample equations provide. Language arts teachers circle similar skills across varied content that includes informational text, short stories, novels, and writing prompts. This content is chosen locally and embedded into the scripted curriculum that reflects the state standards.

**Summary of Findings for Research Question 1**

Research Question 1 focused on curriculum development and the effects of standardized assessments on the process of development. Teachers were asked how much input they have in the developmental process of curriculum, how the content for curriculum is identified, and whether or not they use formalized benchmarking programs to predict testing results. Predicting
test results based on instructional practices amounted to an adjustment of standards focus within the curriculum and dictated the pacing leading up to the assessment dates. Strengths and weaknesses of students are identified by looking at overall school reports and individualized student reports from the previous year’s PARCC results. All three schools reported, in both language arts and math, they assemble at least once at the beginning of the school year in order to examine students’ results and decide which skills will be major points of focus. Pacing of the curriculum was a major theme in this process. Teachers are using this process to determine essential and non-essential skills for the current school year’s respective students. Knowing there are time restraints due to April testing dates, it is possible for standards to not be covered through instruction but presented on the standardized assessment.

Although there is no input or feedback into the skills required at the end of each grade level, teachers have freedom of choosing content that fits into the standards required. Content varies from school district to school district, but similarities in standards and skills required do exist. For example, if middle school students are in a Pre-Algebra course, they will most likely be taught the same equations and formulas across schools. But they might have different textbooks, engage differently with technology, work more or less frequently in groups or have different projects and assignments. As related to language arts teachers, the variety of content across schools is greater and more subjective. Language arts teachers get to pick their own informational and narrative texts for reading and topics of relevance to middle school students for writing. Teachers overwhelmingly felt as though this brought across creativity and the art of teaching in a highly prescribed set of standards.

It is important for authentic formative and summative assessments to be written into the curriculum (Wiggins, 1989) so that language arts and math teachers in the same school are
assessing students with the same methods. Formative and summative assessments developed by
the teachers are the main sources for identifying student strengths and deficiencies. Teachers,
through this data, know when to “spiral back” in the curriculum to revisit standards not mastered.
This requires a near-matched pacing of the curriculum and collegial collaboration to ensure
pacing occurs and students across grade levels are learning the same skills and concepts.

In order to assist with the preparation and practice for standardized assessments, school
districts vary in the types of formalized programs used by teachers and students. Formal
programs and practice can come from textbook supplements, workbooks, computer software, and
PARCC and State Department of Education websites. Schools also have the opportunity to
purchase benchmarking software that has the intention of predicting where student deficiencies
might show on the standardized assessments. These formal programs are an added expense to a
school budget and weigh into decisions of curriculum development and pace. The three school
districts used in this study all use resources that simulate PARCC questions for the purpose of
testing practice. The schools do not utilize a long-term, formal benchmarking system to
consistently track growth in each of the standards. One of the schools uses a formal
benchmarking system but only once in the beginning of the school year. Data received are used
to compare scores and student ability tracking. It is also used to determine starting points for
review of skills lost over the summer months.

**Summary of Findings for Research Question 2**

Research Question 2 focused on effects of standardized assessments on instructional
design and content delivery with the knowledge that 30% of a year-end, summative evaluation
will be weighted towards standardized assessment results. Pacing of instruction presented itself
as the overwhelming effect of standardized assessments on instruction. Teachers feel there is not
enough time to “spiral back” as much as is needed for skill mastery. The state assessments in New Jersey are administered starting in mid-April, and teachers must fit a large amount of content into a timeframe of a little over seven months in order to feel as though students are properly prepared for PARCC. This creates pressures to stay on pace and cover all of the prescribed material regardless of 100% skill mastery. Teachers feel there is more emphasis being placed on ensuring presentation of all material rather than a clear focus on differentiating instruction so that students have the ability to master their individualized deficiencies.

Differences in philosophies of what determined standardized assessment success emerged when New Jersey stopped using NJASK and started using PARCC to deliver assessments. This change now had New Jersey educators measuring growth of each student from school year to school year rather than defining a pre-determined proficiency level for the entire school (NJDOE, 2013), which altered strategies for instruction. This testing modification ran parallel to changes in federal direction from No Child Left Behind (2001) to the Every Student Succeeds Act (U.S. Department of Education, 2015). One teacher interviewed in this study commented in-depth on this theme and gave insight into the “growth model vs. proficiency model.” This teacher thought positively of the growth model since it takes into consideration ability levels of students from previous school years. However, this teacher still made mention of factors not being considered such as socioeconomic status and limited English proficiency.

Factors such as pacing of instruction and focusing instruction on deficiencies of standards forces teachers to remove content that might have been taught prior to PARCC. Pressures now exist to cover a prescribed amount of material prior to testing, and this means teachers are attempting to ensure mastery of skills prior to test administration. One of the teachers interviewed mentioned the removal of some projects that used to be assigned because there is
simply no time to stop and focus on one topic for more than two or three days. The most memorable quote to this researcher came from Teacher 4 when it was stated, “We are always trying to chase after results.” Placing evaluative weight on standardized assessments forces teachers to meet deadlines and move forward with material. This happens regardless of whether or not students have the foundational skills to move forward to the next concept.

Creativity in delivering lessons is overwhelmingly not affected by the teachers selected for this study. All 12 teachers interviewed agreed that lessons must be fun and engaging for the students to learn. This is the art of teaching that must come out when teaching to the standards. Teacher 10 made mention of continuing to do a stained-glass project in order to teach percentages and fractions. All language arts teachers made mention of being able to select their own content for informational text, short stories, novels, and writing prompts. This allows for a greater amount of creativity than mathematics teachers who reported on following the structure of the textbook and using many of the sample equations and multi-step questions. Two math teachers in different schools reported starting with Chapter 3 of the textbook in September in order to have time to cover all material prior to the date standardized assessments are administered. Common Core Curriculum Standards (2010) narrowed the skills required, which meant eliminating content from local curriculum.

Grades 3-8 language arts and mathematics teachers in New Jersey face 30% of their summative, year-end evaluations being tied to PARCC results. This was a grand topic of conversation for all of the teachers interviewed. All 12 teachers mentioned the word unfair and spoke about the inhibiting factors that are overlooked by AchieveNJ (NJDOE, 2013) legislation. When weighing results of standardized assessments, teachers mentioned that the test is a “snapshot of time” within the school year and does not account for major reasons why students
might not score proficiently on assessments. Although testing can take up to three weeks for a respective school, the timeframe for each student averages three days total for each subject area. This information sheds light on the fact that almost one-third of a teacher’s evaluation is tied to these three testing days.

Teachers overwhelmingly focused on factors of demographic data, socioeconomics, transience, ability levels, learning disabilities, limited English proficiency, testing indifference, and apathy as primary to predicting testing results (Maylone, 2002; Turnamian, 2012). Teachers studied felt that the 30% mandate of tying evaluations to testing results causes undue pressures to teach to the tested material and a lack of stronger focus on standards that students need more time to master. When teaching, this mandate is thought about by the teachers, and methods of playing to the “numbers game” (Teacher 4) become a focus and teaching strategy. Concerns of fairness for lack of a 30% mandate on other subject areas also surfaced with teachers in this study; however, language arts and math teachers understand why skills of reading, writing, and mathematics are singled out and weighed as the most critical in a child’s education and career preparedness.

Standardized testing also does not focus on the social/emotional growth and successes of students (Niesz, 2010). Teachers in this study spoke about themes of social and emotional growth and how standardized assessments force us to spend quality time focused on the development of a highly structured, quantitative manner of evaluating student achievement. The qualitative measurements of educating a child, such as long-term focus on cooperative projects and peer interactions, are not valued or weighted in standardized assessments (Nichols & Berliner, 2008). The growth measurements outlined in No Child Left Behind and Every Student Succeeds take only quantitative measurements of assessment results into consideration when
evaluating accountability and effectiveness. In order to prepare a child for the world’s challenges, schools today must educate the “whole-child” and prepare him/her for contentment by teaching healthy behaviors and habits in their lives (Center for Disease Control, 2018). This weighs into the notion of “College and Career Readiness” as mentioned in the PARCC acronym. Teachers know that preparing for college and career has many subjective, social aspects to success outside of testing results and a prescribed standard for skills to be taught.

Summary of Findings for Research Question 3

Research Question 3 focused on the change to a computerized assessment and whether or not instructional time is being taken away from language arts and math skills so that students are deemed proficient with technology skills to navigate the testing format. Sixteen different technology skills were mentioned by the teachers interviewed as key to navigating through the PARCC computerized assessment and answering questions correctly. The most significant skills having the ability to impact answer choices included typing fast enough in a timed setting, expanding text boxes to write more, turning pages forward and backwards to skip and go back to questions, dragging and dropping answers into a text box, underlining and highlighting evidence in text, and using digital tools to replace hand-held tools such as calculators and rulers.

Teachers felt that the students have adapted well to these skill requirements and are now familiar with the testing navigation. By the time the students arrive to middle school, a large majority of them have used software programs that mimic PARCC questions and have taken the PARCC assessment in Grades 4, 5, and 6. All three schools chosen for this study use Google Doc software to receive and deliver assignments with students, and this helps with typing skills practice and practice reading information on a computer screen. Overall, a significant amount of time is not being spent teaching technology skills, and teachers focus only on getting to
computers specifically for PARCC practice approximately 2-3 weeks prior to the testing dates. Specific time spent for each of the students on computers for PARCC practice ranged from one time to six times in the 2-3 weeks prior to testing. Students were often mentioned as being “digital natives” and “tech savvy.”

School schedules are impacted greatly by PARCC testing and the shift to a computerized testing format. There are not enough devices for all students to test at one time, which forces school administrators to develop creative methods of time and classroom management. Students are not testing within the same structure of time and place, but standards are being created for growth. There is evidence to suggest that students perform better academically in the morning hours (Pope, 2016). This information suggests that a student who takes a standardized assessment in the morning hours has the potential to focus and comprehend questions better than a student taking the assessment in the afternoon. Schools are also testing students in different total amounts of days that are either much closer together or more spread apart. The lack of devices for all students to take a standardized assessment at one time can have a direct effect on results respective to each child and when the assessment is administered. Teachers from two of the schools participating in this study stated that things are getting better with experience and purchases of additional devices and greater bandwidth capabilities will help to get testing done in 1-2 weeks in the future.

Bandwidth infrastructure is also affected by a computerized standardized assessment. Teachers in this study were often asked not to use technology requiring bandwidth during testing hours. This was designed to prevent software glitches in the testing navigation websites so that student results are properly recorded and saved. Two schools reported a total district shutdown of technology use for approximately 5-6 weeks due to infrastructure concerns. This has the
potential to vastly impact 21st century learning goals and forces teachers to redesign lessons to reflect a lack of availability with classroom technology. Schools are being forced to increase bandwidth and the number of devices available to students. This can be a large cost for smaller school districts that have smaller budgets to accomplish these goals.

**Limitations of the Study**

Middle School language arts and mathematics teachers are the only chosen subjects since they are the only two content areas in Grades 3-8 tying high-stakes test scores to year-end, summative evaluations. The findings for this study cannot be assumed for school districts outside of New Jersey. Similarly, no findings can be related to any grade levels other than 7 or 8. The results of this research apply only to similar themes arising from the teacher interviews. No specific PARCC data were garnered to correlate with or support the problem. The data gathered for this research are also a “snapshot” of time, documenting the transformation of evaluative systems based on student test scores in the 2017-2018 school year. It is expected that all participants in this study responded honestly about their perceptions of standardized testing effects on development of curriculum and delivery of instruction. No results garnered for this research can be assumed outside of middle school language arts and mathematics classrooms in New Jersey. Many of the testing references, mandates, and legislation directly affecting teacher evaluations are solely specific to New Jersey’s Department of Education.

**Findings Related to Theoretical Framework**

All teachers selected for this study had opportunities to reflect upon the open-ended questions posed. Time was given for them to respond openly and honestly about how a culture that stresses the importance of standardized assessments directly impacts decisions they make with development of curriculum and delivery of instruction. This opportunity allowed the math
and language arts teachers selected to reflect upon their own practice and better understand any impacts that a 30% standardized assessment growth mandate has on their measures of effectiveness and levels of accountability. Reflection on an educator’s practice allows time to develop self-awareness and recognize opportunities for professional growth (Osterman & Kottkamp, 1993).

All 12 teachers interviewed openly displayed a willingness to be skeptical of positive effects of a 30% relation of standardized assessment results to a formal, summative evaluation. All of the teachers interviewed were highly skeptical of this practice and felt that this was the overwhelming aspect of standardized testing cultures that needed to change. Standardized assessments were determined to be an accepted part of the public education process and one objective method by which to gauge strengths and weaknesses of skills. Teachers interviewed were highly ethical and held themselves most accountable for reading, writing, and math skills taught to students.

**Recommendations for Practice, Policy, and Future Research**

**Recommendations for Practice**

This study asked middle school language arts and mathematics teachers to reflect upon the standardized assessment culture weighing heavily in public schools today. Since language arts and math teachers in New Jersey have had a portion of their effectiveness tied to standardized assessments for over four years, this practice is well known to math and language arts teachers and has had an observable effect on development of curriculum and delivery of instruction. Although there is no quantitative method to measure any effects, noticeable changes have been made to curriculum, and classroom time is being spent teaching to a testing structure and PARCC-like questions. Quantitative data, having the potential to emerge, can focus on a
longitudinal study of teachers’ summative evaluation scores over the timeframe of AchieveNJ (2013) mandates. Since the evaluation percentage has fluctuated from 10% to 30%, there is potential interest to see if a teacher’s effectiveness score went up or down parallel to the percentage changes. Teachers can track their summative evaluation scores over the last four years to determine if significant fluctuations have occurred.

In practice, it is important for schools to ensure that curriculum is geared to the standards. The standards provide a framework for expectations on standardized assessments. In New Jersey, the Student Learning Standards directly correlate with skills deemed necessary on PARCC. Teachers need to have the professional time allocated to become highly familiar with the standards and the expectations of skill development. Teachers also need to be involved in the content selection and pacing schedules of the curriculum. This allows for greater creativity in the classroom and empowers teachers and students to better reflect themes associated with classroom demographics, ability levels, and student age-groups. A strong understanding of the standards with appropriate content will prepare students for success on standardized assessments. Since pacing of curriculum and instruction was raised as a major theme in this study, teachers need time to meet during the school year in order to ensure pacing is occurring and that formative and summative assessments are being examined for validity of skill assessments.

**Recommendations for Policy**

Recommendations for policy will solely focus on state mandates of tying standardized test results to a teacher’s evaluation. Currently in New Jersey, testing weighs 30% of a Grade 3-8 language arts or mathematics teacher’s evaluation. Legislators should understand that there is more qualitative data to be garnered from schools than quantitative data. Social and emotional factors of growth and success are not accounted for in a school’s report card. School report cards
heavily weigh into quantitative factors such as test scores and attendance rates. Schools are solely being judged by legislators by annual standardized assessment results. These results are being used by legislators to measure worth of communities and neighboring school districts. There is no argument that standardized assessments hold value in objectively measuring a child’s skill deficiencies and can be used to measure annual growth. However, many factors such as socioeconomics, English proficiency, and testing apathy are not weighed at all into the equation. Accountability can be determined with factors other than standardized assessments. Formative assessments and benchmarks are two manners in which teachers can measure growth throughout the school year and better target skill deficiencies when they become present. Teachers are professionals and should be relied upon to be the primary measurer of student growth and success. Legislators who mandate accountability based on standardized test results are recommended to remove this level of influence on measuring teacher effectiveness. This recommendation places heavy responsibility on local policy and oversight.

**Recommendations for Future Research**

This study was focused on the effects of a standardized testing culture on development of curriculum and delivery of instruction. The study specifically focused on 12 New Jersey language arts and math teachers in Grades 7 and 8 who currently have 30% of their summative evaluation tied to PARCC results. Future research can expand outside the middle school grade level and focus on Grades 3-6 language arts and math teachers, who also have a percentage of their evaluation tied to test results. Future research can also focus on how accountability is decided upon for teachers outside of the language arts and math content areas to determine whether or not this method of evaluation can be associated with reading, writing, and math. Future research can also focus on school systems not using standardized assessment growth to
weigh teacher effectiveness and make qualitative comparisons to educational practices in
development of content and delivery of instruction.

A move to computerized assessments has also been a major shift in the testing culture of
schools. Research can be expanded to determine how schools with a lack of electronic devices
fare with schools operating under a 1:1 initiative. Factors for study can include master
schedules, time testing for the entire student body, and purchases made in hardware and
bandwidth to keep up to par with 21st century demands of testing.

Conclusions

Standardized assessments are accepted by educators as an objective means of evaluating
skill strengths and weaknesses of students. The skills required are tied heavily into the adopted
standards for each respective state. In order to create a perception of accountability and
demonstrate efforts towards demanding effectiveness, politicians are weighting standardized
assessment results into teacher evaluations. This is evidenced with Governor Chris Christie’s
signing of AchieveNJ (2013), which directly tied testing results to Grades 3-8 language arts and
math teachers’ evaluations. These pressures placed on “one shot,” high-stakes tests directly
impact the development of local curriculum and the delivery of instruction to students. Affected
most heavily, is the pacing of curriculum and instruction and the pressures to cover prescribed
amounts of skills and content prior to testing dates. Teachers feel there is not enough time to
focus fully on skill deficiencies before introducing new, more complicated concepts.

Students are tech savvy as digital natives; therefore, there are no anxieties of having to
teach technology skills to master navigation of the computerized testing format. As schools
become more experienced with administering computerized assessments, scheduled timeframes
for testing have improved and schools have invested in more technology infrastructure. Students
are also more experienced in their approach to computerized classroom formats and get time to practice with supplemental software and websites that offer PARCC-like questions and equations.
References


101


104


