

National Louis University

Digital Commons@NLU

---

Dissertations

---

5-2021

## Are All These Elementary Assessments Necessary? Phenomenological Study of Standardized Assessments in the Elementary School Setting

Cathy Wright  
*National Louis University*

Follow this and additional works at: <https://digitalcommons.nl.edu/diss>



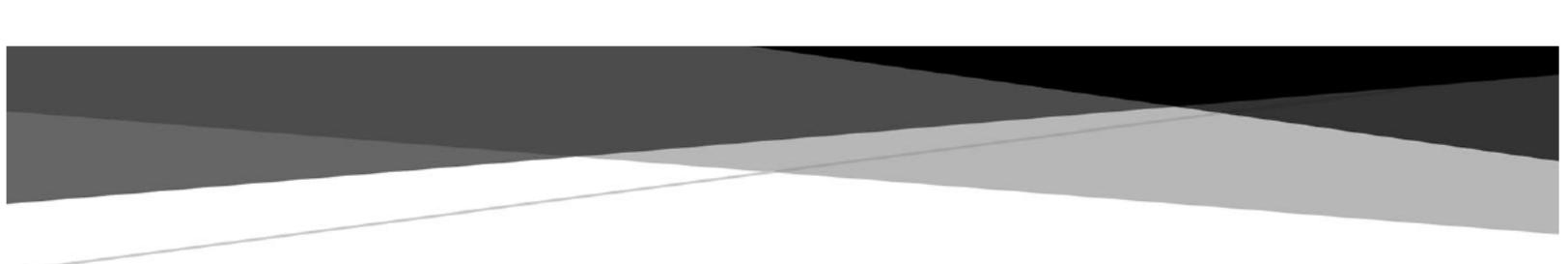
Part of the [Curriculum and Instruction Commons](#), and the [Educational Assessment, Evaluation, and Research Commons](#)

---

### Recommended Citation

Wright, Cathy, "Are All These Elementary Assessments Necessary? Phenomenological Study of Standardized Assessments in the Elementary School Setting" (2021). *Dissertations*. 596.  
<https://digitalcommons.nl.edu/diss/596>

This Dissertation - Public Access is brought to you for free and open access by Digital Commons@NLU. It has been accepted for inclusion in Dissertations by an authorized administrator of Digital Commons@NLU. For more information, please contact [digitalcommons@nl.edu](mailto:digitalcommons@nl.edu).



Are All These Elementary Assessments Necessary?  
Phenomenological Study of  
Standardized Assessments in the Elementary School Setting

Cathy Wright

Curriculum, Advocacy, and Policy

Doctor of Education

In the National College of Education

Submitted in partial fulfillment

Of the requirements of

Doctor of Education

National College of Education

National Louis University

Spring 2020



Are Elementary Assessments Necessary? Phenomenological Study of Standardized Assessments  
in the Elementary School Setting

Cathy Wright

Teaching and Learning - Curriculum, Advocacy & Policy (CAP)

Approved:

*Sara Efron*

Chair, Dissertation Committee

*Kamau Rashid*

Member, Dissertation Committee

*RRM*

Dean's Representative

*James Smith*

Director, Doctoral Program

*N. Miller*

Dean, National College of Education

05/20/2021

Date Approved

## Abstract

All children warrant a rich, meaningful public education that prepares them for the opportunities, responsibilities, and challenges that await them as contributing members of a democratic society and a global economy. Based on testing calendars, teacher testimonials, and applied research from other state and local mandated testing studies, this study found that the length of time students spend taking and preparing for the test is extensive. CBS News (2015) noted between 20-25 hours are spent by each student taking standardized assessments, according to a comprehensive report of 66 of the nation's large-city school districts by the Council of the Great City Schools. Standardized assessments measure about 2.3% of the classroom period for the regular eighth-grader in public school. Between pre-K and 12<sup>th</sup> grade, administrators removed 112 mandatory standardized exams. According to the American Federation of Teachers (2013), educators spend countless hours preparing curriculum, which is becoming lost instructional time (at \$7.15 per hour, equivalent to the per-student cost of adding one hour to the school day). The estimated annual testing cost per student starts at \$700, depending on the grade level, it may increase to \$1000. Less testing would free up funds for more classroom support materials as well as instructional programs. Monies could also support teachers with align classroom curriculum and standards to provide educators with data to drive their instruction.

*This is the value of the teacher, who looks at a face and says there's something behind that and I want to reach that person, I want to influence that person, I want to encourage that person, I want to enrich, I want to call out that person who is behind that face, behind that color, behind that language, behind that tradition, behind that culture. I believe you can do it. I know what was done for me.*

—Maya Angelou

## **Acknowledgments**

A congregation of dedicated professionals, whom I consider the mentors that have shaped me to be a very profound person throughout this process. Without them, I would not have changed my lenses of thinking, processing, and researching words presented in this phenomenon. I want to extend my sincere gratitude to Dr. Efrat, Director of Curriculum, Advocacy, and Policy at National Louis University. You will never understand, nor do I have the words to express how grateful I am for your prompt feedback, gentle guidance, patience, and dedication you showed me. You will always hold a special place in my heart. Dr. Rashid, your good conversations and knowledge across the doctoral program have been a treasured gem. Dr. Elkordy, you rescued me through your gentle guidance in dissecting survey data. You gave me confidence and showed me how to make my topic relevant through data collection. A special thank you to Dr. Terry Smith for her monthly doctoral meetings that kept me on track and provided a plethora of resources during the dissertation journey.

Also, I was part of the phenomenal National Louis support group (Dis Crew). As deaths, a pandemic, and other life challenges happened, we never stopped supporting each other. I did not take off my shoes.

This journey began because I remembered my late father, George Walker, telling me that I would go far in school. My late mother recently passed, saying how proud she was about my journey but could not hold on anymore. My heart aches for them both, but I know they are in a better place. Also, I was blessed with one brilliant brother and five sensational sisters who have been very supportive during this long journey.

Working full-time and pursuing a doctorate can be overwhelming, but Mr. Prado, Ms. Shimkus, my twin/Yeppez, and Ms. Torrez, you made my teaching days full of joy. Keeping me on task, printing documents, and sharing a cocktail or two released so much day-to-day stress.

“The Girls” you have been a sisterhood that everyone wishes they had. I will never forget how you encourage me to travel to a foreign country to work on my dissertation, and it was just what I needed to gain new insight. Your support has made my good time better and my hard times easier.

There were so many sacrifices made during this journey, but the Six Pack stepped in to make sure I maintain direction and a peaceful mind. Dickey, Lisa, Monique, Forrest, you are my guardian angels. When times were hard, I remember Lisa giving me a theme song, “Get your blessing” by Mary Mary. She followed up with Go get it, girl; I knew she meant my doctorate. Being a woman of faith there were so many times Monique offered or I requested her prayers of conviction to sooth my soul. Those prayers offered me safety, peace, and comfort in times of uncertainty.

Over twenty years ago, I left a high-paying job to become a teacher. Although I miss the perks, I thought I would lose the one thing that meant so much to me, My BFF Gail. Well, that did not happen. She has been my ride-or-die for more than two decades. She makes me want to be a better person. She makes me feel inspired, and we inspire each other to do better.

I saved the two most important ones for the last. Aundra Sr. and Aundra Jr. have missed meals, I have missed special events, and you have been my rock. Your encouragement gave me strength, your love gave me power, and your patience saw me through. You lifted me when I felt that this journey was too much; you loved me during my snappiest.

Baby Dre,

“No one else will ever know the strength of my love for you. After all, you are the only one who knows what my heart sounds like from the inside.”

Wright,

Throughout all of life's troubles, you have been there without fail. Life with you has indeed become a beautiful fairy tale.

## TABLE OF CONTENTS

Acknowledgments	iv
List of Tables	xii
List of Figures	xiii
CHAPTER ONE: INTRODUCTION	1
Purpose Statement	3
Significance of the Study	4
Definition of Assessment	9
Historical Timeline	11
The Nineteenth Century	11
The Twentieth Century	12
Assessment from the Perspective of Positivist Theory	19
Seminal Assessment Theoreticians	19
Franklin Bobbit	19
Edward Thorndike	20
Lewis Terman	22
Usage of Assessments	22
Political Assessment Analysis	24
My Assessment Vision	25
Concluding Insights	26
	6

Organization and Structure of The Remaining Chapters	27
CHAPTER 2.	29
LITERATURE REVIEW	29
Assessment Justification	30
Assessment Accountability	33
What Assessment Tests Entail	34
Formative Assessment	39
Summative Assessment	41
Assessment Paradigms	43
Post Positivist Paradigms	44
Pragmatic Paradigms	45
Constructivist Assessment Paradigm	46
Transformative Paradigms	48
Students' Perception of Assessment	48
Criticism of Assessment Test	52
Teaching to the Test	52
Validity and Reliability	53
Assessment's Price Tag	54
Concluding Insights	56
Divulgence Into My Research	57

CHAPTER 3.	59
METHODOLOGY AND RESEARCH DESIGN	59
Critical Stance	63
Research Design	64
Research Participants	64
Sources of Data and Methods of Data Collection	65
Quantitative Data	66
Accounts of My Childhood Experiences	67
Participants	68
Interviews	68
Interview Analysis: What It All Means	70
Observations	70
Triangulation	71
Ethical Consideration	72
Summary of Methodology	74
CHAPTER 4. DATA ANALYSIS	76
Data Collection and Analysis Process	76
Data Retrieval	(-)76
Interview Process	77
Data Analysis Process	78

Interview with the Participants	79
Interviews Summary	87
Observation Data Collection & Analysis	88
Classroom Visit Observations	89
First Observation	90
Second Observation	91
Third Observation	92
Observation Summary/Analysis	94
Survey Data Analysis	96
Survey Analysis of Themes	97
Reduce the Number of Assessments	98
Lost Instructional Time	99
Student Anxiety	100
Less Weight on Standardized Test	101
Chapter Summary	105
CHAPTER 5.	107
FINDINGS AND CONCLUSIONS	107
Data Interpretation	110
Reflections on the Findings	117
Revisiting and Answering the Research Questions	121

Question I– What are all of these standardized assessments needed for?	122
Question II.	123
How do educators perceive standardized assessments?	123
Validity and Trustworthiness of the Findings	126
Triangulation	127
Member Checking	129
Practical Implications of the Study	131
Research Journey	132
Educators' Aspirations and Final Reflections	133
REFERENCES	140
APPENDIX A	149
Appendix B	151
APPENDIX C	154

## List of Tables

Table 1. Themes and Descriptions

98

## **List of Figures**

Figure 1. Student Assessment Information	5
Figure 2. Elementary School Assessments Within a Quarter	89
Figure 3. Instructional Time for Prep	102
Figure 4. Limitations of Standardized Assessments	103
Figure 5. Teacher Created Summative Assessments	104
Figure 6. Teacher Created Summative Assessments Pt. 2	104
Figure 7. NWEA Assessments	106
Figure 8. IAR Assessments	106
Figure 9. REACH Assessments	107
Figure 10. Project-Based Learning Culture	138

## CHAPTER ONE

### Introduction

When I decided on a career change 15 years ago to become a teacher, I had it already set up in my mind that I would enrich, engage, and enlighten students with a fun and contagious way to learn. Bloom (1956) believed that the way we question or formulate students' tasks could be varying levels of complexity. Teachers can use this taxonomy to address, connect and engage students in topics at varying levels (Konen, 2017).

My parents were having marital problems, and because of that, I was uprooted from a northern state to a southern state. I clearly remember the awkwardness of fitting in socially and educationally. It felt like I was being measured on everything. Coming from a northern state, the school/teachers' message I felt was that I should have known more. Based on my standardized test scores, they put me in low-level classes. Separating myself from other students made me feel dejected. I was being labeled as a slow learner in a new environment crushed my energy to learn.

While being continuously assessed as a student in the southern state (which my parents were unaware of), I lost so much of my joy of learning. I recalled that while receiving my low test scores, the teachers made me feel so belittled. I remember my elementary school-age years being nothing but Assessment after Assessment. My teachers never spoke to me about the findings, which was very frustrating as a student. No one knew that I suffered from test anxiety. What, I asked myself, was the purpose of all these tests? What did they mean for me as a student? After learning that I had scored lower than expected, I was moved to the back of the class and given work that kept me busy, not challenged. When I relocated back to Illinois, I was given the same test again, and their findings were quite different. The teachers placed me in my accurate grade, but I had to attend summer school. I believe this was when I knew that I wanted

to become a teacher. I promised myself that I would never let my students feel such negative emotions due to standardized assessments. Yes, I graduated from elementary school and found a high school that appreciated all of my attributes that make a person well-rounded. However, as I pursued college and my profession as a teacher, I found the assessment factor had become more intense. This intense use of assessments led me to ponder, “Do we need all of those assessments?”.

As I got older, I began to establish a love for teaching. I took the initiative to tutor younger siblings, created bulletin boards in our basement, and invited the neighborhood kids for a read-aloud. My actions exhibited behavior that expressed my feeling I was destined to be a teacher. I desired to be a teacher who would offer students the opportunity to exhibit more than a test.

I had a new vision in 2003 when I began my teaching career in education that would start with the recognition that teachers are passionate, skilled professionals whose focus is on effectively engaging students, ensuring their learning, and shaping their minds. I decided that I would be a teacher who knows how to inspire and lead change effectively while preparing students for the many challenges of their future. I know that teachers play a crucial role in creating successful education changes, and I resolved that as an educator. I would not accept failure within any circumstances; I concluded that I would remain positive, adaptive, humble, resourceful, creative, and always think about my students and search for ways to improve my teaching. More importantly, I would be an educator that would allow students to express and display ways to show their mastery of learned skills. Excessive testing that can cause embarrassment is not something that I want my students to endure.

I was very excited about securing my classroom. I had developed a scope and sequence that would allow students to engage and collaborate throughout the curriculum. I remember my

disillusion as though it was yesterday; the district had mandated two standardized assessments, Dibels and NWEA. The Dibels assessment had to be administered to each student individually. This assessment tool took anywhere from 10 to 30 minutes, and the NWEA was conducted over four days with a whole group. After securing a classroom teacher position in a large district, I realized that assessments are out of control, and I do not have much authority regarding their distribution.

In Diane Ravitch's (2010) book, *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*, the author explains how the United States is overruled by testing. Our educational system's drive to assess, label, and sort kids and make decisions based on dubious quality data has gone too far; it is time for a course of correction. According to the Oxford Bibliographies (Leeds-Hurwitz, 2016), social construction assumes that people construct their meanings that give encounters with others or various products they or others create; it also implies that they do this jointly in coordination others rather than individually.

Over the last few years, many mandates of standardized assessments have plagued the classroom, which has limited me as an educator to engage students regularly with an enriched curriculum. I have seen many educators leave the field or lose their passion for teaching because of all of the assessments they have to conduct that is locally and district level mandated.

### **Purpose Statement**

This phenomenological study aims to delve into and understand the effects of standardized assessments on elementary students. My interest in this phenomenon focuses on the loss of classroom instructional time among elementary students. The lost instructional time due to an abundance of standardized assessments and the ineffective ways of reviewing assessment data makes it tough to teach the necessary curriculum. The data gathered from standardized

assessments are supposed to provide teachers information, helping them view what they communicated well and what needs to be taught differently.

My goal is to determine how standardized assessments impact students' educational, emotional, and social growth. I aimed to contribute to making changes in the current approach, the number of standardized tests conducted during the elementary years, and how they should be changed.

### **Significance of the Study**

It has been noted that early every student has suffered the experience of spending hours preparing for a significant assessment to realize that the material that he or she had studied was different from what the standardized assessments chose to emphasize (Ravitch, 2010). This experience (Guskey, 2000a) teaches students two unfortunate lessons. First, students realize that hard work does not pay off in their learning environment because the time and energy that they spent studying had little or no impact on test results. Second, they learn that they cannot trust teachers. When students begin to dive into an assessment, many perceive that the curriculum obtained would allow them to succeed in their evaluations. When administering standardized assessments, educators are blindsided with what the assessment goals are for students. These are not the lessons that loyal teachers want their students to learn.

According to Valerie Strass from the Washington Post (October 24, 2015), on average, 1,110 minutes were dedicated to the standardized assessment process for the elementary students in reading and math grades 3-6, and 1,134 minutes were dedicated to this process for students in Grades 7-8. This amount of time is essentially doubling the standard set by the state legislatures. This amount is an underestimation of the time testing takes. For one thing, science testing, mandated for students in the 4<sup>th</sup> and 8<sup>th</sup> grades, is not included in the calculation. Also, the 180-

day school year is used as a basis for establishing instructional hours in a year, even though previous research demonstrates that students do not receive the assumed 180 days required for instruction and factoring in time spent on test prep and practice tests.

During their teaching career, I have taught grades 1, 2, 3, 4, and 8. The amount of assessments mandated in their state, Illinois, is extensive. As previously stated, as a new teacher, I have developed a scope and sequence of instruction, but when the district announced that on day one of school that I, as a teacher, must administer three standardized assessments within a two-month window, it became very challenging to teach the curriculum designed for the eager to learn students. My personal experience as a teacher is typical of what most classroom teachers go through. Figure 1 lists the types of assessments required for students in grades K-8, along with the allotted times that students need to complete the exam successfully.

**Figure 1**

Student Assessment Information

Assessment	Dates	Required?	School Type(s)	Grades	Approx. Time	Assessment Results Used For (Other than Guiding Instruction):
<b>REACH Performance Tasks BOY</b>	September 17 - November 2	Required	District Only	PK-12	40-60 mins per task	REACH Teacher Evaluation System
<b>Kindergarten Individual Development Survey (KIDS)</b>	Observation Window: September 4 - October 31 Data Entry Window: October 24 - November 7.	Required	District, Charter	K	Varies	

<b>TRC/DIBELS, mCLASS Math BOY</b>	September 4 - October 12	Additional (Recommended for “opt-in” schools*)	District Only	K-2	30 mins for literacy; Math varies	
<b>NWEA MAP for Primary Grades (MPG) BOY</b>	September 4 - October 5	Additional (Recommended for “opt-in” schools*)	District Only	K-2	45 mins per subject	
<b>Fountas &amp; Pinnell BAS BOY</b>	September 4 - October 12	Additional (recommended for “opt-in” schools*)				
<b>NWEA MAP BOY</b>	September 4 - October 5	Additional (Optional except for students without a prior year’s spring score who will test for SQRP baseline)	District, Charter	3-8	3 hours	School Quality Rating Policy (SQRP)  REACH Teacher Evaluation System  Selective Enrollment Schools Application
<b>Teaching Strategies GOLD Fall</b>	Finalize Checkpoint by October 12	Required	District Only	PK	Varies-individual	
<b>TRC/DIBELS, mCLASS Math MOY</b>	December 10 - February 1	Additional (Recommended for “opt-in” schools*)	District Only	K-2	30 mins for literacy; Math varies	
<b>Fountas &amp; Pinnell BAS MOY</b>	December 10 - February 1	Additional (Recommended for “opt-in” schools*)	District Only	K-2	Varies	
<b>NWEA MPG MOY</b>	December 10 - January 25	Additional (Recommended for “opt-in” schools*)	District Only	K-2	45 mins per subject	
<b>Assessment</b>	<b>Dates</b>	<b>Required?</b>	<b>School Type(s)</b>	<b>Grades</b>	<b>Approx. Time</b>	<b>Assessment Results Used For (Other than Guiding Instruction):</b>

<b>NWEA MAP MOY</b>	December 10 - January 25	Additional (Recommended)	District, Charter	3-8	3 hours	
<b>REACH Performance Tasks MOY Qtr 3</b>	February 4 - March 1	Required for teachers without a BOY score	District Only	PK-12	40-60 mins per task	REACH Teacher Evaluation System
<b>Teaching Strategies GOLD Winter</b>	Finalize Checkpoint by February 1	Required	District Only	PK	Varies-individual	US Department of Education Reporting
<b>Las LINKS</b>	Feb 25 - April 5	Additional (Recommended)	Dual Language Schools Only	3 - 8	115 min (3-5 grade)-125 min (6-8 grade)	
<b>State-Required Elementary School Reading and Math Assessment</b>	Dates are presently unknown, subject to ongoing procurement, and added as soon as they are available. Dates will not start before March 4, 2019, and will most likely not extend past April 26, 2019.	Required	District, Charter, Options	3-8	TBD	IL School Report Card
<b>Illinois Science Assessment (ISA)</b>	April 8 – April 26	Required	District, Charter, Options	5, 8	1 hour	IL School Report Card
<b>REACH Performance Tasks MOY Qtr 4</b>	April 8 - May 10	Required for teachers without a BOY or MOY Q3 score	District Only	PK-12	40-60 mins per task	REACH Teacher Evaluation System
<b>NAEP</b>	January - March (school-specific dates)	Required	District, Charter	4, 8 (selected schools & students)	90 mins	U.S. Department of Education Nation's Report Card
<b>STAMP 4Se: (5th Grade) 4S: (8th Grade)</b>	April 1-May 3	Additional (Recommended)	District, Charter, Options	5, 8	STAMP 4Se: 85 min STAMP 4S: 2 hrs	Pathway to Biliteracy (5th and 8th grade)
<b>TRC/DIBELS, mCLASS Math EOY</b>	May 6 - June 14	Additional (Recommended for K-2 "opt-in" schools)	District Only	K-2	30 mins for literacy; Math varies	

<b>Fountas &amp; Pinnell BAS EOY</b>	May 6 - June 14	Additional (Recommended for K-2 “opt-in” schools)	District Only	K-2	Varies	
<b>NWEA MPG EOY</b>	May 13 - June 14	Additional (Recommended for K-2 “opt-in” schools)	District Only	K-1	45 mins per subject	
<b>Assessment</b>	<b>Dates</b>	<b>Required?</b>	<b>School Type(s)</b>	<b>Grades</b>	<b>Approx. Time</b>	<b>Assessment Results Used For (Other than Guiding Instruction):</b>
<b>NWEA MAP EOY</b>	May 13 - June 14	Required**	District, Charter	2-8	3 hours	School Quality Rating Policy (SQRP)  REACH Teacher Evaluation System  Student Promotion (grades 3, 6, 8)  Selective Enrollment Schools Application
<b>REACH Performance Tasks EOY</b>	May 6 - June 7	Required	District Only	PK-12	40-60 mins per task	REACH Teacher Evaluation System
<b>Teaching Strategies GOLD Spring</b>	Finalize Checkpoint by June 14	Required	District Only	PK	Varies-individual	US Department of Education Reporting
<b>Algebra Exit Exam</b>	May 28 – May 31 (Charter/private school date TBD)	Additional (Recommended)	District, Charter	7-8	90 minutes	High School Credit  Course Placement

The information in Figure 1 is an eye-opener; it reveals that students are over-assessed.

The research aimed to not deter assessments, as assessments offer students viable feedback that allows them to make attainable goals for themselves. Assessments are also a critical factor in

students' learning. Students need to see how they are doing in class. Moreover, assessments allow students to determine course materials, understand learning, motivate, course learning being met, and reachable instructional needs. Assessments continue to evolve as policymakers turn to assessments to improve education.

Under the standards, assessments, and accountability, one finds 33 states with English Language Arts, 31 states with Mathematics, and 26 states with science have enacted a particular policy for their grade levels. The Chance for Success Index captures the importance of education in a person's lifetime. While early foundations and adult outcomes contribute to the index, formal education indicators are the driving force behind state ranking. The report represents state-specific summaries of key findings across all six policy and performance areas that comprise the report's state-grading rubric.

### **Definition of Assessment**

There are various definitions of assessments, such as educational Assessment, documenting, usually in measurable terms, knowledge, skills, attitudes, and beliefs (definition net, 2019). Upcraft and Schuh (1996) state that an assessment is an effort to gather, analyze, and interpret evidence that describes institutional, divisional, or agency effectiveness (p. 18). In education, the term "assessment" is often used to describe the measurement of what an individual knows and can do (Banta & Palomba, 2015). Definition Net (2019) defines educational assessment as the process of documenting, usually in measurable terms, knowledge, skills, attitudes, and beliefs. I embraced the definitions of assessment offered by the Education Reform (2015) that assessments are the wide variety of methods or tools educators use to evaluate, measure, and document the student's academic readiness, learning progress, skill acquisition, or educational needs.

One of the most dominant assessment tools in the United States is standardized assessments designed by testing companies and given to large populations of students designed, administered, and scored in a standard or consistent manner (Concepts, 2015). A standardized test is criterion-referenced, which means that students must meet specific criteria that are articulated. The standardized tests are administered in all states to comply with the state or federal mandates, require students to meet or exceed state standards as measured by test items (Efron & Ravid, 2020). Educators know the necessity of gauging students learning, and for this purpose, they use various assessment techniques throughout the school day.

#### Historical Context

In the early 20<sup>th</sup> century, public education held many changes rooted in the best thinking of the day (Edutopia, 2008). Several of these newfangled ideas focused on efficiency to mass-produce students who could read, write, and compute at a basic level. It was only sensible to develop standardized tests that could scientifically measure the “product” rolling off this educational assembly line (Boss, 2011). Fast forward to the 21<sup>st</sup> century, and basic literacy, numeracy, and content knowledge are no longer enough.

The roles and goals of standardized assessments have changed over time. To understand what educational assessments are, we must first have a clear understanding of their purpose and how they were historically defined and implemented.

The earliest evidence of standardized assessments was in China during the Han Dynasty, where the imperial examinations covered the six arts, which included music, archery, horsemanship, arithmetic, writing, and knowledge of the rituals and ceremonies of both public and private parts (Thomas, Murray, Hugh, Chambliss, Muhammad, 2017). In ancient Greece, Socrates tested his pupils over discussions. Responses were not scored as right or wrong (Matthews, 1994). They just proceeded to further dialogue. Numerous intellectual elites in the 5<sup>th</sup>

and 4<sup>th</sup> centuries B.C.E. cared more about discovering the route to higher knowledge than producing a correct response (Matthews, 2006).

According to the Center of Racial Justice Innovation (2013), in 1779, Thomas Jefferson proposed a two-track educational plan with different tracks, in his words, for “laboring and the learned.” Scholarships would allow very few of the laboring class to advance, Jefferson says, by “raking a few geniuses from the rubbish.” As the Industrial Revolution (and the progressive maneuver of the early 1800s that followed) clutched school-age children obtained through farms and factories and put them behind desks, standardized examinations emerged as an easy way to test many students quickly (Fletcher, 2009).

### **Historical Timeline**

Detected by the National Education Association (Alcocer, 2001), the standardized assessment timeline began in the United States in 1838. The American educators articulated ideas that would soon be translated into formal assessments of student achievement. Following this, I describe significant historical occurrences in that timeline.

#### **The Nineteenth Century**

From 1840 to 1875, there were established several main currents in the history of American educational testing, including formal written testing begins to substitute spoken examinations given by teachers and schools around the same time as schools altered their mission from servicing the elite to educating the masses. Pre-Civil War, schools used externally commanded printed examinations to assess student progress in exact curricular areas and to help in a variety of administrative and policy decisions. In 1890, Harvard President Charles William Eliot proposed a cooperative system of standard entrance examinations that would be acceptable

to colleges and professional schools throughout the country, in place of the separate examinations given by each school.

### **The Twentieth Century**

During 1900 – 1932, there were 1,300 achievement tests on the market compared to about 400 tests of “mental capacities.” This included 92 High school test, vocational tests, assessments of athletic ability, and a variety of miscellaneous tests are developed to supplement the intelligence test, and statewide testing programs were becoming more common.

In 1905, French psychologist Alfred Binet began developing a standardized test of intelligence, which would eventually be incorporated into a modern IQ version assessments dubbed the Stanford-Binet Intelligence Test. During World War II, standardized testing became typical practice: aptitude quizzes called Army Mental Tests were conducted to assign U.S. servicemen jobs during the war effort. In 1912, Stanford Professor Lewis Terman marked the beginning of large-scale individual intelligence testing in characteristics of the Binet-Simon tradition. Later, in 1914, the National Education Association endorses the kind of standardized testing that Rice, the first women officer elected to the National Education of Association, had been insisting for two periods. The scheduling was exquisite: on one front, there was the “push” of new technology that promised to be valuable to testing, and on the other, a heightened “pull” for methods to bring order to the chaotic schools. Finally, during 1908 and 1916, Edward Thorndike and his pupils at Columbia University developed a standardized achievement test in arithmetic, handwriting, spelling, drawing, reading, and language ability.

In addition, in 1916, the College Board begins to develop a comprehensive examination in six subjects. These examinations included performance types of assessments such as essay questions, sight translation of foreign languages, and written compositions. Shortly thereafter,

between 1917 and 1918, 6,500 children are given the Stanford-Binet. The new test was written by Arthur Otis (one of Lewis Terman's students who would eventually be credited with the invention of the multiple-choice format). Consequently, in 1917, the American Psychological Association recruits Lewis Terman and a group of colleagues to help the Army exams during World War I ignited the swiftest expansion of the school testing movement. By 1918, there are over 100 standardized tests created by various researchers to measure achievement in the principal elementary and secondary school subjects/

In the Fall of 1920, the World Book published nearly half a million tests, and by 1930 Terman's intelligence and achievement test (the latter published as the Stanford Achievement Test) had combined sales of some 2 million copies per year. John Dewey (1922) laments the victory of the testers and quantifiers with these words; "Our mechanical, industrialized civilizations concerned with averages, with percent. The mental pattern which mirrors this social scene subordinates education and social arrangements based on averaged gross inferiorities and superiorities.

In 1925 the U.S. Bureau of Education conducted a survey that showed that intelligence and achievement assessments are increasingly used to classify students. In 1926 the first SATs were administered and established as the Scholastic Aptitude Test by the College Board, a nonprofit set of universities and other learning establishments. The initial test lasted 90 minutes and consisted of 315 questions testing vocabulary and basic math knowledge and even including an early iteration of the famed fill-in-the-blank analogies. The test flourished and by 1930 assumed its now-familiar form, with separate verbal and math tests. In 1929, the University of Iowa developed the first vital statewide testing layout for high school students, directed by E.F. Linquist. By 1930 multiple-choice tests are strongly entrenched in the schools. Not surprisingly,

the quick spread of multiple-choice tests kindled conversations about their deficiencies. Critics accused them of encouraging memorization and guessing, representing “reactionary ideas” of instruction, but to no avail. Efficiency and “objectivity” are worn out.

Consequently, in 1935, high-speed computing was initially devoted to testing. Electronic data dispensation equipment was used to process massive numbers of the test. One statement showed that the cost of administering the Strong Inventory of Vocational Interest dropped from \$5 per test to .50 cents per test as a result of the computer. The in 1936, the first automatic test scanner is developed, a rudimentary computer called the IBM 805. It remained essentially the same until 2005 when the analogies were made away with, and a writing section was added. By the late 1930s, Iowa tests are being made available to schools outside the State; Iowa also introduced in 1958 computerization to the scoring of tests and production of reports to schools.

In 1960, standardized assessments were designed as those assessments in which the condition and content were the same for anyone taking the test, regardless of when, where, or by whom the test was given or graded (Popham, 1999). According to the National Education Association, Standardized assessments began in the U.S. from 1838 until the present. During this time, the assessments had a variety of purposes and designs. Standardization limits the quality and quantity of what is taught and learned in schools (McNeil, 2000).

In 1965, the elementary and secondary education act, in particular, unlatched the way for new and escalated uses of norm-referenced tests to gauge initiatives. In the 21<sup>st</sup> century, nevertheless, the SAT and the ACT were just part of a gauntlet of tests students may have faced before reaching college. The College Board also offered SAT II tests created for individual subjects from biology to geography. The marathon four-hour Advanced Placement examinations, which some universities obtain for students who wanted to opt-out of initial college-level classes

– remain popular: nearly 350,000 took the U.S. history A.P. test last year, the most popular subjects offered. There is also the PSAT, administered in the junior year as preparation for the full-blown SAT and as an assessment for the coveted National Merit Scholarships; the twenty-first century has the most students tested each year of grade school.

From the early 1980s to the present, the educational reforms can be traced almost as cause and effect to *A Nation at Risk*. One example is *The Imperative for the National Commission of Excellence in Education* (1983). This report sponsored by the U.S. Department of Education presented a rhetorical call to reform with its infamous assertion that “if a preceding unfriendly power had attempted to impose on America. The mediocre instructional performance that exists today, we might have viewed it as an act of war” (Reimers, & Villegas-Reimers, 2014).

In 1983, *A Nation at Risk*, the report of President Ronald Reagan’s National Commission on Excellence in Education, recommended: (1) a stricter set of academic basics for high school graduation, (2) higher standards for universities, (3) a more extended school year or school day, (4) merit pay for top teachers, and (5) more citizen participation.

Its notification is considered a landmark event in modern American Educational history. This report inserted that American schools were failing, and it touched off a wave of local, state, and federal reform efforts. The report’s study of an education system that had “lost sight of the basic purposes of schooling, and of the high expectations and disciplined effort needed to attain them” (p. 12) rallied for improvement advocates to press for stricter accountability measures, including increased testing.

C.H. Edson, the assistant professor of educational policy studies, University of Oregon, Eugene concluded that this report very closely resembled the famous report issued by the

Committee of Ten in 1893 (McDill, Natriello, & Pallas, 1985). In 1983, *A Nation at Risk*, designed by President Ronald Reagan's National Commission on Excellence in Education, recommended: (1) a stricter set of academic basics for high school graduation, (2) higher standards for universities, (3) a more extended school year or school day, (4) merit pay for top teachers, and (5) more citizen participation. This imperative for Education Reform is the report of American President Ronald Reagan's National Commission on Excellence in Education. Non-public school personnel dominated both 1893 and 1983 groups. Both reports presented recommendations that were intuitively based rather than based on empirical or evaluative data. Both groups recommended longer school terms. Both reports endorsed a philosophy of social Darwinism—survival of the academic fittest. One difference between the reports is that the Committee of Ten established the high school's academic, general, and vocation duration. In contrast, *A Nation at Risk* implied that high school should be an academically elite institution. Jehlen (2002-2019) interjects several assessments omitted from the National Education Association timeline, such as the *A Nation at Risk* (1983) report.

### **Twenty-First Century and No Child Left Behind**

In 2001 the primary law for K–12 general education reform was published. The No Child Left Behind (NCLB) educational reform is the expansion of state-mandated standardized testing to assess school performance. The No Child Left Behind law—2002 improved the Elementary and Secondary Education Act—successfully ascended the federal part in holding learning establishments responsible for student outcomes. The NCLB law—which developed out of concern that the American education system was no more extended internationally competitive—significantly raised the federal role in holding schools responsible for all students' academic progress. It puts a particular light on ensuring that states and schools boost the achievement of some students, such as English-language learners, students in special education,

children from low socioeconomic backgrounds, and children from diverse racial and ethnic communities, whose achievement, on average, trails their peers. States did not have to submit with the new necessities, but if they did not, they risked losing federal Title money (Klien, 2015).

Donor and Schockley (2010) express how this law brought Republicans and Democrats together to expand opportunities for American children of all backgrounds and provided all our children with the quality education they deserve while preserving local control. No Child Left Behind, stated the authors, introduced sweeping changes that have affected every public school in the United States. The most extensive program with NCLB, Title I, allocates federal financial support to local education agencies, primarily school districts, to address the needs of educationally disadvantaged children. Almost 30,000 schools in the United States failed to make adequate yearly progress under the NCLB Act in the 2007-2008 school year (Hoff & Manzo, 2007). Ravitch (2010) notes that the same children left behind in 2001 are still left behind in 2015. NCLB had not accomplished its stated goal of leaving any child behind.

In 2015, Congress proceeded with the Every Student Succeeds Act to replace NCLB. ESSA moved in the opposite direction—it seeks to pare back the federal role in K-12 education. Race to the Top was a \$4.35 billion United States Department of Education’s aggressive grant developed to spur and reward innovation and state and local district K-12 education reforms. Funded as part of the American Recovery and Reinvestment Act of 2009, it was announced by President Barack Obama and Secretary of Education Arne Duncan on July 25, 2009. States taking part in the grants were awarded points for enacting specific educational policies, instituting performance-based evaluations for teachers and principals based on multiple educator effectiveness measures tied to targeted professional development and feedback, adopting common standards. However, adopting the Common Core State Standards was

unnecessary; adopting policies that did not prohibit or effectively prohibit the expansion of high-quality charter schools, turn around the lowest-performing schools, and build and use data systems (Pulse, 2017).

The development of a Commission on the Future of Higher Education, also known as the Spellings Commission, was announced on September 19, 2005, by the U.S. Secretary of Education, Margaret Spellings. The nineteen-member commission was charged with recommending a national strategy for reforming post-secondary education, focusing on what colleges and universities are preparing students for the 21st-century workplace and a secondary emphasis on how well high schools support post-secondary education. In the study, released on September 26, 2006, the commission looks at four key areas: access, affordability (particularly for non-traditional students), the standards of quality in instruction, and the accountability of institutions of higher learning to their constituencies (students, families, taxpayers, and other investors in higher education). After the report's publication, implementation of its recommendations was the responsibility of the U.S. Under Secretary of Education Sara Martinez Tucker (appointed August 2006). We have arrived in the year 2010 where the Common Core State Standards are debuting. This educational initiative details what K-2 students throughout the United States should know in English language arts and mathematics after the school grade. These standards were sponsored by the National Governors Association and the Council of Chief state school officers and seeks dependable academic standards over the states as well as ensure that students graduating from high school are prepared to enter credit-bearing courses at two- or four-year college programs or to enter the workforce (Common Core State Initiative, 2019).

### **Assessment from the Perspective of Positivist Theory**

Underlining all these historical reform waves was the perspective of positivist philosophy. This philosophical theory established the most dominant approach to high stakes testing and standardized Assessment (Ritt, 2016). To understand the current discourse on Assessment and its practice, we should understand the positivist theoreticians' theoretical foundations in the turn and the first part of the 20 century. The positivist paradigm of searching for social reality is based on the conception that knowledge is objective and scientifically measured. One can best obtain an understanding of human behavior through scientific observation and reason. Stated differently, the only objective, observable facts can be the basis for Science (Ntgrty, 2016). According to the positivist paradigm, proper knowledge is based on senses and can be gained by observation and experiment. Positivist thinkers lean strongly on determinism, empiricism, parsimony, and generality (Ntgrty, 2016).

### **Seminal Assessment Theoreticians**

#### ***Franklin Bobbit***

According to John Franklin Bobbit (1918), the immediate effect of standardization, a university professor and writer whose work influenced the development of the curriculum field in the United States, Bobbit advocated for scientific method in curriculum making and making social efficiency the curriculum. Bobbitt's legacy pertains to four areas: First, he was one of the first American educators to advance the case to identify objectives as the starting point for curriculum-making. Second, his so-called scientific approach to curriculum-making served as a precedent for various educators during the next half-century in spelling out the procedures for designing courses of study. Third, along with other early-twentieth-century efficiency-oriented school reformers, Bobbitt decided that the curriculum should be differentiated into numerous programs, such as academic, preparatory, vocational, terminal, and students ought to be

channeled to these tracks based on their ability. Finally, Bobbitt, one of the First American scholars, developed the curriculum as an instrument of social control or regulation to address modern society's problems. He also viewed the curriculum from the point of view of social needs rather than academic study. He formulated five steps in curriculum making: (a) analysis of human experience, (b) job analysis, (c) deriving objectives, (d) selecting objectives, and (e) planning in detail the experience to attain these objectives (Null, 1999).

Bobbitt believed that the learning objectives, together with the activities, should be grouped and sequenced after clarifying the instructional activities and tasks. The curriculum that Bobbitt advocated included components of general education for all youth but was, for the most part, differentiated into several very particular vocational tracks. Determined no doubt by the then-popular mental testing movement, Bobbitt believed that schools should assign children to these specialized curricular tracks, based on assessments of their intellectual abilities, which fore-told their ultimate destinies in life ("Franklin Bobbitt (1876–1956) - Social Efficiency Movement, Bobbitt's Contribution", 2019).

### ***Edward Thorndike***

Edward Thorndike (1874 –1949) played another very influential role in developing standardized assessment tests. Thorndike was an educational psychologist who identified three learning laws: readiness, exercise, and effect. He also set the law of effect, which means that any behavior followed by pleasant consequences is likely to be repeated. Any act followed by unpleasant consequences is likely to be avoided.

Despite claims that I.Q. Tests such as the Stanford Binet and Wechsler batteries are objective. Thorndike's history of developing such tests reveals them to be inherently culturally bound (Cope, 2016). According to Thorndike, tests, whose validity are based on teacher's

assessments, are necessarily bound by those teachers' cultural value. He laid to foundations for the current scientifically based measurement of student achievement. By the 1890s, most teachers were white, middle-class females. The standardized tests being invented were made so that the test results correlated with the assessments of white middle-class teachers of their middle and working-class students. Thorndike did not trust the teachers, but his scientific 'objective' measurements reflected his prejudice and cultural expectations (The National Academy of Sciences and Engineering Medicine, 2014).

Thorndike developed dozens of tests and became the leading theorist of connectionism active learning, an early form of behaviorist psychology. It was the perfect accompaniment for standardized tests and Snedden-driven curriculum theory. Connectionism, learning is the formation of a bond or an association between a stimulus and a response. Bonds are formed through two laws: Exercise practice and effect (reward). Since rote learning had been criticized for decades, Thorndike understandably emphasized the rewards or punishments that fleshed out the Law of Effect. According to this principle, the *desired effect* occurs primarily in response to the satisfying situation, while responses that produce an unpleasant effect are less likely to happen again.

### ***Lewis Terman***

Lewis Terman was an influential psychologist known for his version of the Stanford-Binet intelligence test and his longitudinal study of giftedness (Cherry, 2019). His research conducted in the year 1921 is the longest-lasting longitudinal study ever conducted. His work added essential contributions to understanding how intelligence influences life's success, health, and outcomes. According to Cherry, Terman was also a noted eugenicist, once citing Galton as a prime influence. Bobbit and Thorndike were also proponents of eugenics. He administered

English tests to native Spanish speakers and unschooled black students. He concluded that the ensuing low scores were the result of inheritance and had a racial basis. Terman was an unashamed defender of a meritocratic society where economic success, power, and the other benefits of life would be distributed to individuals according to merit. For Terman, “merit” included intelligence, though he never defined the term. Many of Terman’s critics have seen his support for a meritocracy, a ruling or influential class of educated or skilled people, as highly problematic because they see it as elitist, anti-democratic, or misguided (Warne, 2019). Terman’s support for meritocracy becomes problematic because of the lower mean performance of low-income individuals and most non-white groups on the intelligence tests.

### **Usage of Assessments**

Standardized testing is considered essential, and these tests assess what is taught at the national level. There are three initial reasons for Standardized tests: Comparing among test takers, Enhancements of ongoing instruction and learning, and Evaluation of instruction (Popham, 2016). A standardized test is designed to permit a reliable comparison of outcomes across all test takers because everyone is taking the same test. It is quite often that the children tested have not all been exposed to the same materials found on the tests. However, age and/or grade level are poor indicators of what children have learned (Weiss & Garcia, 2017). A 1983 inquiry of arrangement between textbook content and the standardized test found that “in no case was even 50 percent of a test’s content satisfactorily addressed in any textbook” (Popham, 2016, p.31). That is a poor relationship between what was in the test and in the textbooks that were a significant resource to support students’ readiness for the test. Quite often, teachers use test data to drive instruction and learning. These assessments are designed for ranking and are commonly not suitable for helping teachers improve their instruction or modify their approach to individual

students for various reasons, such as they are administered at the school year-end when most classroom instructional activities are near completion and teachers are not in receipt of the outcomes until two or three months later. By this time, their students have usually moved on to other teachers. More importantly, teachers' results usually lack the level of detail needed to target specific improvements (Barton, 2002; Kifer, 2001).

Despite the importance of assessments today, minimal teachers are given much formal professional training in assessment creation or study. A new survey showed, for example, that fewer than half the states require competence in assessment for licensure as a teacher (Stiggins, 1999). Missing specific training, teachers rely heavily on the assessments offered by the publisher of their textbooks or instructional materials.

Assessments are also used to identify academic weaknesses and strengths so that educators can provide specialized academic support, educational programming, or social services. Standardized assessment performance has used statistical techniques called value-added models to estimate teachers' impact on their students. Education Reform (2015) states that in a word, the act of assessing student learning not only takes many forms, but it generally requires an assortment of sophisticated strategies and techniques.

### **Political Assessment Analysis**

Although the above theorist laid the foundations of a positivist approach to Assessment, many state initiatives such as NCLB and a Nation at Risk are. There has been a battle growing over renewing landmark education laws. I positions that the two parties of Democrats and Republicans had agreed about standardized tests and accountability. According to Andrew Rudalevige of Education Next (2003), President Bush and various democratic leaders such as Sen. Edward Kennedy put together the bipartisan coalition behind NCLB in 2002. There have

been many complaints from several states concerning the NCLB that felt the federal government had overruled testing for students, especially ESL (English Second Language) students. In March of 2007, various Democratic senators signed a letter stating that the NCLB law testing is “unsustainable” and wants an overhaul. The NCLB law has drawn opposition from the right because they were opposed to federal interference and from the left because of too much testing. A bipartisan bill was crafted in the wake of 9/11 when there was a premium on lofty ideals and unanimity. These secured out more sober assessments of the federal government’s capacities (Myers & Wallach, 2015). Under the Bush administration, there have been some suggestions for flexibility for states, proposing to give school officials more discretion in using federal money, along with federal power.

There was a bill sponsored by Peter Hoekstra and 50 conservative Republicans in the house that would substantially weaken Washington’s control by allowing states to opt out of law’s testing requirements without losing federal money. Many felt that the law was intended to strengthen, depending on President Bush. Many representatives thought that the government should give states significant flexibility while holding them accountable for ESL students’ progress. Although Representative John B. Larson voted for the law in 2001, he has stated that he regrets this decision (Usher, 2012).

After eight years, new political holders formalized new educational reform ideas. The Obama administration implemented the Race to the Top educational reform. This program offered unprecedented resources \$4 billion to states committed to reshaping their education systems and ensuring every student would graduate college – and career-ready, regardless of disability, race, zip code, or family income (Weiss, 2013). Although this program provided a more significant sum of discretionary funding for education than had been available to states

ever before, Race to the Top was not supposed to be just about money. It was to ensure equality and success for all students.

Unfortunately, that is not the case. Wealthier students get higher SAT scores due to being able to afford to take the test several times, which has been known to increase a students' score. Moreover, wealthy students are still more likely to have taken standardized tests like the SAT more than once. Test scores today say a lot about what our labor force will look like over the coming decades. Our current students' skills will dictate our economic future in the long run. Understanding the implications of higher skills—as measured by regular standardized tests—provides a way of assessing how our country as a whole will fare in the coming years.

Furthermore, the economic costs of not paying attention to the message of stagnating schools are enormous. The absence of improvement in our nation's schools not only translates into significantly lower economic outcomes for our children but also signals a loss of our international prestige and influence (Hanushek, 2018). Why are we risking losing our country's top leadership position in the world economy and the futures of our next generations in one fell gathering?

### **My Assessment Vision**

This dissertation research comes from a dark place in my life where discrimination, hatred, and teacher misguidance were prevalent. After engaging in conversation with teachers, parents, and students, assessments are a heavyweight that no one wants to tackle. I has gained considerable knowledge of how and why it is necessary to give assessments but not quite sure if anyone understands the severity or structure. Yes, there needs to be accountability. Linda (2004) states that students cannot meet the new economic demands if they do not encounter much more

challenging schoolwork. Many argue and schools cannot be stimulated to improve unless their students' actual accomplishments or deceits are raised to public attention.

I support assessments because students need assessments to determine if they are doing well in a class or subject area. Assessments help provide information to help educators, policymakers, and parents. In some retrospect, it can help motivate the learning process, but at what point do we say enough is enough. Multiple assessments are draining, and the purpose gets lost. What I am not in agreement with is the frequency and the number of assessments given. I am not in agreement with multiple assessments given using the same data points. I cannot entirely agree with every eight years when new government new assessments are designed when there are already assessments in place to measure students' growth. I am not in agreement with not allowing educators ample time to view the assessment data, give students good feedback, and provide the next steps for growth. I agree with not holding the classroom assessments to the same high standards as standardized assessments.

### **Concluding Insights**

The Economic Policy reports (August 2010) assert that elementary students take too many standardized assessments. What is more, no evidence adding testing time improves student achievement. What we need is value control in learning to quality assurance. Traditional approaches to teaching and assessment involve teaching some given material and then, at the end of the teaching, working out who has and has not learned the material – akin to a quality control approach in manufacturing. In contrast, assessment for learning involves adjusting instruction as needed.

In contrast, knowledge is an exciting place. It is a quality assurance approach. Valued assurance also involves a shift of attention from teaching to learning. The emphasis should be on

what the students are getting out of the process rather than what the teachers are putting into it. (Ratvitch, 2009), the result is profound demoralization among teachers, who understood from the start that testing is a measure, not a goal, and that test scores largely reflect family income, over which teachers have no control. There is considerable redundancy in the tests. For instance, it is not unusual for school systems to administer multiple summative exams towards the end of the school year that assesses student attainment in the same subject. Students spend a reasonable amount of time taking tests, but its extent depends on the state, the district, the student's grade level, and learning needs and aspirations. It was clear from the research I reviewed for this paper that the time needed on average to take mandatory tests amounts to about 25 hours or so or between four and five days per school year, about 2.34 percent of a typical 180-day school year (Strauss & Corbin, 2015). This is not a huge part of a school system's total instructional time. However, in practice, testing time can be separated over additional days. Additional instructional time may be lost in downtime (e.g., state NCLB exams may be given in sections with one subject taking multiple half-days). There is a long tradition of the United States' desire to measure every school system aspect quantitatively. I often marvel over whether the successive amount of testing is the right way for students to show their academic knowledge in the classroom.

### **Research Questions**

My experience as a student and as an educator and the historical and social, and political background of the assessment trends discussed in this chapter serves as a context to the research questions that guide this dissertation research. These research objectives are to answer the following questions:

**RQ1:** What are all these standardized assessments needed for?

**RQ2:** How are standardized assessment usage perceived by educators?

### **Organization and Structure of Subsequent Chapters**

Chapter Two will provide a brief history of the need for Education Reform within the United States and what Assessment entails, and its criticism. Chapter Three recognizes the phenomenological methodologies of this research design. This chapter will bring forth the participants' selection in the interview process. I provided an Interview Guide to assist the control of the conversation with the interviewees. Data analysis methods for interpreting the research will be explained, and my interim reflections throughout this research.

## CHAPTER TWO

### The Review of the Literature

This chapter will provide an understanding of the base of assessments within the social and educational context classroom, and discuss the assessment reform achievement gap, assessment accountability, what assessment test entails, formative assessments, summative assessments, assessment paradigms, post-positivist paradigms, constructivist paradigms, transformative paradigms, pragmatic paradigm, students perception, teacher perceptions, criticism of assessments, teaching to the test, assessment validity and reliability, and the assessment price tag. Different theoretical perspectives of Assessment and evaluation will be addressed. This research will explore the number of standardized assessments given, the loss of classroom instructional time, assessment validity, and lost data. Chapter Two is designed to inspire interesting thoughts and declare empirical and factual data that “serve as evidence for the arguments and assertions” proposed throughout this study (Efron & Ravid, 2013, p.21).

Assessments are a valuable tool designed to inform what a student needs and inform where there are deficiencies. In the previous chapter, I recalled their purpose and inspiration to bring awareness to this subject. Although assessments are a crucial component of a student’s academic life, there have been many drawbacks. This review attempts to summarize the prevailing literature on standardized assessments and their impact on elementary-aged students and teachers.

The last decade has witnessed growth in recognizing the need for significant changes in educational assessment practices (Archbald & Newman, 1988). This review's efforts are to give a short overview of how educators and students are affected by standardized assessments and the historical data of how assessments were formed, and how they seem to keep being altered year after year. The different points the critics have raised leads to the question: How is data from the

standardized assessments used in the elementary grades? Are all these data points needed? Mainly, are all these assessments necessary?

### **Assessment Justification**

Assessment for obtaining knowledge is best described as a process by which teachers use assessment information to fine-tune their teaching strategies and by students to adjust their learning strategies. The primary purpose of assessments is to improve students' learning and teachers' teaching, as both respond to the information it provides (Guskey, 2003). Assessments are an ongoing process that arises out of the interaction between teaching and learning. What makes Assessment for learning effectively is how well the information is used.

Standardized assessments have a role to play as an audit. However, many policymakers and parents forget or do not know that these tests have a very narrow focus and purpose as audits (Edutopia, 2008). Although teaching is one of our most essential professions, teacher morale is falling. Our front-line educators are burdened by tests and isolated in their classrooms. They get discouraged. Fifty percent of new educators vacate the profession within five years, and the problem is worse in our most troubled schools (Garcia & Weiss, 2019).

Education cannot only be about tests: Great teachers have passion, and passion must be nurtured. Teaching basic literacy and “voluntary compliance” was the moral law of the first common schools. Their mission was to teach the young to work in fast-growing *refectories* (Greene, 2001). We have since evolved, and the importance of assessments has become a way to inform students of their faults rather than pave a way to their success.

#### **Assessment Reform and Achievement Gap**

It seems indisputable that U.S. education requires reform. Ravitch (2010) argues that the standardized test is normed on a bell curve that will not close the achievement gap. It is known that black students naturally mark one standard deviation under white students on standardized

tests-roughly the alteration in performance among the average 4<sup>th</sup> grader and the average 8<sup>th</sup> grader (Fryer & Levitt, 2004). Historically, what has come to be identified as the black-white test score crack has emerged before children enter kindergarten and have tended to widen over time (Fryer & Levitt, 2006). There are many beliefs that assessments do not support efforts to overcome the achievement gap. Education Reform (2013) notes that the unequal or inequitable distribution of educational results and the disparity between what students have learned and what they were expected to learn at a particular age or grade level qualifies for an achievement gap.

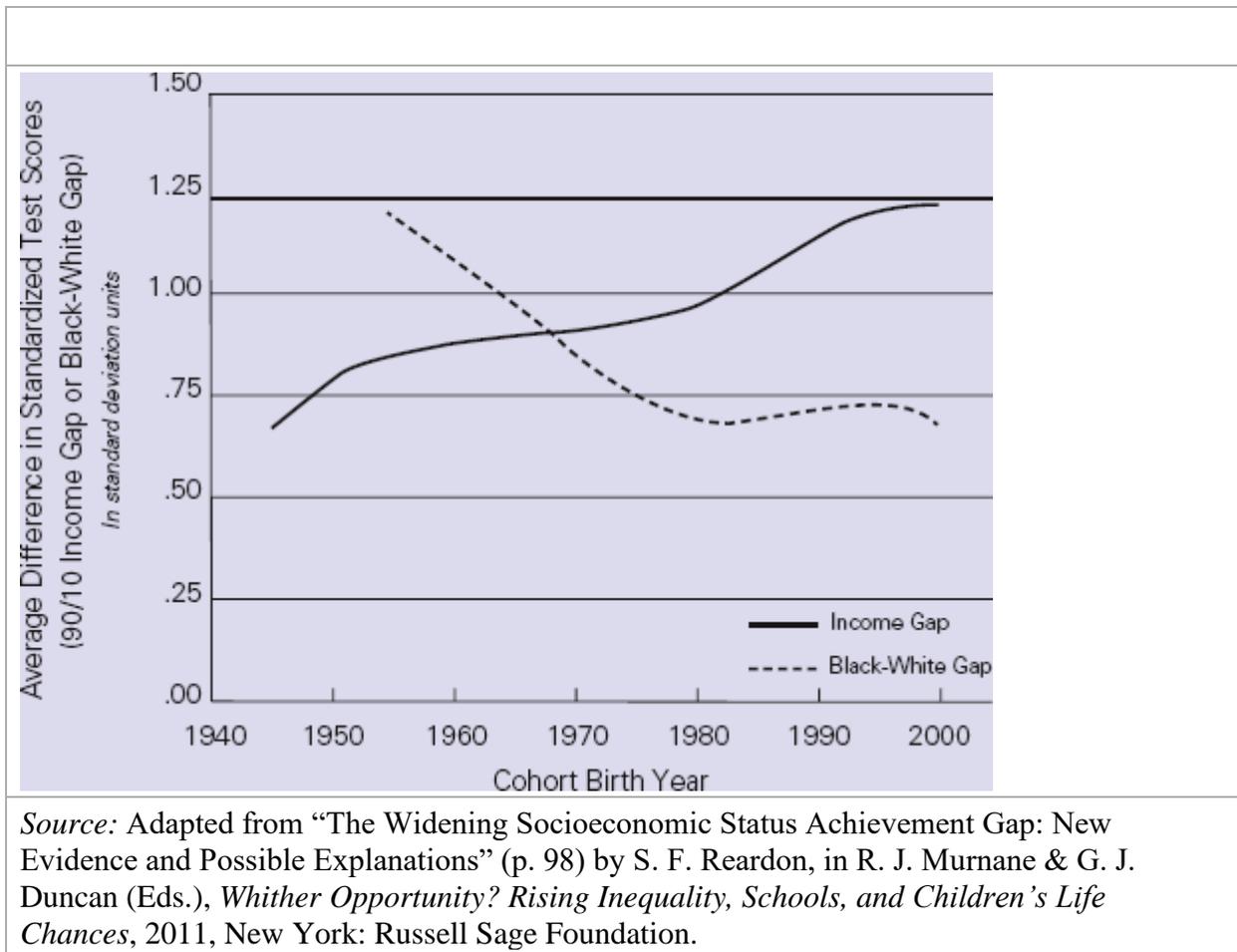
The achievement gaps are broadly defined as the differences in academic performance between groups of students of different backgrounds. They have been documented concerning students' ethnicity, Race, gender, English language learner status, disability, and income status (Aiken & West, 1991). Some assessments have been adjusted to support these gaps. Diane Ravitch (2009) explains that the bell curve reflects the demographics of test-takers. The opportunity to learn is critical, especially for our neediest students. For example, as a group, students labeled as economically disadvantaged or poor never score higher on standardized tests than their non-dis-advantage peers in any state on grade level currently tested under NCLB (Carter & Welner, 2013). Reardon (2011) conducted a complete study of the relationship between academic attainment and family revenue in the United States over the latter 50 years. He used data from 122 nationally representative studies that included family income and student performance on standardized tests in math or reading. Because each test measured reading and math skills on a different scale, Reardon related that he checked all test scores and expressed the income achievement gap in standard deviation units.

Figure 1 encapsulates two critical trends in U.S. history over the last 50 years. In the 1950s, racial inequality was high in virtually every life domain (education, health, earnings,

residential segregation). Simultaneously, monetary inequality was inferior to it had ever been in the last century (Piketty & Saez, 2003). By the start of the 21<sup>st</sup> century, racial inequality was much worse (although far from eliminated) in wages, health disparities, and residential segregation. Meanwhile, economic inequality reached historic highs (Saez, 2012).

**Figure 1**

*Income Achievement Gap and Black-White Achievement Gap in Reading for 1943–2001 Birth Cohorts*



For young people born in the 1950s, 1960s, and early 1970s, the reading achievement gap among those from high-income families (at the 90<sup>th</sup> percentile of the income circulation) and

those from low-income families (at the 10<sup>th</sup> percentile) was about 0.9 of a standard deviation. As illustrated in Figure 2, this gap began to widen, beginning with the cohorts born in the mid-1970s. For those born 20 – 25 years later, the gap in standardized test scores was roughly 1.25 standard deviations--40 percent larger than the gap several decades earlier. According to Rubenstein (2008):

Today's standardized assessments can be useful for spotting significant trends or gauging the effectiveness of state programs overall. However, when used in high-stakes accountability, as the sole indicator of an individual student's achievement or the quality of a single school or school district, these tests can be imprecise. Creating and scoring such tests is complex.

### **Assessment Accountability**

Students are given assessments for various reasons. It has left teachers, administrators, policymakers, and parents wondering if standardized assessments ensure teachers' accountability, promote student achievement, inform instruction for educators, and evaluate instructional methods in schools. These wonderings have sparked a new conversation on assessment reform. Valerie Strauss (2016) noted in the Washington Post that the United States has embarked on reforming its public education system and civic institution. She notes that this has been based on fair principles and the trust that standardized assessments are the best way to assess students, teachers, principals, schools, districts, and states. Standard-based reforms have become widespread across the United States, particularly in the wake of NCLB (Rand Report, 2008).

However, there are also positive assessment efforts. The 2015 innovative assessment in the Every Student Succeeds Act (ESSA) demonstration project, in which up to seven states can build new systems, has had the most significant victory for improving assessments. One of NCLB's biggest complaints was the test and punitive nature of the law – the high-stake

consequences attached to student standardized test scores (O'Brien, 2016). The ESSA, for the first time, states must use more than academic factors in their accountability system. At a minimum, one gauge of school success or student support – such as attendance, school climate, or access to A.P. or other advanced coursework—must be included in measuring school performance. However, academic factors must still make up at least half of all indicators for accountability purposes (O'Brien, 2016). This assessment initiative provides flexibility for states to redesign systems of assessments to better align with student-centered learning. To develop an accountability structure that explains not just what results were reached but what decisions led to those outcomes, states should consider measuring the effectiveness of coordination among and between each level of systems: states, districts, and schools (Jimenez & Scott, 2017).

In designing healthy interaction systems to close the achievement gaps between states, districts, and schools, the critical questions states must ask are: What are the reasonable, short-term outcomes that states, districts, and schools can expect? How are these measured and by whom, and how often are the results reviewed? State's answers to these questions should inform their development of the metrics, benchmarks, and processes foundation to their accountability systems (Jimenez & Scott, 2017). Knowledge and skills shape learners' future lives, so learning assessment must be accurate, reliable, and fair. Assessments have proved to be problematic and controversial because of their multivalent functions.

### **What Assessment Tests Entail**

As a result of inadequate accountability, teachers have been teaching content then test students. "Teaching to the test" is an informal term for any education method whose curriculum is deeply absorbed in preparing students for a standardized test. This sequence of teaching and testing is known to anyone who has been a student. Tests in theory measure what students have learned.

However, there can be other more difficult reasons as to why schools use tests. At the school level, educators create a test to measure their students' understanding of specific content or the practical application of critical thinking skills.

We can differentiate assessment tools between formative, summative, -standardized commercial achievement tests, and teacher-made tools. While assessments are often associated with traditional tests—especially the standardized tests were designed by testing companies and given to large populations of students—educators use an assorted array of assessment tools and methods to measure everything from a four-year-old's readiness for kindergarten to a twelfth-grade student's comprehension of advanced physics. These tests are constructed by professional item writers and undergo extensive pretesting before administering them to students across the district, state, or nation (Efron & Ravid, 2019). Efron and Ravid noted that the majority of these tests could have been divided into two types: norm-referenced and criterion-referenced.

According to the Education Reform (2015), norm-referenced tests, a form of a standardized assessment, is designed to compare and rank test takers to one another. Norm-reference tests report measures whether the test takers performed better or worse than the hypothetical average student. This resulted by comparing scores against the performance results of a statistically selected group of test-takers, typically of the same age or grade level, who have already taken the exam. The Glossary of Education Reform (2015) also states that Criterion-referenced tests/assessments are created to quantify students' performance in contradiction of a fixed set of predetermined criteria or learning standards. These are concise, written descriptions of what students are expected to know and do at a specific stage of their education. In the elementary setting, the criterion-referenced test evaluates whether students have learned a

specific body of knowledge or acquired a particular skill set. In some districts, criterion-referenced tests were created to make significant decisions for student's advancement.

There are various forms of assessments within an elementary classroom. At the Illinois state level, schools are required to administer five standardized assessments. There are currently six standardized assessments at the district level that I teach in that are being administered multiple times throughout each school year, including:

1. Northwest Evaluation Association (NWEA)
2. Dynamic Indicators of Basic Early Literacy Skills (DIBELS)
3. REACH Performance Task
4. Benchmark Assessment System (BAS)
5. Text Reading and Comprehension (TRC)
6. Skilled Assessed (mClass math)

Grades K-2 assessments are REACH, KIDS, NWEA (ELA), NWEA (Math), Dibels, and TRC. Recognizing Educators Advancing Chicago students (REACH) is a new system. It evaluates teachers' student growth and the quality of teaching practice. This assessment is administered at the beginning and end of the school year.

The Kindergarten Individual Development Survey (KIDS) was developed in 2015 but was piloted for five years before being an assessment mandate. It is an observational tool designed to help teachers, administrators, families, and policymakers better understand children's developmental readiness to enter kindergarten. This assessment is given once during the school year. Therefore, according to the Illinois State board of education (ISBE), KIDS is a core assessment that states that every child in Illinois deserves to attend a school wherein all

kindergarteners are assessed for readiness (KIDS: Every Illinois Child Ready for Kindergarten, 2019).

In 1973, educators and researchers from the Pacific Northwest discussed how to create a more efficient method of measuring student achievement in schools. This achievement should support students and educators worldwide by creating assessment solutions that precisely measures growth and proficiency and provides insights to help tailor instruction (nwea.org). Northwest Evaluation Association (NWEA) is a non-profit organization that has evaluated over 4.5 million students. NWEA has an occurrence in 49 foreign countries, 50 states, and 3400 districts. This assessment is the only district-mandated exam within all schools. Students are liable to take this assessment at the beginning, middle, and ending of the school year (NWEA: Advancing growth for all students, 2019).

The text Reading and Comprehension (TRC) program is a technology-based reading assessment that combines performance on brief indicators of foundational skills with performance on a running records measure. TRC allows teachers to record detailed running records of student reading and assess student comprehension using a series of leveled books. TRC also allows teachers to evaluate initial written perceptions and reading habits in pre-readers. Teachers and school leaders use TRC results to make instructional decisions and carry out/monitor interventions and support (Learning, 2019).

According to the Institute for Research and Learning Disabilities at the University of Minnesota (1989), in the late 1980s, the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) was designed as a customer of actions and measures for assessing the acquisition of early literacy skills from kindergarten through sixth grade. Dibels is the seventh assessment used in Illinois; it is a short one-minute fluency measure used to display early literacy and early

reading skills frequently. DIBELS consists of seven measures to function as phonemic awareness indicators: alphabetic principle, accuracy, and fluency with connected text, reading comprehension, and vocabulary. DIBELS was created away to identify children experiencing difficulty in the acquisition of basic early literacy skills to provide support early and prevent the occurrence of later reading difficulties (Young, 2006). These assessments are administered three times a school year.

In their district, intermediate grades 3-5 state assessments are REACH, NWEA (Math and Reading), Illinois Science Assessment (ISA), NAEP, and PARCC ("Illinois State Board of Education," 2019). Illinois Science Assessment complies with federal testing obligations. Illinois will manage a science assessment for students enrolled in a public school district in grades 5, 8, and once at the high school level. The high school assessment utilizes a course-based model with content aligned to Biology I ("Illinois Science Assessment (ISA)," 2019). The assessment will be given in an online format aligned with the Illinois Learning Standards for Science, including the Next Generation Science Standards (NGSS) in 2014. The National Assessment of Educational Progress (NAEP) is the most extensive ongoing and nationally symbolic assessment of what U.S. students have obtained and can do in various subjects. NAEP is a congressionally mandated project given by the National Center for Education Statistics (NCES) inside the Institute of Education Sciences (IES) of the U.S. Department of Education. The first national management of NAEP happened in 1969. The National Assessment Governing Board (NAGB) is a self-governing, bipartisan panel that sets rules for the NAEP and is accountable for creating the framework and test specifications ("Illinois State Board of Education," 2019). The National Assessment Governing Board, whose supporters are chosen via the U.S. Secretary of Education, contains governors, state legislators, local and state school officials, educators, business

representatives, and the general public. Congress formed the 26-member Governing Board in 1988 (NAEP, 2013). Teachers who teach grades 6-8 only administer the REACH, NAEP, and PARCC assessments. The Partnership for Assessment of Readiness for College and Careers (PARCC) is a group of states working together to develop assessments that measure students are on track to be successful in college and careers (Pearson, 1998-2015).

***Formative Assessment.*** Although assessments are indeed used to identify academic weakness and strengths, assessments are also typically designed to measure specific elements of learning—e.g., a student’s perceived ability or readiness to learn; the successful acquisition of particular skills and knowledge; the understanding and recall of facts; or the ability to analyze and comprehend various types of texts and readings (Stiggins, 2002). Using a range of assessments allows the teacher to gather more evidence along the way. In addition to the state and district-level mandated standardized assessments, educators administer formative, summative, and diagnostic assessments. These assessments allow educators to formulate report card grades.

Formative assessment, including diagnostic testing, is a range of formal and informal assessment procedures conducted by teachers during the learning process to modify teaching and learning activities to improve student attainment ResourceEd, (2017). Roskos and Neuman's (2012)'s article asserts that formative assessments are often taken for granted, and educators are not using them to their proper development of identifying misconceptions, struggles, and learning gaps along the way and assessing how to close the gaps. It includes practical tools for helping shape learning. It can even boost students’ capabilities to take ownership of their knowledge when they comprehend that the goal is to advance learning, not apply final marks (Trumbull & Lash, 2013). Formal and informal assessments provide significant insight into students’ progress, curriculum effectiveness, and teaching strategies.

Formative assessments are not a judgment evaluation but rather are conducted in the flow of the instructional process.

The formative assessment's central purpose is to identify the gap between where students are and where they need to go, such as reading development. When gaps are significant, teachers face problems of having insufficient instructional time, support, and proper resources to address learning needs, not to mention maintaining students' motivation to learn. Formative assessment is a gap-minder because it helps the teacher stay alert to gaps in individual students' reading development and adjust instruction as needed before moving on. It is essential not to create a single snap at some predetermined benchmark. The district I teach in uses formative assessments for assessments which is for the wrong reasons.

Formative assessment must turn out to be a strength in educational discourses and practices. Black and Williams (2009) inform us that the first reference to the term "formative" has roots in curriculum development and evaluation. They also recognize that Cronbach (1963) refers to the idea of using assessment as a tool for refining curricular programs. Scriven (1967) builds on Cronbach's work by recommending the term "formative" to clarify evaluation roles. Bloom (1971) applies Scriven's definition to the procedure of learning by using the term to describe a way of improving student learning. Bloom et al. (1971) link the idea of formative evaluation to the instructional approach of mastery learning as an instructional process that includes data to improve both teaching and learning. During the 1980s and 1990s, educational researchers continued to expand on the ideas and theories proposed, and the term "formative evaluation" was replaced by the term "formative assessment." Sadler (1989) proposed a theory of formative assessment that builds on the definition previously offered. However, it highlights the student's role in the assessment procedure and interpretations student self-assessment as

critical to improved student learning. Gipps (2012) documents how the educational community views assessment, including a shift from a psychometric view to assessing assessments and assessment data by educators to monitor instruction. In the 1990s and the initially 21<sup>st</sup> century, the Assessment Reform Group in the United Kingdom focused on developing formative assessment practices and providing a definition of formative assessment (Gipps, 2012).

Assessment improvement group members Wynne Harlen and Mary James believed that a difference between formative and summative assessment was needed to meet these two roles in the field (Garies & Grant, 2017).

Educators and students should use formative assessments during instruction that provides actionable feedback. The actionable feedback is used to adjust ongoing teaching and learning strategies to improve curriculum learning achievement. This is not the case. Educators are advised to give formative assessments attached to a grade. Formative assessments are renamed quizzes. There are no formative assessments given. The conferring component is not happening as much as it should with students on various curriculum tasks. Summative assessments are often administered at the end of a unit. These assessments are excellent support for students but often come too late in the learning process. When preparing to teach students, there should be courses on assessment implementation along with assessment design. These components would support teachers in developing a holistic view of assessments.

***Summative Assessment.*** On the other hand, summative assessments focus on making judgments at some point in time after instruction. This assessment aims to assess student learning at the end of an instructional unit by comparing it to a few standards or benchmarks. Summative assessments are frequently high stakes, which means that they have a high point value. (edglossary, 2013).

Tests, assignments, or projects are designed to determine whether students have acquired academically what they were anticipated to learn. In other words, what makes a classroom assessment “summative” is not the development of the assessment, assignment, or self-assessment, per se, but the way it is used—i.e., to regulate whether and to what degree students have learned the classroom curriculum they have been taught (Education Reform, 2013). Summative assessments are usually compared with formative assessments, which collect detailed information that educators can use to inform instruction and student learning while it is occurring. In other words, formative assessments are frequently said to be for learning, while summative assessments are of learning, or as assessment expert Robert Stake (2010, p. 169) put it, “when the cook tastes the soup, that is formative assessment. When the customer tastes the soup, that is summative assessment.” While there is minimal confrontation among educators about the need for or utility of summative assessments, debates and disagreements tend to center on the fairness and effectiveness of this type of assessment, mainly when summative assessment results are used for standardized assessment purposes. In these cases, educators, experts, reformers, policymakers, and others may debate whether assessments are being designed and used appropriately or whether standardized assessments are either beneficial or harmful to the educational process. Ultimately, teachers need to use their professional judgment when assessing students and determining scores. Teachers can decide that a summative assessment should be formative and then reteach and support students' learning before attempting another summative assessment. (Miller, 2020).

### **Assessment Paradigms**

Classroom assessment has been a crucial and integral part of any curriculum. Assessment plays the part of reporting students' obtaining knowledge (assessment of learning), providing diagnostic and formative information for teachers to inform their instruction (assessment for learning); more recently, Earl (2013) proposed the notion of assessment as learning, which puts students at the center of assessment. Students in this assessment paradigm act as critical connectors between assessment and learning through self-reflection and self-regulation. At this age of responsibility, it is acknowledged that assessment is a powerful lever that can either boost or undermine students' learning. Hitherto, much of the regular institutional practices show that assessments remain inhibitory or void rather than constructive as these lack the formative assessment aspect (Ghaicha, 2016). This denotes that assessment is either not well understood or not done in a principled educational framework across all educational levels.

Various assessments involve different beliefs about the nature of knowledge, the level of teacher and student involvement, the criteria for evaluating student achievement, and the effects of these assessment frameworks on classroom instruction (National Research Council, 2019). Student assessment enables instructors to measure their teaching effectiveness by linking students' performance to specific learning objectives. As a result, teachers can institutionalize effective teaching choices and revise ineffective ones in their pedagogy. Various assessment frameworks use similar procedures and data collection methodologies, and many of these "alternative" assessments do not adhere to traditional criteria of standardization, reliability, and objectivity (Kibble, 2017).

Merten and Wilson (2019) note that they have recognized four paradigms that provide a valuable structure for examining different world views that are functioning in today's evaluation world: postpositivist, constructivist, transformative, and pragmatic paradigms, along with four

sets of philosophical assumptions in paradigms. These paradigms are constructed by the following: Axiology which is the study of value (ex: your societies, cultures, and religions that teach one moral of what is right and wrong); Ontology which is the study of being (ex: Is there one reality that you can discover or multiples depending on experiences); Epistemology, meaning knowledge (example, how does the evaluator relate to the stakeholder); and Methodology which would be applying a systematic method of gaining and measuring assessment data, (ex: do you need two groups or can you document progress by intensively studying one group).

### **Post Positivist Paradigms**

This paradigm focuses primarily on quantitative designs and data; it may use mixed methods, but quantitative methods dominate. Positivists hold the ontological belief that one reality exists and independent of the observer (Sale Lohfeld & Brazil, 2009).

Post-positivism is a significant paradigm that guides many evaluators in their work. The Logical assumption of this paradigm is intertwined with the methodological assumption, in that the conduct of “good research” is a fundamental requirement for ethical good. Good research is described as that which reflects “intellectual honesty, the suppression of personal bias, and careful collection of empirical studies” (Jennings & Callahan, cited in Christians, 2005, p.159).

Although postpositivist research scientifically strives to explore the phenomena, it believes, unlike positivist research, that the absolute truth is nowhere to be found (e.g., Wildemuth, 1993; Guba and Lincoln, 1994; Phillips and Burbules, 2000). Post-positivism does not aim to disapprove of the scientific/quantitative elements of positivism in the research. Instead, it emphasizes a proper understanding of any research study's directions and perspectives from multi-dimensions and multi-methods (Guba, 1990; Fischer, 1998). The post-positivist

paradigm promotes the triangulation of qualitative and quantitative methods that explores the diversity of facts researchable through various kinds of investigations but respecting and valuing all through various kinds of investigations but respecting and valuing all findings as to the essential components for the development of knowledge (Watson, Clark, & Tellegen, 1998).

### **Pragmatic Paradigms**

The pragmatic paradigm focuses primarily on data deemed to be useful for stakeholders. It advocates for using mixed methods and focusing on the assessment procedures, not on the underlying purposes of the assessment program or the epistemological stance (Serafini, 2000). Serafini (2008) informs that teachers in this paradigm are still being asked to objectively measure students' abilities and report information in numerical form to external audiences. In effect, the procedure or method of collecting data in and of itself does not determine the assessment paradigm.

Researchers have adopted the pragmatic paradigm because it reflects the assumptions that underlined their work (e.g., Morgan, 2007; Teddlie & Tashakkori, 2009). Morgan (2007) borrows from Dewey, James, and Mead to explain what researchers/evaluators do in pragmatic terms. Early pragmatics emphasized the ethics of caring as their axiological assumption (Merten & Wilson, 2019). However, contemporary pragmatists' ethical assumption is more closely aligned with the utilitarian theory of ethics, which holds that something's value is a function of its consequences (Christians, 2005). Pragmatists see the value of the evaluation in terms of how it is used and its results. Tashakkori and Teddlie (2003) assert that pragmatists avoid spending a great deal of time arguing about metaphysical terms such as "truth" and "reality." Pragmatics do not proclaim that they will discover the truth.

Some researchers identify the pragmatic paradigm as the philosophical framework that guides their choice of mixed methods, which is the underlying methodological assumption of pragmatism. The method should match the purpose of the study (Patton, 2002). The evaluator chooses a method based on what is suitable for a particular study in a particular context with a specific stakeholder group.

### **Constructivist Assessment Paradigm**

Constructivist paradigms focus primarily on identifying multiple values and perspectives through qualitative methods. They may use mixed methods, but qualitative methods dominate. This paradigm assessment is based on constructivist theories of knowledge (Fosnot, 1996), student-centered learning. This is when portfolios are seen as a vehicle for promoting student and teacher reflection, self-evaluation, and goal setting. Assessment in a constructivist paradigm can inform instruction, guide the following steps, and help students monitor progress toward achieving their learning goals. When creating assessment opportunities using an inquiry-based learning approach, educators should consider the following concepts (Gregory, Cameron, and Davies, 2011):

1. Planning assessments should be planned simultaneously as instruction, and they should be joined effortlessly into the learning cycle. Scheduled assessment should be used to inform instruction, monitor, following steps, and help both educators, and students
2. Establish criteria for assessment and evaluation should be shared with students or co-constructed with students before learning.
3. Secure that assessments are ongoing throughout the learning cycle, varied in nature, and administered over time.

When educators use various sources of evidence, the reliability and validity of the evaluation of student learning are improved. To guarantee valid and reliable assessment and evaluation, educators are invigorated to collect evidence of student learning from various sources on an ongoing basis and various settings. Sources can include conversations, observations, and products collectively referred to as “triangulation of evidence.”

Constructivists have called attention to the importance of researchers’ awareness of their values and their reflections on how their values influence the research process and outcomes (Ponterotto, 2005). Constructivists hold that there are multiple socially constructed realities (Guba & Lincoln, 2005). This assumption reflects a relativist view of reality, in which individuals construct reality by reflecting upon their experiences and interaction with others. In a research context, reality is seen as a hermeneutic process constructed by the research participants through interactive dialogue with me (Ponterotto, 2005). The epistemological assumption for constructivists is already reflected in the ontological assumption (Mertens & Wilson, 2019). The constructivist epistemological assumption is that researchers and participants interact through meaningful dialogue and reflection to create knowledge (Guba & Lincoln, 2005). A researcher needs to be immersed in a community’s everyday activities for a long time to have sufficient opportunities to engage in reflective dialogue with participants. Lincoln (2010) reminds us that constructivists are not limited to qualitative data collection.

### **Transformative Paradigms**

Transformative paradigms focus primarily on identifying multiple values and perspectives through qualitative methods; they may use mixed methods, but qualitative methods dominate. This paradigm pulls together many strands of philosophy that focus on issues of power

and on addressing inequities in the name of furthering human rights and social justice (Mertens & Wilson, 2019).

A critical social theory is concerned in particular with issues of power and justice and the ways that the economy; matters of race, class, and gender; ideologies; discourses; education; religion and other social institutions; and cultural dynamics interact to construct a social system.

While the transformative paradigm shares much in common with the constructivist, pragmatic, participative, critical paradigms, particularly in terms of its strong stance away from the post-positivist approach, what ultimately differentiates it is the focus on the dimensions of culture power, privilege, and social justice. The significant overlap between paradigmatic positions illustrates what Guba and Lincoln (2005) refer to as postures' fluidity and the shifting of categories. Researchers and evaluators confront a single paradigm's methodological limitations in capturing the multiple voices that demand our attention.

## **Perceptions of Assessment**

### **Students' Perception of Assessment**

Students and teachers must comprehend the assessment criteria since they are most affected by them. Students need to be unblemished about what they aim to learn, which signs and standards are appropriate to gauge progress and inform future learning. Evaluating student achievement through assessments is not a new strategy. Good teachers have always judged and monitored their students' progress through observations, experiments, written assignments, and research projects. There seem to be two theoretically related types of students' insights about the standardized assessment atmosphere: individual and aggregate.

The unique insight of the classroom assessment environment refers to an individual student's self-perception of the various aspects of classroom assessment activities. The educator creates assessment purposes, assigns assessment tasks, sets performance criteria and standards, and provides feedback to monitor outcomes (Brookhart,1997).

The aggregate perception of standardized assessments atmosphere refers to the generally shared insight of students in a classroom about the many aspects of the classroom assessment activities done by the educator; and it can be reviewed for each class by the average levels of individual students' perceptions within the classroom (i.e., the class average perceived assessment environment). (Alkharusi, 2008). One may dispute that the aggregate perception of the assessment environment is a compelling attribute for describing the classroom's social influence. From the perspective of social theory, standards are created to give class members some power to regulate their actions over others when they affect the class (Coleman, 1987).

As I stated previously, assessments always brought on sweats, nervousness, and scrambled thoughts. When I became an educator, myself wanted to alleviate any of these feelings of anxiety from their students. While I taught class routines and expectations during the first few weeks of school, myself also taught students how to perceive assessments. I hoped to inform students that the tests will assess various reading and math strategies that are taught in class, but there will be times that the state mandates assessments. Therefore, this researcher and other teachers have little to no control over the content, and that there is no need to feel nervous about the content. That is why every time students take a test, I encourage students to clap. The clapping is a sign that the students will do their best no matter who is giving the test. Many of my students have expressed how clapping brings on a good feeling about taking a test.

Another approach I has tried is to not test students on Fridays. For some reason, students feel that tests occur on Fridays. I s test on Thursdays, allowing them to grade their work and reteach and retest on Friday. This has set the classroom up to be successful and allows students to feel good about what they are trying to accomplish. My students are frustrated when it comes to standardized assessments. When they are not sure about a particular strategy, the culture asks a question, but they cannot receive any support from the proctor or I during these tests. I often sees frustration, confusion, and nervousness on their faces. I once asked one of their students how they felt about taking assessments, and they replied, “I only like taking the classroom assessments because I know what is going to be on the test.” He went on to say that “I do not like those longer computer tests (standardized assessments) because they ask me questions and to solve problems that I never heard of.” The students state that they began to shut down and start a guessing game of picking the correct answer. I then assured him that we would work on strategies to break down questions to locate and identify the correct solution to a problem.

### **Teacher Perception of Assessments**

New Jersey Governor Chris Christie has suggested educational reform in New Jersey that persuaded teachers to “teach to the test” and have their students act at the possible cost of their income and job safety. The reform calls for performance-based pay that depends on students’ performance-based pay that depends on students’ performances on standardized assessments and their informative gains. However, students vary founded on cognitive, developmental, and psychological abilities, so it is unfair to teachers with students with difficulties on the test (Arco, 2015). According to an NEA (2014) survey, most teachers reported feeling considerable pressure to improve test scores (Walker, 2014). Legislation like NCLB has high stakes for testing, potentially combining student performance to teacher salaries and job stability and dictating what the teacher teaches (Nixon, 1989). This pressure, stress, or burnout is frequently wielded by

administrators who draw attention to the test by looking at scores with the entire staff and applying praise or reprimands based solely on test scores. This pressure often results in drill-and-practice-type instruction.

Burnout, which is a broad term, represents a response to the chronic emotional strain of dealing extensively with others in need (Maslach, 1982), date there is still no university accepted definition of burnout (Dworkin, 1987; Farber, 1991; Handy, 1988; Jackson, Schwab, & Schuler, 1986). Emotional Exhaustion, Depersonalization, and condensed Personal Accomplishment are three elements of burnout that have been empirically validated at the elementary, intermediate; and secondary levels (e.g., Beck & Gargiulo, 1983; Friesen, Prokop, & Sarros, 1988; Friesen & Sarros, 1989; Gold, 1984; Iwanicki & Schwab, 1981; Jackson et al., 1986; Schwab & Iwanicki, 1982a, 1982b). Teachers exhibit signs of emotional exhaustion when they perceive themselves as unable to give of themselves to students, as they did earlier in their careers; of depersonalization when they develop undesirable, cynical, and occasionally callous attitudes toward students, parents, and colleagues; and feelings of diminished personal accomplishment when they perceive themselves as ineffective in helping students to learn, and in fulfilling other school responsibilities.

Overall, teachers who fall victim to burnout are likely to be less sympathetic toward students, have a lower tolerance for classroom disruption, be less apt to prepare adequately for class, and feel less committed and dedicated to their work (Farber & Miller, 1981). Assessments affect decisions about grades, placement, advancement, instructional needs, curriculum, and in some cases, funding. Edutopia (2008) wrote that assessments stimulate us to ask these tough questions: “Are we as educators teaching what we ponder we are teaching?” “Are students

grasping what they are supposed to be learning?” “Is there a way to teach the subject better, thereby promoting better learning?”

### **Criticism of Assessment Test**

Assessment data has been used in many negative ways throughout the years, such as teacher evaluations and student promotion policies. This situation can be very disheartening when taking a test that does not connect with the common core state standards (CCSS) nor classroom instruction. When there are negative consequences to taking the test, it exhibits damaging behaviors such as teachers teaching to the test and misjudgment of assessment validity or reliability.

### **Teaching to the Test**

Howard Nelson from the American Federation of Teachers (2013) states that “teaching to the test” disfavors higher-order learning. While it is likely to use a standardized test deprived of letting its substances determine classroom curriculum and instruction, often, what is not tested is not taught, and how the subject is tested often becomes modeled to teach the subject. Ratvitch (2009) notes much cheating is going on due to putting so much emphasis on the test scores; there is much gaming in the system. Instead of raising standards, it is lowered standards because many states have 'dumbed down' their tests or changed their tests' scoring to say that more kids are passing than are.

An unbelievable magnitude of pressure is placed on teachers to ensure an increase in achievement on standardized assessments. The curriculum's scope and sequence are sometimes modified to cover specific test-taking skills and topics (Burns & Drake, 2004). When standardized assessments are the primary factor in accountability, schools use the test to narrowly define curriculum and focus instruction. Accountability can create immense stress to

perform, leading to the misuse and misinterpretation of standardized assessments (Halloway, 2001). This pressure, anxiety, or burnout is usually wielded by administrators who draw attention to the test by looking at scores with the entire staff and applying praise or reprimands based solely on test scores. This pressure often results in drill-and-practice-type instruction.

### **Validity and Reliability**

Having visibility of validity and reliability is very important in revealing the outcomes and findings of this study. According to Blackburn (2017), “Schools throughout the country are starting to embrace a culture of data, which is the integration of data into the day-to-day operations of a school in order to achieve classroom, school, and district-wide goals.” One of the most severe difficulties with this interrogation is determining what data will accurately reflect those goals.

Most research on the assessment environment has used individual student scores as the unit of analysis rather than the average score of students at the classroom level (Church, Elliot, & Gable, 2001). In her book, *Now You See It*, Cathy Davidson (2011) criticizes standardized assessments. She explains our youth as “assembly line kids on an assembly line model” (p. 73), meaning that standardized testing is used as a portion of an informative one-size-fits-all model. She also studies the narrowness of skills being tested and labels children without these skills as failures or as students with disabilities. Education theorist Bill Ayers (1993) has commented on the limitations of standardized assessments. He writes that standardized assessment cannot measure initiative, creativity, imagination, conceptual thinking, curiosity, effort, irony, judgment, commitment, nuance, goodwill, ethical reflection, or a host of other valuable dispositions and attributes. High-stakes standardized tests can measure, and the count is isolated skills, specific facts, and function, content knowledge, the least interesting significant aspects of learning

(Ayers, 1993). Standardized assessments do not allow for creativity, dreams to become a reality, or support an answer through different one-size-fits-all answers. These tests place students in a box of unoriginality. Students are filling in a bubble, filling in a blank, or circling an answer. This process shows no personality. Morgan (2013) claims that learning based on memorization and recall may improve student performance on tests but fails to develop higher-level thinking skills.

Furthermore, teaching to the test often prioritizes linguistic and mathematical intelligence at the expense of a well-rounded education that fosters creative, research, and public speaking skills (West, 2018). Today, more than a decade later, the No Child Left Behind law is uniformly blamed for stripping curriculum opportunities. These opportunities include art, music, physical education, and more, and imposing a brutal testing regime has pushed educators to focus their time and energy on preparing for tests in a narrow range of subjects: English/Language Arts and math (Walker, 2014).

### **Assessment's Price Tag**

Another criticism of the extreme obsession with assessments in this nation is the price tag that it carries. It has been reported that standardized evaluations are costly. The United States spends roughly 1.7 billion dollars annually on assessments. In contrast, in 2001, it was reported that only three companies (Harcourt Educational Measurement, CTB McGraw-Hill, and Riverside Publishing) designed 96% of the tests taken at the state level (PBS, 2001). It is difficult to figure out why more mandated testing will lead to a higher quality of instruction, especially as current accountability spawned educated tests have little utility in the classroom (Popham, 2008). Barbara Miner (2000), a Milwaukee-based journalist specializing in education, explicates the drawbacks of standardized testing in her article written in *The Progressive* (2000). Miner

analyzes Mc Neil's (2000) book *Contradictions of School Reform: Educational Costs of Standardized Testing* that discusses the assessment's positive and negative aspects. Barbara Miner is the co-director of the Center for Education at Rice University and a professor of education. She *writes*, "Educational standardization harms teaching and learning and, over the long term, re-stratifies education by race and class" (2000). Miner (2000) states that test-based education reforms place higher standards for students of color (Minarechova, 2012).

The discriminatory properties of standardization are immediately evident in reducing the quality and quantity of the educational classroom curriculum for students who have been noted to score low on standardized assessments. The NCLB and other standardized assessment tests measure objectives and how schools are meeting educational state standards. For students in low-income areas, the impact has been upsetting. The testing process disrupts instruction for the remainder of the testing days. Also, it forces a reallocation of resources for all students, regardless of whether they are in a testing grade (Corbin & Strauss, 2015). Apple (2004) notes that other countries, who have already implemented NCLB-like policies, show that, although the forces of conservative modernization might have excellent intentions, we make a big mistake when we assume that "in our unequal society there is a direct relationship between policy intentions and policy results." Apple emphatically notes, "There is not" (p. 202). Under NCLB and RTTT, standardized testing became a multibillion-dollar industry. Its lobbyists were stationed in crucial state capitals and congressional halls to ensure it remained (Ravitch, 2009). Ravitch informs us that now that she has been looking at assessments through a different lens, she realizes that "high standards" and "rigorous testing" do not promote equity. Instead, they produce high rates of failure for many students and widen the gap between individuals at the lowest and persons at the highest. States and local school districts that obtain RTTT funding will

receive additional federal funding. The NCLB requires states to make reforms to continue receiving the federal funds they are already getting. Nevertheless, the two programs address many of the same issues. NCLB mandated a target of 100 percent proficiency which was out of reach in every state. RTT demand that teachers suffer punishment if they could not produce higher test scores was even more demoralizing than NCLB, as it forced colleagues into a game of survival, competing against one another (Ratvich, 2010).

### **Concluding Insights**

One of the outstanding features of assessment studies in recent years has been the shift in attention towards more significant interest in the interactions between assessment and classroom learning. This focus moved away from concentration on the properties of restricted forms of the test, which is only weakly linked to students' learning experiences. Effective reporting is essential to communicate summary statements of achievement to students, their parents, and other school practitioners. Such records can support cooperation with parents, ensure consistency of support after student transitions to higher education levels, and provide a basis to make decisions about a student's future educational career. However, where there is a lack of transparency and consistency in how marks and report cards are constructed, the effect of such reporting will be counterproductive for student motivation and future learning.

Clear central reporting guidelines can help build a shared understanding of marks' meaning and the criteria used to establish student performance. They can also clarify that student behavior, effort, and motivation should not be mixed into performance marks. Therefore, I need to take a more in-depth analysis of the standardized assessments elementary students receive and the loss of their instructional time. This information will provide administrators, teachers, parents, and students insight into how receiving an academic snapshot of one day; one test has

too much weight placed on their results. It is not developing a holistic view of students' capabilities, nor is it providing teachers with data to improve their classroom curriculum.

It is necessary to engage individual students in ways that they learn best, which may mean taking a couple of weeks longer in one subject while moving quickly in areas they have demonstrated proficiency. Although this change will not happen without input from parents, students, and possibly a school board, it is a conversation that needs to be addressed. It is important to remember that we cannot assess our way to academic excellence.

### **Research Divulgence**

This journey started when I decided to continue their educational journey by critiquing their craft by sharpening their skills in writing, obtaining ESL endorsement, becoming a reading and language guru, and entering into the Curriculum, Advocacy, and Policy doctoral program. While doing all of that, I continue to maintain a strong presence within education by teaching future adults every day. Teaching future adults has allowed them to view firsthand the damaging effects standardized assessments bring into the classroom. This understanding gave them the drive to engage in in-depth research of standardized assessments; I needed to increase critical thinking's altitude towards what these assessments are being used for. The following chapter discusses the methodology and the process of how I will carry out their research through a critical research lens.

Critical theory teaches that knowledge is power. This perception means that understanding the ways one is oppressed enables one to change oppressive forces. Critical theory orientation is towards reviewing and altering society. It changes from traditional theory, which focuses only on understanding or explaining society. Critical theoreticians aim to dig beneath the surface of social life and uncover the assumptions that keep human beings from a complete and

accurate understanding of how the world works. Collecting this research will be from developing a relationship between collecting data from observations established based on mathematical calculation. These theories related to a naturally existing phenomenon can be proved or disproved using calculations.

## **CHAPTER 3**

### **Methodology and Research Design**

This study examines the use of standardized assessments within the elementary setting developed by me to understand if these standardized assessments are necessary and their usage perceived by educators. The intent is to understand how standardized assessments within the elementary classroom can overlap and extract valuable instructional time, reducing students' opportunities to learn new strategies and skills. Essential to this research is applying critical theory and how it influences or interacts with the experiences and interpretation of positionality when examining standardized assessments' thoughts and decisions. I am interested in developing a design for assessing elementary students and teachers unpacking the results.

Although I collected some quantitative data, in its essence, this is qualitative research. The qualitative research approach is exploratory and seeks to explain 'how' and 'why' a particular assessment phenomenon operates as it does in a specific context (McLeod, 2019). Due to the development that has become known as the pluralization of life worlds, qualitative research is relevant to studying social relations (Flick, 2014). Quick social alterations and the resulting diversification of lifeworlds increasingly confront social researchers with new social contexts and perspectives. This research will bring awareness to educators, administrators, and parents with a new perspective on standardized assessments which their student/child is being

administered. Qualitative data is mostly non-numerical, covering images, videos, text, and people's written or spoken words. As such, qualitative research often investigates i) local knowledge and understanding of a given issue; ii) people's experiences, meanings, and relationships; and iii) a social process and contextual.

According to Cresswell (2003), a phenomenological study describes a small number of individuals' lived experiences regarding a phenomenon. A phenomenological study was best suited to dive into this assessment framework because it endeavored to contribute to theory and practice. Also essential to this research is applying critical theory and how it influences or interacts with the experiences and interpretation of positionality when examining standardized assessments' thoughts and decisions. I am interested in developing a design for assessing elementary students and teachers unpacking the results. In contrast, quantitative data is often gathered through surveys and questionnaires that are carefully developed and structured to provide numerical data that can be explored statistically and yield a result that can be generalized to a larger population. As I delve into this phenomenon, it is important to keep the following questions at the forefront of our minds:

1. How do you feel about standardized assessments?
2. What do you do with the result from the assessments?
3. What weighs more than classroom or standardized assessments?
4. If you had the option to remove a standardized assessment which one would you choose?

There are a variety of educational, passionate, and self-worth questions for the interviewees to answer. Most of this study follows the qualitative approach, and the qualitative framework guides the study's design. However, 20% of the data is collected through quantitative

measures. To support and strengthen the complete picture of various areas, such as teachers' and students' perspectives, create frequency tables and charts. Keeping all this information at the forefront of diving into this research of my siblings' and I's unfairness during our elementary school years of learning, I had to choose a route that would connect them with theory. My challenge was to view others without incorporating my personal biases, character, and beliefs. I had to become an outsider interpreting and connecting with other educators' educational experiences, personalities, natures, and critical thinking approaches. I was intrigued to discover how others that encounter the same standardized assessments embrace this culture.

### **Phenomenological Study**

The purpose of this phenomenological inquiry is to understand the phenomena within the equity of standardized elementary assessments. According to Creswell (2007), a phenomenological study “describes the meaning for several individuals of their lived involvement of a concept or a phenomenon” (p. 57). At this stage in research, the assessments will be generally defined as the process of being educationally evaluated. I will collect data from the interviewees, denoting similar themes of what they experienced and how they experienced it. This process will be taking a constructivist epistemological stance. As a researcher, I am placing myself within the qualitative paradigm. I must set aside all preconceptions, judgments, or prejudices towards their topic to try to be objective as I analyze the data participants convey to an inquiry.

Additionally, one of the principles of all qualitative investigations lies in the participants' perceptions as protagonists of the studied phenomenon (Padilla-Diaz, 2015). The most appropriate data collection strategy for phenomenological research is the profound interview. Existing literature (Kvale & Brinkman, 2009; Marshall & Rossman, 2010) concedes that the

phenomenological should be open or semi-structured. These two types of interviews call me to address the phenomenon profoundly, providing a space where the informants can express their experiences in detail, approaching reality as faithfully as possible. The detailed descriptions or interpretations of the participants' profound phenomenological interview should represent experience reality as possible. The phenomenological interview's primary focus is describing the meanings of phenomena from the participants' perspective (Rubin & Rubin, 2012). These authors point out that it is recommendable to carry out some additional interviews or to verify the information obtained, allow the participants the opportunity to provide further detail or expand on the information offered, and, lastly, for the participants' final approval.

As a teacher, I felt that it was essential to reach out to other educators who have experienced assessments within the elementary setting. By doing this, I can have a conversation about whom the assessments are created for. Is the assessment benefiting the student or the teacher? If we, as educators, are not clear about what the assessments are addressing, what will we do with the data? It could be simply wasting much time, energy, and resources of our own and the students.

Phenomenology researchers encounter research to share their own experiences while revealing insights and worldviews regarding the phenomena discussed in the study. This open sharing allows me to openly contemplate and concedes her own biases, values, and beliefs extracted from such lived experiences. According to Creswell (2013), This ability to position oneself within the writing reflexively is crucial. Phenomenological researchers position themselves and their work on various approaches. There is the "3+1" framework; this framework enunciates three scopes on which phenomenological scholars differ: (1) concentrating primarily on the common or the particular' (2) managing "*pre-understandings*" using mainly description

or interpretation (Hopkins, Pratt, & Regehr, 2016). It is important to not adhere to a single position about these dimensions, but rather I should intentionally and reflectively shift across the various positions depending on the study's purpose and context.

John Creswell (2015) stated, "How we write reflects our interpretation based on the cultural, social, gender, class, and personal politics that we bring to research. All writing is 'positioned' and within a stance. All researchers shape the writing that emerges" (p. 215).

A significant amount of this study was done through interviews and personal experience as an educator in the elementary world. As an educator within the elementary education field, I would like to be the voice of those who do not have the ability or courage to discuss or evaluate standardized assessments that do not contribute to students' success path. As a teacher, student, and educator, I have a story that will bring awareness to administrators, parents, and other educators. This research will explore the history and delve into their research questions through interviews with other educators within the elementary school realm. Reflection and conversations with their peers will provide clarity and lucidity.

### **Critical Stance**

Another stance that underlines this study is Critical Theory. Critical Theory has a narrow and broad meaning in philosophy. German philosophers and social theorists believe that "critical" theory may be notable from a "traditional" theory according to a precise practical purpose: A theory is severe to the extent that it seeks human "emancipation from slavery," performances as a "liberating...influence", and works "to create a world which stratifies the needs and powers" of human beings (Horkheimer, 1972, p. 246). Because such theories' goal is to clarify and alter all the situations that enslave human beings, many "critical theories" in the broader sense have been developed. In a narrow sense, the critical theory provides the descriptive

and normative bases for social inquiry to decrease domination and increase freedom in all their forms.

Critical theory is a social theory focused on critiquing and altering society as a whole (Crossman, 2007). Its goal is to dig beneath the surface of social life and uncover the assumptions that keep human beings from a complete and accurate understanding of how the world works. Critical theory is a research orientation in which investigators study power imbalances and justice and seek to enable individuals to overcome constraints related to ethnicity, class, gender, and other social contexts (Creswell, 2013). Reviewing the history of assessments and sharing my story and the interviews of their colleagues in the elementary standardized assessment world provides them with the ability to examine their stance and their pedagogical research practices, which I have experienced. That is why I am choosing the critical theory approach as a perfect opportunity to explore the views and opinions of their colleagues.

## **Research Design**

### **Research Participants**

Using a criterion sampling technique, participants for this study were selected based on the three criteria. First, each participant must experience the phenomenon the study is describing. This phenomenon was the experience of standardized assessments within the elementary setting. Second, each participant must have been teaching for at least ten years. This professional experience offers them the ability to speak about various assessments that have been introduced over the years. Third, I purposely selected participants in the surrounding suburban school district. The educational district that I am currently in does not allow for outside research to be performed. So, I have chosen participants who have been amid standardized educational assessments and are willing to express their experiences through interviews and enriching

conversations. Connecting with these educators will allow them to understand and view their thoughts on standardized assessment effects within the elementary sector. The exchange of interviews and observations will be with three participants in the Northwest school district, their insight on standardized assessments, which are very prevalent, and their influences on classroom instruction. The educators I have chosen were based on their expertise in the world of elementary education. I had the pleasure previously to engage with them in various educational professional development seminars. Each educator has a different lens of assessments that may carry similarities and differences within their classrooms.

### **Sources of Data and Methods of Data Collection**

The data collection for this dissertation will be obtained through interviews, surveys, and observations, along with some notable accounts of my childhood experiences. Through the interviews, I attempted to document the participants' events and conversations memories, giving turning points, linear trends, and progression within standardized assessments. Before data collection, I explored their own experiences with standardized assessments. Conferring to Merriam (2002), this action is done for two specific reasons. First, the process made me examine the dimensions of the phenomenon. Second, this process made me aware of her assumptions, viewpoints, and prejudices. After the completion of these two actions, I bracketed all prejudices and assumptions. Merriam described bracketing as the process that "allows the experience of the phenomenon to be explained in terms of its intricate system of meaning" (p. 94). Groenewald (2004) described two types of bracketing, including bracketing conducted by interviewees. The author discussed the phenomenon without confusing jargon and bracketing, in which researchers set aside their perceptions. For this research, both forms of bracketing were employed. During this phenomenological study, data was collected using two interview techniques and one

observation for each participant. The gathering of data through interviews (a conversation that would inquire about standardized assessments new and old) within the classroom in comparison with essential ingredients of the historical methodology along with my values and positionality allowed them to gain a better holistic view of the influences on standardized assessments within the elementary classroom.

### **Quantitative Data**

A minimal component of the research is based on quantitative data. Quantitative data in the natural, social sciences, and sometimes other fields is the organized empirical investigation of noticeable phenomena via statistical, mathematical, or computational methods. I have selected various educators from an educational network currently teaching in the elementary school setting and exposed to the phenomenon. A survey was developed through google and was sent to teachers in various elementary grade bands from kindergarten to eighth grade.

There was a total of at least ten questions on standardized assessments, and a few of the questions included:

1. On average, during a 10-week quarter, how much of your instructional time are you prepping (breaking away from scope and sequence) for the standardized assessments?
2. Which of the following ways do you standardized assess your students quarterly in your class? (check all that applies)
3. To what extent does Teacher Created Summative assessments support your classroom instruction? Please answer on a scale of 5 to 1, with 5 being the most helpful and 1 being not helpful at all.

### **Accounts of My Childhood Experiences**

I am in the middle of six siblings. Their younger siblings always talked about the younger days and the different games they played in their early childhood days. During the study, this researcher sought assistance from their siblings for clarification or parts of memorable events that they remember happening during their childhood—this activated knowledge of those years and a different perspective and insight as an adult. I collected data samples until I have reached a level of theoretical saturation. French sociologist Daniel Bertaux, who has used biographies in his social mobility research, describes “Saturation of Knowledge” as learning a great deal from the interviewees within the first few interviews (Curtis & Curtis, 2011). After several interviews, I recognized patterns in the interviewees’ experiences (1981, p. 37). I sought out patterns and connected the social endeavors to expand my understanding of elementary standardized assessments.

### **Participants**

The selected participants of this study are three women, two African American women, and one Filipino. The three participants work in the same Northwest school but reside in Illinois. All of the participants have been in the educational field for ten to twenty years. The sample in qualitative studies is selected purposively and deliberately. It is based mainly on the participants’ experiential knowledge about the topic and their ability and willingness to share it with me (Efron & Ravid, 2013). For this study's purpose, Arena, Ana, and Pia (names have been changed to protect their privacy) will be the names of the participants. Afrena and Ana will constitute the African American educators, whereas Pia will represent the Filipina descent. Afrena and Pia have served in the educational field for over twenty years. They all have taught at the primary level tackling all subjects. Their duties included assigning homework, grading assessments, and documenting progress. They are about to instruct a variety of subjects and reach students with

engaging lesson plans. Ana is a professional with advanced preparation and experience in reading who is responsible for literacy performance for readers in general and struggling readers.

## **Interviews**

Each participant has consented to a semi-structured one-on-one interview that was audiotaped and then transcribed. These interviews took as long as 45 minutes to an hour in one sitting. This semi-structured interview is a meeting where the interviewer does not strictly follow a formalized list of questions. The participants were asked more open-ended questions, allowing for a discussion with the interviewee rather than a straightforward question-and-answer format. The interviews were conducted in an unobtrusive, convenient location based on the needs of the interviewee. I made the presumption that each interview is done at their home school, where all of the participants are employed. It has been noted that sometimes nonverbal communication is just as telling as speaking. During the interviews, I maintained descriptive notes, where I recorded their subjective reflections and insights about what was happening in the setting (Efron & Ravid, 2013, p. 90). At the same time, I observed, conserved, and interacted with the participants “with no display of inferring feelings or responses to what is being said.

Every interview followed an interview protocol. First, I set the stage for the interview session. First, I aimed to establish a comfortable interview environment and confirmed all equipment's working status before the participants' arrival. I made all necessary introductions to ensure that each participant is comfortable and relaxed; then, I discussed the ethical issues involved with the research study. Each participant was informed of their involvement in a research study from which they could withdraw at any time without consequences. Following this, I asked each participant to read and sign a consent (Exhibit 3) for research conducted through National Louis University. I provided each participant with a copy of the interview

guide, a review of the study's purpose, and the associated research questions. Depending on the length and time and the depth of the actual conversation, I conducted more than the initial interviews. If a break in conversation appeared, it allowed for modification or additional questions for clarification after analyzing the first interview transcript. It is highly noted that the interview process should not exceed more than two interviews. I did not have a preference on whom I should interview first. I wanted to be as flexible as possible, so I based it on a first-come, first-serve availability. I needed to see the three teachers on the same day due to my work constraints and the length of traveling to the school.

### **Interview Analysis: What It All Means**

The qualitative analysis examines similarities, differences, and constant comparisons (Lunenburg & Irby, 2008, p. 202). Using the constant comparative method, once the interviews were transcribed, parts of the conversations were noted for correlation to the phenomena. Realizing that only some information and parts of the interview conversation shared with them may be relevant to this research, I used the Interview Guide questions to serve as their compass to determine the breadth of the integral parts that will help establish the noteworthiness of their participants' perceptions of standardized assessments within the elementary classroom.

### **Observations**

Observations are one of the critical tools for collecting data in qualitative research. It is the act of noting a phenomenon in a field setting through the fifth sense of the observer, often with an instrument, and recording it for scientific purposes (Angrosino, 2007). I observed students in all three classrooms and the educators in their setting, observing their interactions and conversations during their engagement of standardized assessments. These observations were not taken place until all interviews were completed with each participant. The observations took

place in the participant's school homeroom. The observation lasted from 45 to 60 minutes long. They were nonparticipant/observer as participant observation, where I am an outsider of the group under study, watching and taking field notes from a distance (Creswell, 2013). This allowed them to record data without involvement with activity or people. It is noted that during observation, as an excellent qualitative observer, one may change one's role during an observation, such as starting as a nonparticipant and then moving into the participant role or vice versa. The participants chose the time of day when standardized assessments were given, and I could enter their class to observe the behaviors.

When the observation was completed, I compiled descriptive and reflexive notes. Descriptive notes recorded the description of activities taken during the observation, and the reflective notes reflect the process, reflections on activities, and summary conclusions about activities for theme development.

### **Triangulation**

"Triangulation" is a term that is often discussed within publications of qualitative studies. Typically, scholars mention "triangulation" in discussions related to how to assure the study's "trustworthiness" and "validity" (Seale, 1999; Tracy, 2010). Triangulation in qualitative research reflects the multiple ways of establishing truth surrounding keywords like reliability and validity. Triangulation is a means of inspection within the integrity of inferences one draws (Schwandt, 2001, p. 257), even from the minor response or action. In developing a holistic view of standardized assessments within the elementary form through an appealing voice, a triangulation involving their participants' different points of view, multiple data sources, and observations was formed by their participants and myself. I also brought forth the validity by

sensitizing themselves and the audience toward acknowledging our experiences with past occurrences through current daily situations that remain personal challenges.

Within my facilitating and organizing of the interview data, I separated them into subthemes of each phenomenon. In this analysis, I developed a holistic description to capture the participants' experiences and highlights in their own words, along with any statements and quotes. In turn, these words developed into clusters of meaning categorized by themes, patterns, and structures (Creswell 2013; Hesse-Biber & Leavy, 2011), that showed commonality with each other. Once again, this process employed triangulating the interview data with data gathered through other methods (Marshall & Rossman, 2016, p. 151) and comparing essential information from the participants and my own experience.

### **Ethical Consideration**

To prepare for their interviews and observations, I have developed a consent form and questions that embodied the project's purpose and those I felt would gain the most mileage in response. With the Institutional Research Board (IRRB) clearance, I can delve into their research. The interview process needed to explain my role and their interviewees' intent expecting full cooperation and participation. It was also imperative for the interviewees to know that their participation is strictly voluntary. The participants had the right to review and assist in editing any time during the interview by asking them to stop taping or ask that they speak "off the record." I explained to each participant that she had the right to terminate her involvement if necessary.

Ethical issues are relevant to research in general and are faced in every stage of the qualitative research process (Flick, 1998, p. 48). Ethics involved research acts as a moral guide of beneficence for the inquired to consider the welfare and effects studied on the people they

research. With qualitative research, trustworthiness gives the study merit (Lukenchuck, 2013, p. 110). One aspect of developing a phenomenological study is epoch bracketing is part of an ethical discussion. Moustakas (1994) consists of identifying a phenomenon to study, bracketing out one's experiences, and gathering data from some people who have experienced the phenomenon. It was essential for them not to exhibit any behavior of acceptance or disapproval while interviewing the participants. I did not want to create an environment of right or wrong feelings. I wanted the participants only to know that they understand the essence of standardized assessment experience. Once all data has been obtained, it was at that point that my life experiences and historical aspects played an intricate part in shaping the standardized assessment world. I sensed it is necessary to embrace their personal experience through a short autoethnography. In doing so lies the question of ethics in assuring that the reality expressed is accurate.

In conjunction with triangulating data sources and research techniques (3 Triangulation of Data Sources and Research Methods, 2018), I also used member checks after transcribed interviews. This process allowed the participants to confirm the findings. As set by the Institutional Research Review Board (IRRB) at National Louis University, the ethical standards for conducting research involving human participants were followed with fidelity. As outlined in IRRB's guidelines, all participants were provided an informed consent form with the option of complete anonymity or agreement of disclosure (i.e., name, affiliate organization or institution, and other identifying pieces of information shared through interviews). All participants received a name that would protect their real identities. As ethics in research primarily concerns and controls the tempo of the procedures that were applied toward protecting those who participated in the study, participants were made aware that alternatively, had the right to withdraw from the

study at any time for any reason, if so desired (Smith, 2003). I have attached the letter of consent to this study (Appendix B). Appendix B represents the detailed consent letter for each participant informing them of what they agree to do and making them aware of any risk that may exist.

### **Summary of Methodology**

The chosen research approach for this project involved a predominance of research performed through a personal connection. Critical Theory is a social theory toward critiquing and changing society as a whole, contrasting to the traditional method only to understand or explain it (Crossman, 2007). Supported by bibliographical research and small stories, I offered an accomplished justification of phenomenological knowledge using these methodologies as the best way to help readers understand the topic as the participants had embraced it.

I have begun receiving clearance from IRRB by taking the collaborative institutional training initiative training and passing. This training is dedicated to promoting the public's trust in the research enterprise by providing high-quality, peer-reviewed, web-based educational courses in research, ethics, regulatory oversight, responsible conduct of research, research administration, and other pertinent topics of interest of member organizations and individual learners. Ten modules had ten to fifteen questions with module. This training was very informative and required at least six hours of participation. Upon approval from the institutional Research Board (IRRB), I began the interviews and observations.

## **CHAPTER 4**

### **Data Analysis**

There are cultural biases built into testing, and that was one of the motivations for the concept of affirmative action – to try to balance out those effects.

- Sonia Sotomayor

In this chapter, I analyzed the data collected from interviews, a survey, and observations.

All of these data collection strategies enabled them to answer the research questions:

1. What are all these standardized assessments needed?
2. How do educators perceive standardized assessment usage?

This phenomenological inquiry-aims to understand the standardized assessment journey as perceived by the participating teachers. Additionally, within the analysis of the data, I interweaved my own standardized assessment journey. As preparation for analyzing the data, I created separate files by sorting the data according to participants or data sources. I sorted their data first by methods such as interviews, observations, and a questionnaire. I arranged and coded various themes for all methods used. To better understand the data documentation, I had to break it down into smaller units. Following, I describe the data collection and analysis process of each data method I used.

### **Data Collection and Analysis Process**

#### **Data Retrieval**

Research fills a vital role in society: How researchers make discoveries, perceptions are confirmed or refuted, ideas and insights emerge, and a theory is developed or refined. “The researcher's role in qualitative research, in general, and phenomenological inquiry in particular, is to attempt to access the thoughts and feelings of the study's participants” (Sutton & Austin, 2015).

The open dialogues I conducted with the participants of this study to gain insights into their perceptions of standardized tests took place in a school setting. While it was easy for them to engage their interviewees in open conversations, I had to learn how to extract information without becoming too close with the subject or agree with the participants' points of view. I allowed them to reveal their thoughts and insights without interfering with and reinforcing their positions.

### **Interview Process**

It took them one hour and fifteen minutes to arrive at the Northshore School. I left my house around 5:45 a.m. when the skies were pitch black, but the birds were chirping; the traffic was light as a Sunday morning, only to arrive at the destination at 7:05 a.m., traffic had picked up, and the sun was shining brightly.

Upon arrival at the school, the host teacher greeted them at the back door, where the property tour began. The noise level was minimal, art sculptures were hanging from the ceiling, and classwork was on the bulletin boards. I always enjoyed looking at the bulletin boards because, in their district, certain things must accompany displaying students' work. These bulletin boards displayed no rubrics nor any common core state standards. The board unveiled student work with minimal feedback along with their grade percentages. The pale tan walls were full of life and meaningful work. I was able to walk on both levels and examine the quiet school climate before students arrived and during students' noisy sounds entering the building. Some of the students, I noticed, arrived early to meet with the art and music teacher to touch upon their latest art display, which was on sale (I bought two pieces of artwork) or go over a musical piece they had been practicing. I did not notice anyone in math, science, or language arts who came for additional support.

The bell rang, and all those wondering students and teachers made their way to the gym. During this time, I was welcomed into the school's morning meeting, where the entire faculty and student body had gathered. The principal introduced them and explained that I was there to observe classrooms and speak with students and teachers concerning a particular topic. I was surprised by the wonderful welcome that I received. For the three teachers' interviews, I implemented a semi-structured conversational interview method. I interviewed in person each of the participants at least twice. There were also multiple follow-up phone calls to clarify the transcription of certain interview recordings sections. Even though the interviews were done during their 60-minute prep periods, it did not seem very long; there were often only five to ten minutes left. In abiding with the ethical research rules (i.e., the guidelines of IRRB), all names and identifiable attributes have been changed. I met with three participants: Afrena, Pia, and Anna (not their real names). Afrena, their first participant, had an early morning prep, so we met to discuss the previously prepared standardized assessment questions.

### **Data Analysis Process**

I gathered the data by recording the interviews in audio form and transcribed the teachers' interview responses. I analyzed the data to discover patterns and identify reoccurring themes. Generally, qualitative data analysts face the task of recording data via a variety of methods (interviews, observation, field notes), coding and categorizing (using a variety of clustering and classification schemes), attaching concepts to the categories, linking and combining (integrating) abstract concepts, creating theory from emerging themes, and writing an understanding. Metaphors are helpful as interpretive tools in this process, serving a heuristic (guiding) role or explaining the theory elements (Bogdan & Biklen, 2003). I coded everything in the transcripts with labels that depicted the meaning of each of the segments. Then I identified themes around

which I grouped similar segments. To organize these themes visually, I created anchor charts to hang on their office walls. In the left column of the chart, I wrote the name of the category (the theme), and across the top row, I wrote the name of the interviewee in its column. Then I saturated the chart with the words spoken by each interviewee, words, and expressions connected to the category. I did that by working from the top to the bottom of each column and moving from one interviewee to another.

After coding the interviews to note themes that emerged from the data and having represented this data in a chart, I read each row of the chart across from left to right to note the frequency with which the category was touched upon by each interviewee. The data I gathered proved to be rich and exciting, and the analysis task at certain times was quite eye-opening. In many ways, many of the participants described in their standardized assessment experiences were similar to what I read in the literature. In the following section, I present the interviews with each of the three participants.

### **Interview with the Participants**

The study participants are educators who have been immersed throughout their work as classroom teachers in various elementary standardized assessments. They are dedicated, passionate, caring educators that have been in the profession for over 40 years combined. Educators are primary educational sources that can reflect and give in-depth information about the standardized assessment process. Educators comprehend that the business of coming to see our students as learners is too important to leave to chance—and that the peril of not undertaking this inquiry is not reaching a learner at all (Powell & Kusama-Powell, 2011). For years, these participants, who have been, amid the results of constant standardized educational assessments,

were willing to express their positions by sharing their experiences through interviews and enriching conversations.

### **First Interview**

Afrena is an African American woman who stands about 5 feet tall. She is a very short woman, around 4'9", but I was taller than her due to their heels. I will never forget her red cat eyeglasses on her small frame face. She was so well-poised, from her well-manicured hands to her long sleek black stylish ponytail. She spoke in a low soft tone; all decked out in a maroon a-line skirt accompanied by a white top and maroon boots, not a wrinkle in sight. Afrena has been an educator for twenty years in public and private schools. She has taught 1<sup>st</sup>, 3<sup>rd</sup>, 4<sup>th</sup>, 6<sup>th</sup>, and 8<sup>th</sup>-grade students. She is very active within her school family, serving on various educational committees, always willing to support new learning endeavors, and growing within her educational craft. She has classroom sizes ranging from 18 to 32 students.

Before the interview began, I presented her with the consent letter to review and sign, allowing their time to read and address any concerns. Upon signing, we could start the interview. I intended to meet with all the interviewees after school had ended, but they all felt stressed about that time and stated that they wanted to meet earlier, so I obliged. Afrena had an early morning prep, so she was chosen to go first. Since the music teacher was absent, we met in that room. We sat at a kidney table side by side. These kinds of tables are used for guided reading or small math groups.

Upon retrieving their recorder, I noticed her brown eyes get a little larger. I reassured her that the interview process is all confidential, and this was to support them in transcribing the interview. I started our conversation with a few getting-to-know-you questions to loosen up the participant, hoping that she would release the tension I saw in her eyes. After that third question,

I could tell that she was ready to proceed. These questions allowed them to connect with the interviewee and develop a platform for a friendly dialogue. I was so glad that I was not taking notes because I felt that she watched their every facial movement. When I would write something down, she would project her body over to see what I was writing, so I stopped taking even minimal notes and focused totally on our discussion. The following excerpt from our interview explains how she feels about the standardized assessment and its impact on instructional time:

Afrena engaged in a very passionate discussion on lost instructional time. She went on and on how she has not executed her entire grade level curriculum due to so many standardized assessments. Spending so much time developing a scope and sequence can be very daunting when all CCSS has not been exposed to the students. Afrena felt abandoned during instructional time. For two weeks, the interviewee, Afrena, has allotted at least 3½ hours of instructional time to prepare for the upcoming test. She has deviated from her scope and sequence of teaching. "I never thought teaching would turn into so much testing," Afrena shared with them in looking back at her testing schedule. School schedules are meticulous and demanding. With all of the standardized assessment requirements placed on students and teachers, she said it can be challenging to find time for the quantity and quality of instruction students should receive. Afrena seemed quite disgusted when speaking about the ongoing assessments, "The district is demanding more assessments every year." She related how over the summer, she reviewed the upcoming curriculum, reviewing various books, seeking out educational games, and looking for field trips, all intended to support students' development, all to be cut down because of the standardized assessments. Pia, the second interviewee, would whisper during her interview, "it is crazy how some districts are trying to lengthen the school day or year to regain instructional time

lost." As a nation, Pia's concern was that we are headed in the wrong direction due to the frequent assessments. Strauss (2015) would agree with Pia as he noted that if all that testing had been refining us, we would have been the highest-achieving nation in the world.

### **Second Interview**

It is the third period now, and I am walking into this aqua blue room. There are six tables and twenty-eight chairs, along with one teacher's desk. The teacher sat with her spiked two-tone blond highlights and blue eyes, a green knee-length skirt, a yellow cardigan, and white gym shoes. She has more freckles than I have ever seen, along with her outspoken earrings and necklace. Ms. Pia is of Korean descent 5'5" woman. She brings to the interview her educational experiences of 20 years. Pia has taught in the international spectrum for nine years and 11 in the United States public arena. She has a strong Korean accent, and you can tell she is aware of that because she speaks slowly and nods her head to make sure you understand what she is saying. Pia has taught kindergarten, fourth, and seventh grades. She has a very nurturing presence about herself. Students were very excited to see her and converse about their classroom learning when walking through the halls together.

While walking towards her desk, she stood up to greet the researcher, and we began to speak about how their day was going. I felt this was the right way of easing any tensions she may have about being interviewed. This was her prep period, so we only had sixty minutes to conduct the interview. I pulled out their recorder and consent form for her to review before we began. I believe that having a one-on-one five-minute talk before the interview and asking getting-to-know-you questions can release a great deal of tension. She spoke at significant dimensions about how she felt about standardized assessments. Although the conversation was in-depth, I noticed that she kept looking at the door and whispering at times. It gave me the impression that,

for some reason, she felt that someone was listening, but still, despite it, she wanted to speak her truth. Pia was very passionate about how many of her students break down with tears when taking an assessment.

In a fast-moving culture full of pressure to accumulate good grades and broad experience portfolios, student stress and anxiety move down the chain and affect middle, elementary, and high school students. Pia stated, "so many of my students speak to me in conferencing, noting their fears of not doing well on their standardized assessment." Pia noted the tears, fears, and frustrations of many of her students. She relates that many students feel that these standardized assessments will hold them back from the next grade. Pia explains, "It brings tears to my eyes when I see students break down because of the standardized test." She is very compassionate towards her students. She goes on to reply, "I have witnessed students who do not have an IEP shut down, and I know they are capable."

### **Third Interview**

My third and final interviewee was an African American 5'8" woman with a powerful voice, Ana. Walking in the hall with her is like walking with the principal. When the students see her coming, they begin to straighten up from any foolishness. I would not think that students would have this reaction to her based on her relaxed appearance. She wore baggy blue jeans, a school sweatshirt, and Nike gym shoes. Her squinted eyes hardly showed the hint of hazel that they contain. Ana mentioned that most people think she is Caucasian due to her long hair and bright skin tone. She carries a southern dialect from her home in Alabama. Although the school walls are tan, the classrooms carry vibrant colors, and hers are green. The room has 28 desks with 24 students. Ana has contributed 22 years of teaching elementary-level students in public schools. Although she holds a type 75 certificate, she feels that she can better serve students

within the classroom. She has had to teach all grades within K-8, either in summer school, after school, or in regular education classrooms. She boasted about every year; she has constructed a drive for school supplies. She has been under the leadership of more than eight principals in her profession. Ana did not speak very highly about the administration within her building. Ana feels that there is not enough administration support provided to classroom teachers.

We sat down at the classroom kidney table, not needing to hold an introductory conversation. Ana immediately signed the consent, and before I began to record, she put up her finger to announce that we needed to shut the door. After the interview, she gave Iher home number to converse further if needed after this week was up. However, she did whisper as if someone was listening to her, sharing her frustrations with standardized assessments. "My students should be promoted because they are not good test-takers," Ana adds. I agree with her that standardized assessments are not a valid form of a students' ability.

Like the other two teachers, Ana felt frustrated that she could not help her-students, who are poorly affected during the standardized tests. She felt very strongly about students' anxiety during a high-stake test; as she stated, "Have you seen them? Have you ever seen them look to the ceiling, look at the floor, stair outside the window, am I supposed to be immune to the frustration I see them go through?" She goes on to say, "Those tests measure a certain type of skills that are not even taught – our focus has come to only focus on what is on the test, year after year after year." This conversation makes I feel like many teachers are losing their creativity and academic freedom in exchange for reading a scripted text. Whenever I attend their school's principal, they and many educators request more academic freedom to stay creative and engaged.

Ana noted that, fair or not, test participation and performance have become school and district performance measures. This situation adds to the stress on students. Standardized tests

have a tornadic consequence on schools. Administrators pass the heaviness on to a previously stressed staff. Teachers fear losing their positions and often wind up leaving the profession altogether due to test preparation pressures. The three teachers I interviewed shared lots of information on the anxiety/stress of their students.

Students are so upset that they throw up on the testing day (Eggers & Kaul, 2018). Ana noted, "this article demonstrates the troubling students are facing with standardized assessments." In my interview with Afrena, she also added that "It is developmentally inappropriate to ask third and fourth-grade schoolboys and schoolgirls to sit still and be quiet and to take a test masterfully for several days in a row." Many teachers have watched students read a passage repeatedly, trying to find the correct answer, squirming, stewing, visibly quaking with energy, effort, and frustration.

Figure 2 is a schedule of standardized assessments provided to me by the interviewees. This schedule outlines various assessments given to elementary public students K- 8 in a ten-week semester from September to November. The dates allow you to denote how much time students are in the standardized assessment mode. Although categories are not given within each assessment, the interviewees spoke about how the data points continuously overlap.

## **Figure 2**

Elementary School Assessments Within a Quarter

Elementary School Assessments	
	<b>REACH Performance Tasks BOY</b> September 16 - November 1
	<b>Kindergarten Individual Development Survey (KIDS)</b> <i>Observation Window:</i> September 3 - October 30 <i>Data Entry Window:</i> October 23 - November 8th.
	<b>TRC/DIBELS, mCLASS Math BOY</b> September 3 - October 11
	<b>NWEA MAP for Primary Grades (MPG) BOY</b> September 3 - October 4
	<b>Fountas &amp; Pinnell BAS BOY</b> September 3 - October 11
	<b>NWEA MAP BOY</b> September 3 - October 4

As the figure demonstrates, students in grade band K -2 are engaged with four standardized assessments. These assessments are given in a one-on-one environment to every student. The classroom teacher must administer the evaluation within the calendar dates given. In grade band 3-5, students have two standardized assessments given during the the12-year-old semester, and in grade band 6-8, on average, two assessments are given during a ten-week semester. This schedule does not consider the accountability of quizzes or tests (formative and summative) created by teachers. It displays where tests have been piled upon each other under mandates from Congress, the U.S. Department of Education, and state and local governments.

## **Interviews Summary**

Afrena's words made I think that it is such an awkward position for teachers to assess where they are the "administrators" of standardized assessments, usually without training or reviewing the materials' responsibilities. When a teacher becomes a "test administrator," procedures are developed for an exam program to help reduce measurement errors and increase the likelihood of fair, valid, and reliable assessment. Specifically, appropriate standardized procedures improve measurement by increasing consistency and test security. There is no teacher support involved. Pia adds that this is done "In addition to the regular teaching responsibilities, teachers must read and understand a 100-page booklet about the test, its administration and procedures". Teachers and parents use test scores to create a baseline of student's academic strengths and weaknesses.

Additionally, communities rely on these scores to judge the quality of their educational system. States and federal lawmakers use these same metrics to determine whether schools are up to par. However, when the economic and emotional stakes associated with standardized tests are disproportionately high, this laudable goal gets distorted. Listening to all of the interviewees discuss how the loss of instructional time has affected them left me thinking what a disservice we are doing to our children today! They even shared a document (Figure 5) on how their district will be decreasing subject percentages. Seventy-one percent of the nation's 15,000 school districts had reduced the hours of instructional time spent on history, music, and other subjects to open to more time for reading and math (Dillion, 2007). In some school districts, educational programs and institutions are on the chopping block if they do not improve test scores. Recess, class meetings, and even whole subjects such as music, art, science, and history are being cut back or eliminated to focus on the only tested areas yearly.

## Observation Data Collection & Analysis

In this assessment or rather, keeping in mind that phenomenology is more a way of thinking and feeling than a formulaic method, the activities with this part of the method I organized the data and mediates meaning, looking for patterns and identifying themes; analyzes and synthesizes; employs "thick" description to capture the participants' lived experiences; highlights significant statements or quotes in order to develop clusters of meanings into themes, patterns or structures; balances parts to the whole and explores the subjective, particular experiences of individuals regarding the phenomenon, as well as the objective, broad experience understandings that the group has in common with each other (Creswell, 2013; Hesse-Biber, 2011).

I visited each of the participants' classrooms twice in one week. Each classroom had at least 23 students present. The observations lasted for one class period of sixty minutes. Although I was introduced during the school's morning meeting, the teacher reintroduced me as a proctor in the classroom. I did not want her to explain that I observed students' assessment behavior, making some students aware of it and changing their demeanor. Many students had already begun to speak and asked for my name while passing in the hallways. Students became very friendly with their greetings and approached I with various forms of questions "Are you a new teacher"; "what school do you teach at"; "do you like teaching" just to name a few. I gathered information by taking notes of what was occurring with student behavior and body language, making a logical explanation of the observation drawn from prior knowledge or experience. Being in the standardized assessment mode (desk-separated) allowed me to have a good view of the class. When this observation was over, I created colored index cards with subheadings to make connections when applicable to the corkboard hanging on my office wall.

## **Classroom Visit Observations**

On day one, I visited each interviewer's classroom not for discussion but students' observation while taking their assigned standardized assessments. Each interview took place during the educator's classroom special. Walking through the busy school's pale hallways of tall, short, outspoken, quiet, ambitious, energetic, friendly, cheerful, and pleasant students was not orderly. While moving through this hallway of minimal teacher presence was at a voice level of 2. When speaking at a conversation level, two students spoke as if they are in a one-on-one conversation with their peers, not as if they are in the cafeteria or on a field trip bus. Although I expected to see students in a single file line being escorted by their teachers, that was not the case. Students in grades K through 2 had the most visible teacher escorts, while the middle and upper-grade students' educators were a lot more lenient with the switching of classrooms. I seemed to mesh in as classes as they passed by. Some of the students caught this strange face passing by, giving a second look with squinted eyes, and others walked right by them as if I had belonged in the building.

### **First Observation**

Entering the first door with a vast window allowed them to peek in before immersing themselves with twenty-eight 7<sup>th</sup> grade students. Some of them greeted the teacher as they walked in; others grabbed a laptop and proceeded with the board instructions of successfully logging into their assessments. There are five minutes of dragging desks apart from each other to ensure privacy. The blue walls, which once housed various strategy posters, are now all covered with anchor chart paper. I asked why they were covered, and the teacher answered that students could not have any type of support while testing. Some of the anchor charts/posters on the wall were created by the teacher, students, or store-bought. A proctor entered the room after I did. This is

part of the assessment administering process that is meant to keep everything according to the rules among the students and the classroom teacher. It has been ten minutes, and the classroom teacher rang a bell to pull everyone's attention to her. She stated the norms of testing and asked if anyone has not entered the right site. No one responds, so she advises that the test has begun. Initially, I sat in the back of the classroom, but I could not see the students' faces, so I began to walk throughout the room just like the proctor. Fifteen minutes into the class, I saw students squirming in their seats. I thought they needed a restroom break, but I was advised that they just had one. At least two students finished this 54-question test in fifteen minutes and immediately put their heads down. This, I was told, is the procedure when one is done before the allotted time. I could not believe their eyes; these upper-grade students were exhibiting distressed behavior. At least ten students were continually looking at the ceiling as if they were observing a mosaic painting. I watched the behavior of one of those students. I noticed how he gazed for a few minutes and then returned to the screen only to make a quick selection of a question. This behavior went on for the entirety of the assessment. When the teacher announced that time was up, many of the students sighed a sigh of relief. The teacher questioned the students on how they felt they did, and they responded, "fine." I was amazed when there was so much squirming and dazing going on. Do not get me wrong; some students were very focused and in deep thought. I would say at least one-third of the class portrayed this behavior. At the end of the session, and the class was over, a student approached the teacher's desk and asked if she could go to the restroom. The teacher asked if she was ok; the student nodded a very quick yes. She did this all while her head was in a downward position. I saw the teacher make eye contact with the aide to follow the student out, and she did. The aide reentered without the student, and she and the classroom teacher engaged in a small conversation. The teacher allowed students to stand up and

stretch before heading to their next class. I stayed behind to ask what the issue with the young lady who left was. The teacher advised them that she was in tears and felt she did not do well on the test. The young lady returned, looking quite refreshed, gathered her things, and headed out with the class.

### **Second Observation**

I had a 15-minute break before heading to their next class. When I opened the wooden brown door, it made such a creaking sound that all eyes were on them upon entering. There were rounds of tables with at least 4-second graders sitting at each one with these fancy cardboard dividers between them. Each student had an iPad that had already been set up for the test. All they had to do is select their names and push yes to start the assessment. The teacher advised students to take their time and read the question at least twice and inform students that he could not provide any help or read any of the questions. I noticed the yellow walls were covered with brown paper to hide any alphabets, posters, mathematical expression/strategies, anchor charts, or sight words. As I began to move about the room, few students kept focusing on their movements, so I focused on two tables for this observation. One of the 2<sup>nd</sup> graders started to doodle on the table when the teacher observed this behavior; she instructed him to stay on task. Ten minutes into the exam, many of the students were asking how many questions there were. There were so many bathroom breaks I could not keep track of. Many students had technology issues: their iPad froze, or the screen would not advance once the answer was selected. No students were crying, just a lot of antsy little ones. A ritual was already in place for students to immediately engage in the Go Noodle exercise assessment. Go Noodle is a web-based video, game, and activity focused on introducing a short burst of physical exercise in the classroom. It seems that students were more excited about the activity than the assessment.

### **Third Observation**

Upon entering this gray steel door with a large glass window, I peek again before entering the classroom. There were no students present, so I quickly went in to find a place to observe before the students arrived. That is when I saw Max coming. This four-foot height, weighing no more than one hundred pounds, little brown hair boy is wearing brown glasses that allow his round blue eyes to see better but not hide flat light brown freckles. He walks in a strolling manner observing his surroundings with his mangled clothing. Not engaging with any of the students, I saw his name tag on the desk.

It is 7:45 in the morning when Max enters the classroom; fifteen minutes later, the morning bell indicates that breakfast is over, and the morning pledge will begin after the 7:50 bell rings. I observed when he makes eye contact with the teacher, he grunts "good morning" and slumps down into his seat, placing his head on the desk. I noticed how the teacher resisted asking him to repeat himself before making a way to the morning meeting carpet. It is vital to understand Max's outside factors and traumas in his life before he even crosses into the classroom.

In his morning's writing prompt, his teacher has shared that Max has expressed his feelings about his parents' divorce while struggling with reading and math. Instead of doing his morning routine of answering the writing prompt, the teacher asks him to join him after the meeting before the test to engage in breathing techniques, which has helped him develop a growth mindset. As Max navigates through these breathing sessions, he observes that Max is ready for the standardized assessment scheduled for the class, or will it be another case of Max rapidly clicking through the test simply to complete it. The teacher tries to honor Max's feelings while explaining that the assessment scores help inform what they should continue to work on

for Max to be educationally successful. Max's performance anxiety and emotional pressure on himself to maintain good grades are something that I can relate to. I did not do well in grammar school, so I also exhibited performance anxiety and emotional pressure in college and high school to prove that I am worthy. The district requires four standardized assessments in ten weeks. These assessments are directly linked to school quality rating, funding, programming, and in more extreme cases, keeping the school doors open.

Once everyone has a laptop, the teacher talks them through the instructions for logging in. He monitored their progress in answering the questions very carefully without giving any educational support. There were no posters on the walls but brown paper-covered anchor charts on the three closet doors. It was so silent you could hear a pin drop in the classroom. No one asked to go to the restroom; there were a few squeamish bodies, especially Max, along with daydreamers looking out the windows. Students seem to be at ease and focused. I was able to walk throughout the room to observe other students, but I always kept an eye on Max. I observed that one-third of the class finished before the allotted time, which included Max. It was undeniable that he rapidly clicked through the test. When the test was over, the teacher allowed students to write about their experiences while playing a classical music piece. She called their numbers, and they returned their laptops quietly and in order. Although he had observed Max's behavior early, it had already devoured him from taking a successful exam. We both observed his writing, where he barely wrote anything within the 15-minute free write. He mostly drew a picture without labeling it.

### **Observation Summary/Analysis**

Students exhibited quite a bit of anxiety and frustration while taking their standardized assessments. Many of them showed signs of relief when the assessments were over. While

observing in that second-grade class, their attention span went so quickly that they did not dive into what was being asked of them. The upper-grade students did not seem to display the urgency of excelling as well. I excused myself from speaking with the teacher so that others would not cloud their thoughts nor observation. I only wanted to reveal what I saw on those days I observed.

When I returned home, I immediately reviewed their notes, noting what behaviors the students had exhibited during the observations. Next, I sorted and organized the data into specific files. These files, upon their review, shared many similarities. All of the classrooms I observed had at least one, if not a few, students who were requesting washroom breaks during the test. It was evident to them 15 minutes into testing that much twitching moving side to side in their seats, was happening. When I would move out of the students' eye view, they would gaze out of the window or the ceiling for periods. I timed a few of them, and they were staring from five to ten minutes before returning to the assessment. Standardized assessments can be intimidating, long, arduous, and confusing. Sitting in the rear of the classroom, I heard a young nine-year-old student sniffing over and over, so I handed her a Kleenex, but she used it for her eyes. It was then I realized she was exhibiting a low cry and merely wiping the tears from her eyes.

Once the most prevalent assessment behavior was identified, I would write the behaviors observed on colored post-it notes, developing columns on the left and tally marks on the right. These behaviors were then compared to the interviews, as well as the questionnaire distributed.

### **Table 1**

### *Themes and Descriptions*

---

Theme	Description
Reduction of Assessments given in one quarter	lower the number of standardized assessments
Loss Instructional time	Preparing for assessments takes away from the amount of time during which learners receive standard based grade-appropriate curriculum from the classroom teacher
The weighting of Standardized Assessments promotion	Some of these assessments are used for the
Student Anxiety	becoming overwhelmed, crying, twitching, shutting down, clicking on any answer to finish

---

I presented these themes or categories that arose from having coded data and developing meaning clusters based on their own "insight, intuition, and impression (Creswell, 2013, p. 182). Walking into these classrooms, I was able to view students twitching, window gazing, and hurrying through their standardized assessments. All of the rooms had students becoming disengaged with their assessment materials very early. The educators seemed very frustrated in not being able to give students some type of guidance. Observing these classrooms gave depth to understanding standardized assessments from a teacher and student perspective.

### **Survey Data Analysis**

After interviewing all participants in this phenomenological study, a survey supported the commonality of a particular group's lived experience. It also added to make the identified themes more credible and, ultimately, valuable—qualitative and quantitative research deep

understanding of the problem differently (Austin & Sutton, 2014). Statistical procedures can help qualitative researchers by noticing through numerical findings trends not seen in the qualitative that point out relationships between variables (Corbin & Strauss, 2008).

I have developed a standardized assessment survey for classroom teachers in conjunction with qualitative interviews and observations. A survey is defined as "the collection of information from a sample of individuals through their response to questions" (Check & Schutt, 2012, p. 160). Surveys permit researchers to obtain a large amount of data in a relatively short period. The survey, which can be found in the appendix, was designed with 20 questions that allow teachers to share their thoughts on standardized assessments through multiple choice or write-in options. These options were ranked based on their percentage responses. I reached out to 40 teachers. A large amount of them came from their school district. I chose several schools that I had visited as well as the one I currently teach at. Those teachers I contacted had shared with them through messaging that they had friends who are educators in the surrounding North and South suburbs who would be interested in answering the anonymous survey. In the expiration, I was able to obtain 67% in responses.

### **Survey Analysis of Themes**

It was three days past the survey deadline date when I reviewed the data. Each question was placed on a sentence strip on their wall. Underneath each strip was the percentage outcomes. I was able to retype all of the write-in-question submissions. These submissions were then highlighted with various colored pencils based on word repetition. Constructed on the analysis of the survey, the following themes emerged: (1) reduce the number of assessments, (2) lose instructional time, (3) student anxiety (4) less weight on standardized tests. The remainder of the chapter will address the phenomenon from themes identified within the research survey findings.

## **Reduce the Number of Assessments**

It has been my experience that assessments have long held a strong influence on educational practice. It is not hard to find a teacher willing to bend your ear about standardized testing volume in schools today. Nearly half of the responses to question 20 in the survey wanted assessments to be reduced. Question 20 was as follows: What kind of changes would you like to see in the standardized assessment process to better support your classroom instruction? The responses to this question included the following:

1. The number of assessments reduced
2. There needs to be a reduction in standardized testing, as three a year is too much and unnecessary. The NWEA is the best of the three because it offers an insight into several skill sets specific to student needs.
3. Eliminate IAR; its redundant NWEA is sufficient.
4. One assessment a year is differentiated for students' needs.
5. Shorter meaningful test
6. Too many assessments cause students to disengage. Just let us teach our curriculum in alignment with the common core.
7. Too many tests! One good one will be enough!
8. Less testing, of course. Complete tests in fewer amount of days to maximize instruction
9. Less testing in general. We lose roughly five weeks of average instructional time per year to standardized tests that both teachers and students do not enjoy. I would not say I like teaching to a test. There is much more meaningful information to teach them how to do well on a standardized test

10. I would like to see less testing for the primary grades. Primary students should have one maybe two assessments for the year. They should be administered the assessment twice a year.

11. Not as many administered

### Lost Instructional Time

Every minute counts when it comes to student classroom instruction. Student success has not reached its full potential until students are given every opportunity to maximize their educational experience. The impact of losing instructional time to conducting standardized tests was questioned in the following survey question:

*20. What kind of changes would you like to see in the standardized assessment process to better support your classroom instruction?*

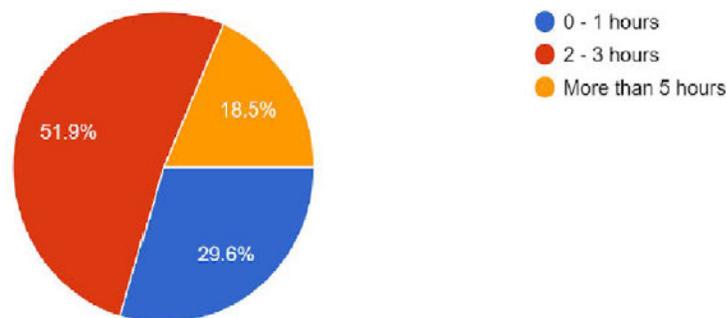
*Ans. "Less testing in general. We lose roughly five weeks of normal instructional time per year to a standardized test that both teachers and students do not enjoy. I do not like teaching to a test. There is much more important information to teach them how to do well on a standardized test."*

### Figure 3

#### Instructional Time for Prep

15. On average during a 10 week quarter, how much of your instructional time are you prepping (breaking away from scope and sequence) for the standardized assessments?

27 responses



Fifty-one percent of the teachers who took the 20-question survey spend 2-3 hours preparing students (exhibit III) for standardized assessment success in Figure 3. The diagram in Figure 3 also shows 29% of teachers spend 0-1 hours, and 18% spend more than 5 hours during a ten-week quarter in preparation for a standardized assessment of teachers. It was also noted in the poll results from the Northwest Evaluation Association that most teachers still think that too much time is spent on testing. These teachers break away from their regular planned lessons to prep students and prepare them for an unforeseen assessment with unknown depth questions.

### **Student Anxiety**

During the school day, opportunities often arise for producing useful assessment information for teachers and students. For example, in a class discussion, some of the students' remarks may lead the teacher to believe that they do not understand the concept of energy conservation. Teachers experience limitations in covering the curriculum; in turn, it can lead to various ailments.

### **Figure 4**

Limitations of Standardized Assessments

17. What are the limitations that you have observed within standardized assessments? (choose all that apply)

27 responses

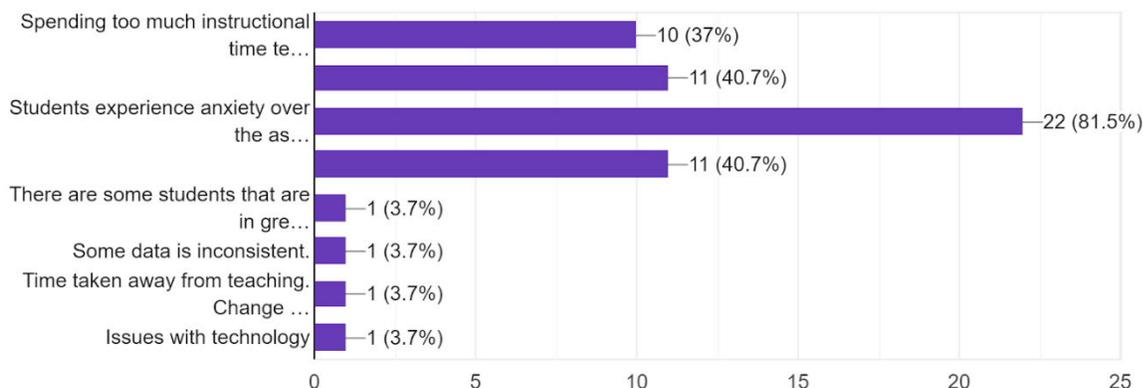


Figure 4 displays the percentages and outcomes of educator's responses to standardized assessment limitations. Question 17 (Figure 4) in the survey given to various elementary teachers denotes that 81.5% of teachers can identify with students experiencing anxiety when taking standardized assessments. At the same time, 40% of the educators noted that the data is not compatible with instruction and the technology not consistent with the assessments.

### Less Weight on Standardized Test

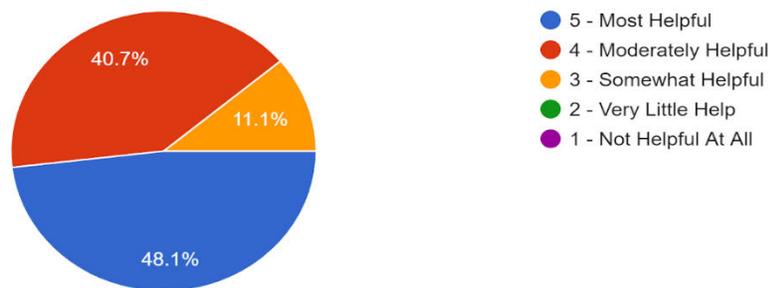
Standardized tests are created to be "fair" assessments that eradicate bias and unfair advantages. However, performance on the standardized test depends on the quality of the individual school system, material covered in classes, and financial status, all of which can be considered unfair advantages. In Figure 5, the outcomes for summative teacher-created assessment percentages are displayed.

### Figure 5

Teacher Created Summative Assessments

8. To what extent does Teacher Created Summative assessments support your classroom instruction? ... with 5 being the most helpful and 1 being not helpful at all.

27 responses



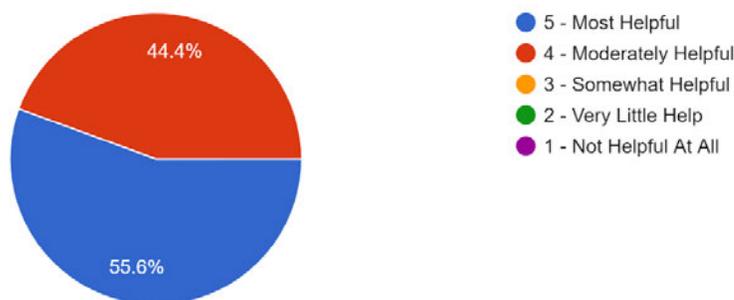
The below figure displays outcomes from educators surveyed on formative assessments strength within the classroom.

### Figure 6

#### Teacher Created Summative Assessments Pt. 2

9. To what extent does Teacher Created Formative Assessments support your classroom instruction? ... with 5 being the most helpful and 1 being not helpful at all.

27 responses



Figures 5 and 6 reveal a high percentage rate of teacher-generated assessments within the classroom. Having teachers create their tests is one way to counter the backlash to "over-testing" and give teachers better data to improve instruction. Based on the survey administered, educators have stated that commercially prepared tests often fail to provide teachers with timely, practical,

or actionable data to drive student improvement. Assessments designed by classroom teachers can better reflect what is taught in class and allow teachers the flexibility to choose the best format—such as presentation, essay, multiple-choice, or oral examination—to assess students' mastery (Lewis, 2019).

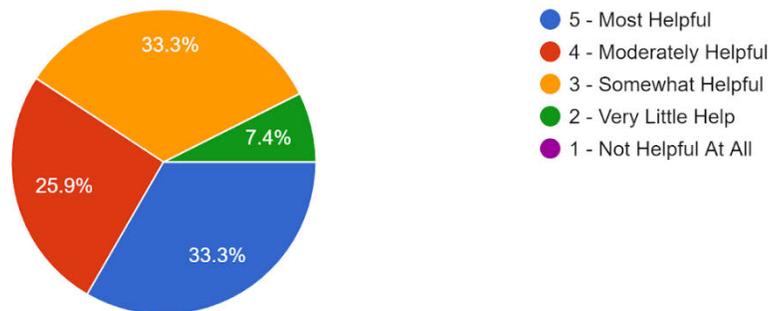
Figures 6, 7, 8, and 12 further support how standardized assessments are not well supported within the classroom. It is a tie in Figure 6, where 33% of teachers feel that the standardized assessment NWEA is most helpful and somewhat helpful, while the remaining 25% is somewhat helpful. Figure 7 revealed that the IAR standardized assessment is most helpful to 3.7%. At the same time, 25.9% believe that the IAR assessment is not helpful at all. Standardized assessments vary on the importance of classroom teachers. The IAR is not that helpful to educators. Students are administered the test in the spring, and the results are given in the fall. During this time, teachers have a new set of students to teach.

### Figure 7

#### NWEA Assessments

10. To what extent does NWEA Assessments support your classroom instruction?  
Please answer on a...ith 5 being the most helpful and 1 being not helpful at all.

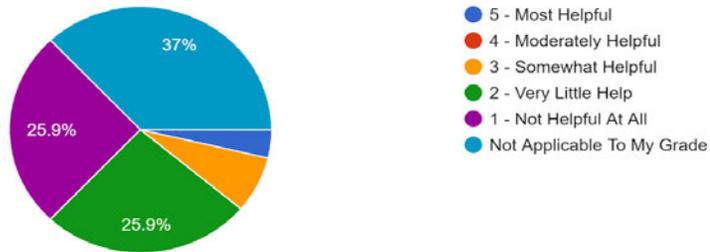
27 responses



**Figure 8**

IAR Assessments

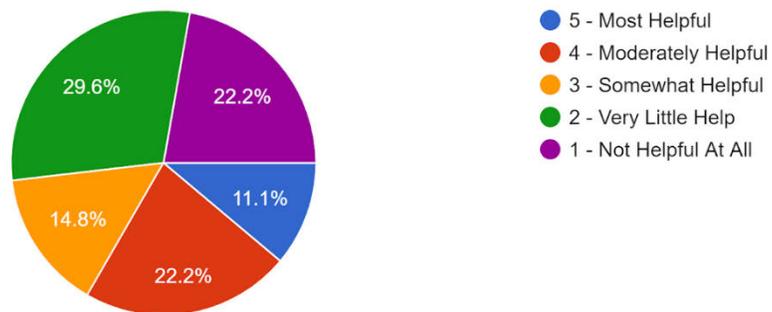
14. To what extent does Illinois Assessment for Readiness (IAR/Formally PARCC) Assessments support your classroom instruction? Please ans...g the most helpful and 1 being not helpful at all.  
27 responses



**Figure 9**

REACH Assessments

12. To what extent does REACH Assessments support your classroom instruction? Please answer on a...ith 5 being the most helpful and 1 being not helpful at all.  
27 responses



Figures 5 and 6 reveals how teachers embrace the results of formative and summative assessments; Figures 7, 8, and 9 tell a different story. Figures 7, 8, and 9 unveil the low

percentage rates on standardized assessments that teachers deem helpful in the classroom. There are many glitches with the testing treatment, but a massive issue for classroom teachers is that they do not help the teachers in their instruction. Folks outside of schools often visualize that one of the test's benefits is to review to see how students are doing and regulate instruction accordingly.

### **Chapter Summary**

As I expected, gathering this data became a very prodigious experience that allowed them to be a vessel for a few educators concerning this standardized assessment phenomenon. I had the arduous task of formulating themes and commonalities into levels of significant groupings. I wanted to develop a shared vision for those who have experienced this phenomenon by conducting surveys, observations, and interviews. The coding and themes were inspired by words connected through phrases and words affiliated with each other through the methodology.

The process helped them identify how the teachers who have participated in the study, those three teachers who answered the questions in depth during our interviews, and those who responded numerically and in short responses in the survey, felt about the value for and the impact of the surveys on their students and themselves. The observations helped them understand the students' reactions when standardized tests are conducted to triangulate the teacher's responses. In the following chapter, I will triangulate the data to capture the same phenomenon's different dimensions. This process is a way of assuring the validity of research through the use of a variety of methods.

As I examined this study's patterns, it became clear that all educators experience some type of frustration when it comes to standardized assessments. They have acted in ways that exemplify many variables that contribute to high teacher turnover rates due to visualizing their

student's frustrations and a large number of standardized assessments. Through the interviews, the educators had expressed their anger and pain of not being able to fulfill their teacher's responsibility of being a resource, mentor, providing support, and allowing students to become engrossed in learning.

Looking back on this experience, it has painted an original picture described by educators through interviews, observations, and surveys, of the standardized assessment phenomenon. All of this has brought forth emerging themes to form the analysis. I feel ready to take the next step in chapter 5, which will re-immense themselves in the data and provide recommendations for the next steps.

## CHAPTER 5

### Findings And Conclusions

*Looking at the past must only be a means of understanding more clearly what and who they are so that they can more wisely build the future.*

-Paulo Freire

In this study, the inquiry which inspired the research was guided by the following questions:

1. What are all these standardized assessments needed for?
2. How do educators perceive standardized assessment usage?

The four chapters up to now were focused as stages in answering these two research questions: In Chapter One, I described the problem and purpose of the study, its significance to the educational field, and its meaning for them on a personal and professional level and provided a contextual background to the issue. I have explored the current discourse on standardized assessment through the literature review in Chapter Two, outlined the specific methods to the research problem. In Chapter Three, and reviewed the process of analyzing the data that reflected the participants' perceptions of this issue and the classroom observations. In Chapter Four, I described the data I collected through the themes that emerged as I analyzed how the participants experienced the assessment phenomena and the insight I gained through classroom observations. I developed those themes within each data collection method into a culminating, composite description that highlighted the essence and underlying structure of standardized assessments within the elementary setting.

In Chapter Five, the dissertation study's absolute path, I interpreted the data, looking for similarities and differences within the meaning gained from various data sources and pertinent literature. The information, understanding, and meanings gained through the data analysis and

interpretation allow them to present the study findings and answer the research questions. I considered the practical significance of the educational field's findings, general and classroom teachers, and I assessed the study's limitations. I end this chapter by looking back on the dissertation journey and experiential paths.

Assessments have been given the authority to open or close gates for students' educational pathways. However, the data of this study has conveyed that we do not need all those standardized assessments. Many teachers do not have enough time to review the measurement data in-depth or understand what the assessments are covering in a way that enables them to prepare students for success thoroughly. This understanding is disheartening because teachers' decisions often focus on tracking and placement, promotion and retention, and awarding or withholding diplomas.

From my perspective, I believe that we urgently need to rethink how we assess students' work. Based on long experience as a teacher, I think there is far too much concern on accountability at the outlay of learning. Many of the methods commonly used also do nothing to reduce the workload for teachers. The truth is that children learn when they have opportunities to learn. Students learn much more when those opportunities are rich, engaging, and relevant to their life than when those opportunities are narrow, constrained, and focused on dry, unengaging test preparation.

The previous U.S. Education Secretary, Arne Duncan, who served during the years 2009 to 2016, recently conceded that too much standardized testing was "sucking the oxygen out of the room" and causing "undue stress." Although some of the nation's educators may have been encouraged by Duncan's words, most policymakers have spent the past decade ignoring teachers' calls to curb high-stakes testing (Walker, 2014).

Nevertheless, the Common Core State Standards Initiative was designed to ensure that students across the country receive a similar quality of education. The initiative sets specific standards for what students should know at the end of each grade to guarantee that every student has an equivalent opportunity to attend college. What is wrong here is that schools develop their curriculum around the common core standards, but rather than focusing on the quality of the education they offer, they mainly teach to the test. Teachers have to teach only subjects that will be tested or center their instruction only on the correct factual answers to questions expected to be on the standardized tests, rather than exploring the meanings of these answers or why they are right. You may ask why this occurs, and the answer is simply because standardized tests are linked to federal funding based on the outcome of standardized assessments. So, again I ask, "Are All These Assessments Necessary?"

This study aimed to provide educators a platform to share their stories and tell their truths as they have experienced them, and the in-depth interviews, the surveys, the observations, and my professional encounters allowed them to achieve that goal,

During the data collection phase, it was shocking to see how many educators, in casual conversations about standardized assessments within their classrooms, vocally expressed their opinions. Still, they did not want to be recorded nor mentioned by name. Although this research was anonymous, many teachers who shared their views about the standardized testing phenomena were leery about the fact that I was taking notes. Even though I was only discussing how standardized assessments should establish measurable and transparent students' learning outcomes, at times, it felt as if I was investigating some hot FBI topic. The participants were hesitant in sharing their ideas on whether the system has provisioned an adequate amount of learning chances to achieve these outcomes, implementing a systematic way of gathering,

analyzing, and interpreting evidence to determine how well the students are learning. Their body language sometimes conveyed as if they felt someone was listening to them. Although we were in an empty room with the door shut, there was continuous staring at the door or looking at the PA system. The PA system is a public address electronic system comprising microphones, amplifiers, loudspeakers, and related equipment. Each classroom that I visited had one. It is commonly used to make announcements throughout the school building and speaking with a specific classroom without disturbing the entire school building. There is a beeping sound when it comes on, but there is no warning when it goes off. Because of this system, I felt that the interviewees would whisper at times or move in closer to whisper their true feelings about the assessment phenomena.

### **Data Interpretation**

In this interpretation of the data, I centered first on summarizing and synthesizing the insights gained from analyzing the information gained from the interviews and surveys discussed in Chapter Four and then highlighting the study's findings. The analysis process helped me identify how the teachers who have participated in the interviews, and those who responded statistically and narratively to the survey felt about the value and impact of ongoing standardized evaluations on their students and themselves.

Qualitative researchers recommend utilizing the triangulation of data that was obtained from multiple sources. This process is done to corroborate the information, cross-examine the identified patterns and trends, and strengthen the credibility of the findings discovered in the data (Efron & Ravid, 2020). Honorene (1970) states that the triangulation process helps researchers make a more profound sense of the data and its information. The triangulation of the different data sources enhances the research findings' strength, rigor, and validity.

Data gathered through in-depth participants' interviews showed that teachers favored doing away with standardized assessments. During the interviews, the teachers expressed a negative outlook toward standardized assessment's ability to provide guidance and feedback to promote student success. In their eyes, the information about students' achievements collected during these standardized tests does not help them reflect on or improve their instructional practices.

Having these in-depth discussions with various educators in different grade levels of teaching brought awareness of their frustration, Frustration for not having enough time to teach students the grade-level curriculum. A disappointment that administering repeating assessments causes their students to be anxious, frustrated for not being appreciated for their teaching craft, and angry for lack of support from their administration. For the teachers in this study, defeat involved feelings of discouragement and feelings of being thwarted from accomplishing specific critical educational responsibilities. They spoke of frustration in their daily work with students, parents, and administrators. They felt frustrated by the politics and policies of the school system.

I learned that many interviewees were frustrated with students not understanding how valuable standardized assessments are for their grade advancement. One of the interviewed teachers related that she was cajoling her students to their advantage. She was trying very hard to push students to complete all the high-stakes tests while imploring them to pay great attention to the words of the questions, instructions, and tasks they are being asked to accomplish so that no one would fail.

Much like the participants in Barnes' (1993) and Britt's (1997) studies, the teachers in this interview did not feel that their teacher-education programs had prepared them for the discrepancies between what they hoped for entering into the teaching profession and what was

possible to accomplish. Many teachers wanted to use different strategies to achieve various educational goals and apply various means of evaluating students' knowledge and assessing students' learning approaches. Standardized assessments do not allow for flexibility that would enable teachers to identify students' strengths and weaknesses promptly.

Through our conversations, I digest that the classroom educators felt unprepared for the "real" teaching world, populated with multiple, ongoing standardized tests as they grappled with their frustrations, disappointments, doubts concerning accountability assessment. The classroom teachers related their uncertainties about resolving issues brought forth because of these high-stake standardized tests. There was a general sense among the interviewed participants that they have sold a false bill of goods.'

Data collected through the interviews and online surveys revealed that some teachers felt frustrated about the time it takes to prepare the children for and administer the high-stake assessments. From their perspective, standardized assessments measure achievement against predetermined goals rather than measuring the students' progress. Teachers have also revealed how excessive pressure from their schools and administrators to improve their students' standardized assessment scores negatively affected them emotionally.

Though it would not be possible to generalize these findings to apply to the total teacher population in the district, due to the small size of the survey sample (17 teachers), this study does offer a glimpse into the way teachers across significant urban district regard, in general terms, the standardized assessments. While the participants, both in the interviews and surveys, seem to support the standardized assessment phenomena, they negatively perceive how the assessment system currently works. From the teachers' perspective, there was also a general feeling that there are excessive standardized assessments in the elementary school setting. The data analysis

also revealed some notable differences in specific aspects of the assessment phenomena that they deemed helpful in their practices. Moreover, in the survey, there was some indication that a particular standardized assessment, the Northwest Evaluation Association (NWEA), has been perceived as a positive measurement tool for delivering successful data to students' success. The study showed that 33 % of teachers found that the NWEA standardized assessment is most helpful, while 25% felt that the IAR standardized assessment was not at all helpful. The NWEA result data is available within 24 hours, but the IAR data is unavailable until the next school term.

Moreover, both the interviewees and survey participants rated their own formative and summative assessments very high over the standardized assessments. The survey revealed that 56% of teachers in this study find their formative assessment, which aims to monitor students' learning and provide ongoing feedback that instructors can use to improve their teaching and advance their students' Learning to be more effective. More specifically, in their eyes, formative assessments help students identify their strengths and weaknesses and target areas that need work. Additionally, the study showed that 48% of the teachers find their summative assessments to be the most helpful for their educational instruction. According to the teachers in the study, the summative assessment goal is to assess student knowledge at the end of a pedagogical unit by comparing it against the common core state standards. Summative assessments are often high in states, which means that they have a high point value.

The participants' interviews validated these ratings of assessments. When the interview participants were asked about their formative and summative assessment preferences regarding standardized tests, they insisted that the teacher-generated assessments were most helpful. The teachers found that their formative and summative assessments were clear and straightforward,

offering a series of examples and detailed steps that the students can easily follow. A new review of studies on this topic reinforces the teachers' position. It shows that a student's grade point average is usually not consistent with the same student's standardized test score (Lombardi, 2019).

The teacher-made assessments are usually criterion-referenced tests designed to assess student mastery of a specific body of knowledge (Wiggins, 1989). Disha M. of [yourarticlelibrary.com](http://yourarticlelibrary.com) noted that a teacher-made test is one of the most valuable instruments in the teacher's hands to solve his/her purpose. She explains that a teacher-made test is designed to solve the problem or requirements of the class for which it is prepared. Teachers appreciate that, for their students, those types of assessments have no ambiguities in the language as in standardized assessments. They explained that the teacher-generated tests' three areas of good quality evaluation could be measured within the educator's designed test by reliability, validity, and bias.

Reliability refers to the stability of the scores acquired. It indicates how consistent the scores are for each individual from one administration of an instrument to another (Efron & Ravid, 2013). Teacher-made tests measure a broad scope within the subject area; they cover a wide range of abilities and rely heavily on facts. Furthermore, there are no assessment biases in teacher-made tests that present one or more items on a test that offend or unfairly penalize students because of their characteristics, such as race, gender, socioeconomic status, or religion. The classroom teacher can connect their students to understand and relate with them before developing an assessment of their educational success.

Moreover, based on the teacher perspective, it became clear that these teacher-generated assessments have constructed and demonstrated validity by ensuring the assessments measure

the intended attribute and other extraneous attributes. Formative and summative evaluations also exhibit formative validity, which seeks to establish the extent to which a test can provide information that can help improve how a program functions. For example, in assessment for Learning, the aid is to collect information that will enhance how teaching is done for the learner's benefit (Knowles & Clark, 2008).

Furthermore, during the interviews, the teachers felt that students are spending too much time preparing for standardized assessments interrupting their educational curriculum's scope and learning sequence. They noted how there are ten weeks in the school quarter, and out of those ten weeks, three to four of them are spent on assessment preparation. Additionally, two weeks are spent on administering the standardized tests, which leaves only five weeks of learning grade-level content using methods and strategies to demonstrate their knowledge in various forms effectively.

According to a report by the Council of the Great City Schools, students devote 20 to 25 hours a year to captivating standardized tests, but it is unclear how much time they spend preparing for those tests (Waldman, 2020). An overwhelming majority of 34% of teachers within the survey believed that students spend too much time taking district- and state-mandated tests. Many teachers participating in that study attested that they would prefer to cut the frequency and length of state- and district-mandated tests rather than eliminate them.

Based on my own experience as a classroom teacher, I agree with the study's teachers' perspective. I, too, believe that teachers spend much time preparing students for school tests. As the study participants asserted, I also think that the right kinds of tests are those tests that focus on students' knowledge and skills and that they need to be well prepared for their futures. These kinds of tests, they argued, can influence teaching and learning in very positive ways. By their

nature, standardized high-stake tests are contrary to the Common Core State Standards, emphasizing the knowledge and skills required to be well prepared for education, work, and life after high school.

Observing children, assessing them, and teaching are inseparable processes. Observation information informs teachers what each child can do and what they are ready to learn next. During this research process, my classroom observations strengthen what I had been experiencing during my struggles with standardized assessments. It also left me with some homework that needed to be created before their subsequent encounter with standardized assessments.

I learned through these observations that the level of anxiety for all age levels of students during standardized assessments is exceptionally high. When observing the classrooms, many students asked to leave the room during the assessments on multiple occasions. The teacher had informed me that this behavior sometimes lingers into the classroom's formative and summative assessments. It is not until the students fully engage in the formative and summative assessments that they become at ease about engaging. I thought back to their classroom and realized that they, too, have seen these behaviors as well. I now understand where they are coming from. While reviewing their observation notes, I began to develop ways to address students' anxiety once their standardized assessments are over. The observations have given Ia compelling insight into the devastating damage these high-stake tests cause to the learners.

## **Findings**

The analysis and interpretation of the interviews, surveys, and classroom observations have led Ito the following findings:

1. The majority of the study educators lose at least three hours daily within a ten-week quarter to prepare for standardized assessments.

2. Teachers would like to see less focus on standardized assessments and more on instruction.
3. Using teacher-made formative and summative assessments is more valuable in evaluating students' educational success.
4. Teachers believe that states should remove the heavyweight focus placed on standardized assessments.

### **Reflections on the Findings**

Besides triangulating the different data sources to reach the research findings, I also triangulates these findings with the literature. Comparing and contrasting the various participants' perceptions and insights gained from classroom observations with the current scholarship discourse helped explain and validate this study's findings. Such triangulation allows for a deeper understanding of the complexity of standardized assessments. It shows how these research findings resonate with the body of current knowledge of the uses and abuses of high-stakes assessments in elementary school's educational experience.

Standardized tests are among the most dominant levers that elected bureaucrats, and other policymakers have for manipulating local schools and classrooms. A rising body of research proposes that standardized tests often change school and classroom practices (Corbett & Wilson, 1991; Herman & Golan, 1993; Madaus, 1988; Rottenberg and Smith, 1991). A standardized assessment is a method of building a consistent principle across states and grade levels (Education Reform, 2015). It requires test takers to answer the same question, and all answers are graded the same, predetermined way. Status and growth information in our assessments can help locate what is working and point to successful programs scaled up. Assessments that bring data that can be used in real-time may make decisions regarding educational policies and

provide real opportunities for teachers and school leaders (Dyer, 2016). Dyer also points out that standardized assessment tools can support teachers by differentiating instruction according to students' readiness and enhance their ability to set academic goals, evaluate programs, and target professional development.

Overall, this study's participants affirmed the importance of the "standardized assessment" phenomenon as an overwhelming process for students and teachers. Similar to the findings in the study by the Council of Great City Schools (2015), from their participant's perspectives, there was an understanding that for standardized assessments to be successful, classroom teachers must be flexible and capable of identifying students' withdrawn behavior. The study's researchers have found that the number of standardized tests that U.S. public school students take has blown up in the previous decade. Most of these schools require too many tests of dubious value.

On the other hand, the research findings show that the assessments are attached to many operations complexities and do not offer their students educational success (National Research Council, 2001). Moreover, all the study's participants agreed that having fewer assessments and providing sufficient time to review the students' high-stake test results would better benefit the students and curriculum. Archbald and Newman (1988) denote that the last decade has witnessed a growth or acknowledgment of the need for essential deviations in educational assessment practices. Findings also reveal that the teachers agreed that various standardized assessments do not measure the test results within the learning scope and sequence (Kliebard, 2004). Findings of a study conducted by the National Center for Education Evaluation Regional Center (2009) show consensus among the participants of not having enough school time to evaluate curricula and

intervention programs, inform instructional practice changes, and target professional development.

The above study also reveals that educators acknowledged the need to minimize the number of standardized assessments; from their perspective, this is a critical element in achieving a less pressured environment that hurts student educational success. This calling for reducing the frequency of high-stakes testing is also heard in the National Research Council (2001). The research findings demonstrate that teachers also utilized various data and information to decide their students' learning success. In many ways, they prefer the teacher-made tests because it aids teachers by reassuring them to reflect on the significance of what they teach and provides valuable results for improving instruction (McCann, 2018). Standardized testing has become prevalent that It is required by the Legislative, Executive, and Judicial branch of government, standardized assessments have become a central feature of American Public schooling (National Research Council, 1999).

Overall, this study revealed that standardized assessments are a very daunting task for educators, mainly due to the variety of operations and political issues that are consistently encountered. I propose that these findings be consistent with the critical theory with its roots in Max Horkheime (1937). The essential argument of critical theory is that all knowledge, even the most scientific or "commonsensical," is historical and broadly political by nature (Belland & Fee, 2012). However, these types of schools often have prosperous and obedient students and families. The test scores are frequently something these "successful" schools like to display to make themselves look superior to other challenging schools in the area. It makes the school appear viable.

As teachers, it is harsh to accept that our brilliant teaching is not the primary driver of our students' test scores. In reality, a far more significant predictor of students' success is socioeconomic status (Buboltz & Merrit, 2015) – and guess who is most likely to fail the high-stakes tests? Low-income students, of course. I ask, "Which students are overrepresented in the ranks of the poor?" The answer: Students of Color, Special Education Students, English Language Learners, to name a few. All of whom the public education system has historically underserved, a euphemism for terribly served. Michael Apple (2001) noted that students from working-class families and schools of color focus their curriculum on standardized test preparation. Students from working-class families and schools of color focus their curriculum on standardized knowledge measured by high-stakes tests. On the contrary, Apple says that more emphasis was put on systematic thinking, experimentation, and collaboration amongst white middle-class students. The most affluent families use their economic, social, and cultural capital to expose their kids to more experiences, attain credentials, and develop the literacies and competencies valued by the higher education market.

Color-coded sifting begins at an early stage when children should be learning through play, not multiple-choice bubbles. However, in the world of high-stakes testing, play and fun time have become a privilege denied to many Black and Brown children (Greenberg, 2015). Greenberg states that teachers of poor students and students of color are more likely to be rated ineffective and forced out because low test scores track poverty rates. As a result, their students' educations will be disrupted by a revolving door that deposits inexperienced teachers in their classrooms yearly.

### **Revisiting and Answering the Research Questions**

Both research questions have been deconstructed throughout the study. I started with an introduction of the assessment topic and its significance in Chapter 1. In Chapter 2, I presented a critical overview of the research and theories related to standardized educational assessment; then, I described the methodology of research in Chapter 3, analyzed the data. in Chapter 4, and synthesizing, interpreting, and presenting the findings in this chapter, Chapter 5. In the following section, I outline, based on the analysis, synthesis, and interpretation of the study data, the research-based answers to the two research questions that drove this dissertation study. My conclusions are presented from the educators' perspective who participated in the research and then examined within the broader educational discourse.

Elementary schools continue to strive to meet educational success nation/statewide, which is measured based on short-term metrics like attendance and standardized assessment scores. However, these schools should also consider long-term metrics such as graduation rates, student retention, drop-out rates, and school suspension. This holistic process would enhance logical decision-making capabilities by analyzing data and assessment measures gathered from multiple sources, not a multitude.

### **What are all of these standardized assessments needed for?**

All the study participants asserted that standardized assessments embrace many challenges for students and teachers throughout all elementary grades of K – 8. This data shows that all these assessments are certainly not needed from the teachers' perspective. Teachers find formative and summative assessments to be more valuable in supporting students' academic success.

The current concerns about appropriate test use represent only the latest round in a continuing discussion over the use of standardized assessments to develop educational goals

(Kliebard, 2004). They caution that if test scores are used to grant prizes or enforce consents, there are several problems: widening the gap in educational opportunities between haves and have-nots, narrowing the curriculum, centralizing educational decision-making, de-professionalizing teachers (Airasian, 1987; Haertel, 1989).

From the introduction of student evaluation in the mid-19<sup>th</sup> century, standardized assessments have served as instruments for accomplishing various policy purposes (Aronowitz, 1992; National Research Council, 1999). Among these purposes were defining the sorts of teaching students obtain, molding the content and arrangement of that instruction, and holding schools and students accountable for their performance.

The tension between policymakers' enthusiasm and experts' caution is symptomatic of two fundamental dilemmas posed by standardized tests when used as a strategy. First, policy and public expectations of testing generally exceed the tests themselves' technical quality and capacity (National Research Council, 1999). In his third version of *Ideology and Curriculum* (2004), Michael Apple argues that ideology is legitimated as schools' knowledge to support social norms. These social customs, in turn, produce students and teachers who are unable to challenge social, economic norms.

One of the most shared explanations for this gap is that politicians, under a constituent burden to recover schools, often decide to use existing tests for resolutions for which they were neither intended nor adequately validated (Bryk & Schneider, 2002). For example, tests designed to produce valid performance measures only at an aggregate level for schools or classrooms are used to report on and make decisions about individual students (National Research Council, 2014). In such cases, profound moments, such as retention in grades, may be unethically forced

on single students. That unfairness is further compounded if the skills being tested do not display or validly measure what students have been taught.

Policymakers sometimes acknowledge these problems and the need for more research. Nevertheless, they often choose to rely on an available test because they see only a fleeting opportunity for action or because they believe that, even with imperfect tests, more good than harm will be done. Given that funding is a crucial component of public schools, many rarely opt-out of standardized assessments.

### **How do educators perceive standardized assessments?**

Teachers have spoken of their frustration about the time it takes to prepare for and administer tests. Teachers feel excessive pressure from their schools and administrators to improve their standardized test scores. Standardized tests measure achievement against goals rather than measuring progress. The very fact that there will be standardized tests change how we are teaching the students new learning material. This study shows how educators believe that high-stakes tests lead to the narrowing of instruction and classroom curriculum. They feel so mainly when they know that tests influence decisions that cause significant changes in a student's life, such as advancement, teachers choose to teach to the test.

Teaching to the test means teaching students only the material they will be tested on and suppressing extra information not vital to high scores. The data obtained from the observation, interviews, and surveys implemented in the study revealed that teachers' pressure to improve their students' test scores resulted in some educators neglecting materials not included on the test. This means that students no longer experience long-term projects, reading physical chapter books, solving higher-order concerns, computer programs, etc. Teachers are teaching students by giving worksheets with questions formatted identically to standardized tests.

It is evident from this study's findings that teachers are becoming incredibly stressed by the emphasis placed on standardized tests. The regulation is known as the "No Child Left Behind Act" has created the importance of testing results even more pertinent in both students' and teachers' lives. Their careers, salaries, and livelihoods may be influenced by how well their pupils achieve on one assessment, one day of the year at a single moment. In some cases, student performance is connected to their teachers' salaries and educators' job stability, discovering which teachers will remain teaching and which ones will be placed on probation. These stress feelings echo the findings of decreased motivation, morale, and increased burnout (Saeki, Segool, Pendergast, & von der Embse, 2018).

Test scores, several teachers asserted, can be affected by a significant amount of variables. How sound students slept the night before, what types of breakfast they endured, how restless students are, the testing room temperature, etc., are out of the teachers' control. However, these factors influence how well students perform on standardized tests, determining the educators' fate.

I remember that during one of the interviews, one of the participants explained that some students would decisively do poorly on standardized tests because they did not care. From their perspective, the test did not affect them, so they did not even bother putting in any effort when taking the exams. However, teachers' careers and salaries are at stake. It is pretty prevalent that teachers' pressure to increase their students' test scores has led to the underside effect of diminishing their love for their profession. They became teachers hoping to improve the learning experience and cultivate the knowledge and development of elementary-aged students. They became disappointed when they realized that accountability based on standardized test scores had turned their classrooms into environments where test skills and memorization received more

attention than student discovery and developing a love of Learning (Franklin & Snow-Gerono, 2007).

Some teachers indicated they are questioning how much longer they can stay involved in public education. Standardized assessments have added significant stress to all teachers. They are looked at by administrators, and they feel judged by their students' scores. They think we push children way too hard now and expect so much. Some so many students feel stressed out from testing. It is heavy emotional pressure on them. I am not sure how much harder can we, as educators, be pushed. Most of the time, many teachers, novice and veteran, feel close to an emotional breaking point.

Although these teachers appear to be interested in "data-driven decision-making" because they appreciate the information that assessments may potentially provide, they are dissatisfied with the prevalent accountability craze in the educational environment. They are frustrated by the total emphasis on standardized test results and the fact that their job security is connected to standardized assessments that do not consider the student's holistic view.

All of the teachers believe that it is increasingly important to understand the influences of current accountability structures in elementary schools and classrooms on teaching and Learning. To fight against the negative consequences of high-stakes tests and the implementation of assessment and accountability structures in line with the new educational policy, teachers, administrators, and teacher educators must be attentive to the perceptual impacts at the classroom level. With this awareness, educators may partner together to create support structures and curriculum and instruction that meet all children's needs in classrooms. Likewise, classroom teachers and teacher educators may be able to demonstrate an unprecedented unity toward

policymakers by voicing appreciation for standards and accountability while at the same time demonstrating limitations of local implementation structures.

### **Validity and Trustworthiness of the Findings**

All researchers strive for trustworthiness in their findings. Trustworthiness supports the validity of the study's findings, regardless of the epistemological approach to the research. As Merriam (1998) wrote, "Every researcher wants to contribute results that are believable and trustworthy" (p. 218). Lincoln and Guba (1985) declared that trustworthiness asserts that the research findings are "worth paying attention to" (p. 290).

To ensure trustworthiness and rigor in qualitative research, Lincoln and Guba (2005) suggested four indicators that must be considered (a) credibility, which involves confidence in the findings; (b) dependability, which involves the consistency of the findings with the data collected; (c) conformability, which involves me acknowledging limitations, bias, and dilemmas encountered and the process entailed in transparency; and (d) transferability, which consists of the relevancy of the findings that can be transferred to other contexts by the reader of the research. Drawing from the criteria suggested by Lincoln and Guba, researchers such as Creswell (2012) and Marshall and Rossman (2016) proposed methods that strengthen qualitative studies' validity and trustworthiness. Five of these methods were used in this study: Multiple sources, triangulation, thick description, and a description of the research process and procedures.

### **Multiple Data Sources**

The validity was reinforced by using the following multiple data collection methods. The first was the use of face-to-face semi-structured interviews of three elementary classroom teachers deemed by the institutions' principal as the most knowledgeable of the phenomena in the elementary setting. The second method was the use of online surveys of the 17 teachers within the elementary setting ranging from kindergarten to 8<sup>th</sup> grade, which served as an external

means of consistent data retrieval. Finally, observations were obtained at the same elementary institution. Field notes were used as reflective commentary by I throughout the research process. The assurance is that each data collection method provided rich, thick data that supported the study's purpose and accuracy, thus the credibility of the findings.

### **Triangulation**

I triangulated different sources, compared and contrasted the different teachers' points of view, and compared what I heard from the teachers to what I saw in the classroom. The triangulation allowed for the converging of information from different sources to corroborate the findings and obtain credible answers to the research questions. Creswell (2009) contended, "in triangulation, researchers make use of multiple and different sources, methods, investigators, and theories to provide corroborating evidence" (p. 208), thus shedding light on new themes or perspectives. Merriam (2009), Miles and Huberman (1994), and Patton (1990) concurred that triangulation as a research method does substantiate themes, issues, and other new perspectives of the research and its findings.

Conducting semi-structured interviews, surveys, and classroom observations on the same research phenomenon allowed I to connect themes of similarities and differences within and between each data source. I used an anchor chart paper to draft themes found in each method of data collection. The next step was to circle and highlight themes of similarity. This allowed me to interrogate different understanding ways while creating a more in-depth picture of the research problem. Additionally, I triangulated findings with relevant literature to examine how an understanding of the impact of assessment on teachers who participated in the study reflects other researchers' findings and insights.

### **Thick Description**

It was vital for Ito provide a thick description. This detailed description provided a vivid illustration of the classroom I observed and the participants' point of view, and the rich story of their experiences. These data were all told from their perspective and their insights. Hammersly (2008) noted that a thick description is often used to characterize the distinctiveness of qualitative research. While interviewing, I captured the participant's thoughts and shift in body language while expressing their thoughts on standardized assessments. I was able to notate through quotes depicting what they had verbally expressed on specific questions. The open-ended survey question in the survey allowed educators to express themselves through their writing. This allowed me to correlate their opinions with those of the interview data and connect various themes. Having the opportunity to view classrooms allowed me to paint the picture of demographics and culture and describe the behavior and reactions of the children who took the high-stake tests and the teachers who supervised them.

When thinking about this research topic, it was essential that the teachers' voices came through very loudly, clearly, and accurately. I wanted to make sure that there were not no misconceptions. Member checking allowed Ito present narrative along with findings from this process to the participants.

The interviews and observations took place in early February, right before the world was hit with a medical pandemic. I had obtained the interviewees' and classroom observation teachers' emails. I sent a thank-you note to everyone following the interview and observation. Then the pandemic happened: all schools were ordered to close and begin remote learning by the United States Governor of Illinois. I was able to reach out and send the transcriptions and narrative to each participant through email. With pandemics plaguing our state, I was pretty stunned by their quick turnaround to responding, ensuring the accuracy of what had transpired.

That gave me more momentum to continue the research. During a remote learning professional development, the principal allowed me to discuss standardized assessments within the elementary school setting, and during this time, I released the data from the survey.

#### Study Limitations

In examining this study's limitations, it is essential to note that the data was gathered from a relatively small number of participants. The online survey polled 27 teachers, while the in-depth interviews consisted of 3 teachers accompanied by three classroom observations. The low rate of willingness to participate in the survey can likely be attributed to the circumstances surrounding the Covid-19 pandemic affecting all school districts during the time the study was conducted.

I initially wanted to conduct the study in their current school district, but the administration does not allow outside researchers to collect data in their district. Another point to contemplate is the possibility that because of their position as a classroom teacher, the respondents did not feel comfortable sharing their views about such a delicate and contentious subject. Sharing information often leads you to wonder where the information will end up. Although I ensured the interviewees that everything was anonymous and confidential, interviewees still became apprehensive at times. On the other hand, though the study's timing was unpredictable, it can also be surmised that individuals who, despite the circumstances, accepted the invitation to participate in the study considered the opportunity for their expression to be heard, and their opinions count essential and meaningful. This was certainly the impression I gathered from the interview participants' contributions.

Another limitation worth noting is that this study lies in the protocol used to conduct the interview. My original intention was to structure the discussion to have the flow of a natural conversation. For this reason, I had resolved to be the least intrusive as possible during the in-

depth interviews. Besides posing questions or clarifying what the teachers had asked, I did not intervene or participate in the discussion. The intent was not to influence the participants' opinions and let the conversation evolve organically within the group's members. However, during the conversations' transcription, it became apparent that there were many missed opportunities where it would have been appropriate for Ito ask the participants to clarify or expand on what they shared during the proceedings. Doing so would have unquestionably added to the discussion's depth and generated more insights and an enhanced understanding of the teachers' views and opinions. Despite these shortcomings, many valuable insights can be gained from it and be generalized to other settings. The results offer a compelling glimpse into the perception of excessive standardized assessments within the elementary school setting, which can be the basis for informed reflection and further exploration and research.

### **Practical Implications of the Study**

As an educator myself, I often think about the effects that standardized assessments have on students. Many of their colleagues that I have spoken with and the teachers I have interviewed and surveyed concerning standardized assessments feel that there are way too many standardized tests. This phenomenon of standardized assessments has gone too far. Standardized assessments will capture what is known about how students think and learn in that domain. Many educators are looking for ways to navigate through the overwhelming pressures of standardized testing. Although standardized assessments are mandated, it is essential that we, as educators, put our students at ease about the process and inform them that it is just a glimpse of who they are, but not lose sight of every teacher's assessment desire.

Additionally, as a critical theorist, I would recommend that educators increase public awareness of the harmful effects of the ongoing high-stakes tests on students' learning and

teaching quality. Teachers should also actively work together to ensure that authentic assessment is implemented and replace the high stakes standardized assessments. As I demonstrated throughout the dissertation, teachers recognize the value of being accountable and the potential positive contribution of assessment. What they reject is the way the evaluations of students are currently done.

### **Research Journey**

"Sit in the back; your score was low" that is how all this research began. My parents had separated for a year or two. I was thirteen years old when my mother uprooted my three younger siblings and me to a new state. A state demoted my siblings and me to a lower grade based on one test for their school entry. My mother went through a rough period with my dad, and she did not protest that rule, nor did she inquire about the situation. That humiliating demotion, based on one test, devastated me then and has left bitter memories. Here I am four decades later, still hearing in my head a teacher putting me on display at a brand-new school in front of the entire class telling me, "sit in the back; your score was low." I did not know then that I would graduate with an Associate's degree in Applied Science, Master of Arts in Teaching, Master of Arts in Reading, obtain endorsements in English as a Second Language (ESL), Social Science, and enter a doctoral program. Once in the doctoral program, I always wrote and spoke about standardized assessment injustices from day one. The hurt is still there. I was not the model student, nor was I an A student. I am a student who goes the extra mile, one who tries her very best. My father, bless his soul, once told me that I would go far in life if I continue to work hard. I only wish he was here to visualize me now address the pain I have been caring for decades by obtaining a doctoral degree. Although it has not been easy, I have given it my all.

When it came to the current journey of writing this dissertation, just when I thought that things are going smoothly with my research observation, surveys, and interviews, a medical pandemic erupts along with a particularly strong visibilities of racism against African Americans. I was already trying to balance my marriage, being a mother, taking care of my elderly mother, and holding a full-time job. I had just completed my interviews and observations for the dissertation study when the Corona pandemic occurred; I thought I would have more time for research when the pandemic was announced. This pandemic opened my eyes even more to the assessment challenges that educators face. In March of 2020, teachers were unable to administer the regularly scheduled assessments due to immediate remote learning for everyone. Teachers became frazzled about preparing a strong curriculum and, at the same time, were relieved of the standardized assessment restrictions. Not administering any year standardized assessments in 2020 allowed me to develop a relationship with my students and target areas of concern for them to be educationally successful. I didn't have to teach to a test nor worry about them taking a standardized assessment. We mainly focused on their current grade common core state standards. Starting a new school year, standardized assessments crept back into our schedules during the middle of the year. I observed during the 2020 year students becoming relaxed and not exhibiting any anxiety. Yes, I gave formative and summative assessments but they were tied into our everyday curriculum.

I come to realize during this pandemic that standardized assessments are not necessary to determine where a student places educationally; formative and summative assessments can guide you there. Teachers are using the common core state standards in developing these types of assessments. I would meet with my students in a google meet chat which allowed me to assess them individually on comprehension and fluency, which gave me more data points. My light

bulb moment was I didn't need those standardized assessments. Providing my classroom formative and summative assessments allowed me to obtain all of the data required to secure a student's success.

I took a deep dive into Maxine Greene's literature: *The Passions of Pluralism* and *Variations of the Blue Guitar*. This literature offered me knowledge on classroom teachers' practice, their questions, and their thinking about teaching and learning for aesthetic education. I leaned heavily on Darling-Hammonds literature *Standards, Accountability, and School Reform*. I found out that these strategies have often had unintended consequences that undermine access to education for low-achieving students rather than enhancing it. It has increased my knowledge of tests, coupled with rewards and sanctions, as the basis for accountability systems.

The fact is that when it has come to my situation, it has been the total opposite; I have had less time. Developing an online curriculum for fourth graders who are not tech-savvy has been a considerable challenge. My students had to be taught how to manipulate their computers for something other than TikTok or video games. However, when you are passionate about something, you make time.

Equally important to me is the Black Lives Matter movement. I know that it has been portrayed as something destructive, but it is not. The Black Lives Matter movement stands for freedom, liberation, and justice. There were so many shootings of police officers on African Americans with no accountability. I was scared; my son was nervous; my husband was frightened; black people were scared. The police brutality against blacks, which we all have experienced, was becoming a common phenomenon.

Being a woman of faith, I had to keep going, even though it seems as if the odds were against me. I made a poster for my bedroom with this scripture "For nothing will be impossible

with God" Luke 1:37. I genuinely believe that with all of the world's uncertainty, nothing is impossible with God.

### **Educators' Aspirations and Final Reflections**

One of the understandings I gained from the study is the possibility of re-examining the goals of education rooted in test-based accountability policies distinguished by its overemphasis on standardized testing with a more holistic approach to teaching and learning. This approach will be at the center of the discussion on educators' aspirations.

One of the things that became clear in the study is that I believe standardized assessments should focus on the students holistically rather than only on content knowledge. More and more educators, legislators, business leaders, and parents identify those educational objectives should reach far past the scope of old-fashioned subject-area domains. For instance, in the report *What Work Requires of Schools* (Secretary's Commission, 1991), the Department of Labor recognizes a broad array of both academic and non-academic competencies as necessary for the modern workplace, including:

1. creative thinking
2. decision making
3. problem-solving
4. learning how to learn
5. collaboration
6. self-management

This emphasis on goals outside the traditional content disciplines is also reflected in some widespread school reform approaches not necessarily associated with the America 2000 proposal

or its related efforts. For example, William Spady (1988), an advocate of outcome-based education, persuasively argues that educators must broaden their target goals to include lifelong learning outcomes. Given these commendations, several districts, and entire states, have focused their care on recognizing significant "exit outcomes" or "learner outcomes." For instance, the Aurora Public Schools in Aurora, Colorado, have noted five lifelong learning outcomes related to their proficiencies (Redding, 1991). They were becoming a self-directed learner, a collaborative worker, a complex thinker, a quality producer, and a community contributor. Several other states, including Maine, Michigan, Minnesota, Pennsylvania, and Virginia, have included similar lifelong learning outcomes among their revised educational goals. This emphasis on goals outside the traditional content disciplines is also reflected in some widespread school reform approaches not necessarily associated with the America 2000 proposal or its related efforts.

In my ambitious assessment world, students would only be subjected to teacher-created assessments such as summative and formative, project-based learning assessments, and one mandated standardized evaluation. This will give the educator three data points to analyze and develop a holistic view of who the student is and their potentials to become. Based on the data from the surveys and interviews, formative and summative assessments are highly favored. When united into classroom practice, it provides the information needed to adjust teaching and learning while they are happening. Formative assessments help teachers determine the following steps during the learning process as the instruction approaches the summative assessment of student learning. Formative assessments are quizzes and tests that assess how individual students are learning components during a course. Summative assessments are quizzes and tests that gauge how much someone has learned throughout a course. Project-based assessments are an

alternative to tests that allow students to engage with their learning more concretely. Instead of merely studying theory, a hands-on project asks students to apply what they have learned to an in-depth exploration of a topic. The below figure provides a chart developed by Magnify Learning, demonstrating the process of creating a project-based culture.

Magnify Learning is a teacher-driven non-profit organization dedicated to expanding high-quality Project Based Learning (PBL). Project-Based Learning offers professional development in customized open workshops, on-site workshops, administrative workshops, and after-school workshops. Their emphasis is on building relational culture and authenticity, an approach that I agree with. Along with formative, summative, and Project-based assessments, I believe it is crucial to have one mandated standardized assessment. Data from the interviews and surveys show that the Northwest Evaluation Association (NWEA) is the most respected standardized assessment. I could not agree more with the outcome. The data analysis from this assessment is ready within 48 hours, which allows a teacher to make immediate adjustments to students' academic success.

**Figure 10**

Project-Based Learning Culture



In 2015, President Obama directed the Department of Education to review its policies and address any states where the Federal Administration may have contributed to the problem of overemphasis on testing and caused burdening classroom time. The Obama Administration provided financial support for states to develop and use better, less burdensome assessment. Obama declared that "moderation, smart, strategic tests can help us measure our kids' progress in school, and it can help them learn" (Richmond, 2015). However, in reality, multiple actors require urging and encouraging various tests for very different reasons that do not necessarily add to a clear picture of how students are doing. The result is an assessment system not very intelligent and not coherent (Layton,2015).

This study revealed that from the teachers' perspective, many of the tests we do use could not measure what should be the seal of a "thinking" curriculum: the refinement of students' ability to administer competence and knowledge to real-world problems. The testing

process may interfere with the kind of higher-order skills that are desired (Resnick 1987, p. 47). For the most part, standardized tests want students to remember or become familiar with fragmented and isolated bits of information. They rarely ask students to apply that information, and they rarely depend upon students to display proficiencies in the "higher forms" of cognition, such as complex reasoning and self-directedness (Marzano & Costa, 1988).

These shortcomings make it clear that if we are to reasonably assess students' aptitude to meet the enduring learning standards and challenging content criteria and consequences, that is the strategy's focus to make the American public education system the best in the world, new approaches to assessment are needed. While there is not enough support for states and school districts to reduce the time spent on testing, there are some glimpses of hope. There may be some signs for new trends of more flexibility from federal mandates who offered more significant support to innovate and minimize testing and reduce the reliance on student test scores through rules and executive actions. Some states have begun to make that change; states like New York and North Carolina limit the number of time students spent on tests. Rhode Island eliminated standardized tests that did not advance teaching and learning and ensured that the administered tests inform teachers' instruction (Dept. of Education, 2015). According to the Department of Education report, the Delaware governor launched in Jan 2021 a review of all tests managed by the state, districts, and individual schools to decrease the testing burden on students and teachers and increase teaching time.

A unique "ah-ha" moment in the study occurred when the interviewees and I discussed the ramifications of standardized assessments in the elementary school setting and our ability to continue enriching and engaging students' education. Our conversations reminded me of the

great Maxine Greene, an American educational philosopher, author, social activist, and educator who began developing a path for assessment success. Her ideas and lectures on education aim to educate and reform teachers and inspire them to alter standardized education's systematic formula. Maxine Greene (1977) contributed a significant component to reforming the current assessment situation by discussing three themes: freedom, imagination, and aesthetic education.

To Greene, freedom is defined as breaking from "automatization and the bland routinization of life" (Greene, 1977, 36-48). To accomplish this personal and individual freedom, we must first know something beyond banality. This is where imagination comes in. Greene talks about what she calls "social imagination," which uses imagination to picture the world as better than it is and imagines the steps that we can take to change it (Greene, 1993). Through these themes, Greene has inspired me, like so many other educators, to facilitate change. She was influential in movements like civil rights, women's liberation, and education reform.

In an era of scripted teaching, yearly standardized tests, and relentless attention to data collection, Dr. Greene would seem to be a relic of little interest. Dr. Greene is a symbol of teaching as lost romance. She simply states that "Teachers can get bored and boring if they forget that they are fundamentally intellectuals working with ideas, not just testing kids to memorize facts"(Arenson, 2001). She inspires teachers and gives them permission not to get loaded down with the humdrum of bureaucracy and standardization. She keeps people teaching longer.

She charges teachers to become aware and realize that we are in a very inhospitable environment. Be aware that imagination is given a very small place in the way people talk about learning today, but it is a way to go beyond, to break through boundaries.

Whereas standardized assessments often initiated through politics are generally designed to survey policy compliance or student achievement across the country or

internationally. Most commonly, standardized tests are used to measure knowledge in specific areas such as math, reading, and social studies. These tests are a method of assessment built on the principle of consistency: all test takers are required to answer the same questions, and all answers are graded in the same, predetermined way. Maxine doesn't believe that constant testing will achieve the standards that many political figures have designed. She dismisses the president's education plan with one of her signature questions: "Is he talking about education or is he talking about an Army camp?"(Arenson, 2001).

Maxine Greene has already designed the need to move forward. Iresonates with her philosophy. As educators, parents, administrators, and even policymakers, we need to realize that students are not robots; they are people who merit the right to explore the world and not have the heavyweight of a standardized assessment on our shoulders to measure our worth. After listening to this study's interviewees, observing the classroom of students during assessments, and analyzing the survey data on standardized assessments brought to light the inequalities of standardized assessments. This clear insight inspires me to embrace Maxine Green's vision and inspire other educators to actively fight against what Greene called "the plagues" of indifference, ignorance, and apathy.

*Sometimes, the most brilliant and intelligent minds do not shine in standardized tests because they do not have standardized minds."*

- Diane Ravitch



## REFERENCES

- Triangulation of Data Sources and Research Methods*. (2018). National Academies of Sciences, Engineering, and Medicine. Changing sociocultural dynamics and implications for national security: Proceedings for a Workshop, Washington, DC
- Aiken LS, West SG. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, California: Sage Publications
- Alcozer, P. (2001). History of standardized testing in the United States. National Education Association. Retrieved from <http://www.nea.org/home/66139.htm>
- Angrosino, M. (2007). *Doing ethnographic and observational research*. SAGE Publications Ltd. <https://www.doi.org/10.4135/9781849208932>
- Apple, M. (2004). *Educating the right way: Markets, standards, god and inequality*. New York: Routledge Falmer. <http://louisville.edu/journal/workplace/issue5p2/applereview.html>
- Apple, M. W. (2004). *Ideology and curriculum*. New York: Routledge
- Arco, Matt (2015). Christie education speech in Iowa, NJ.com,
- Archbald, D., & Newmann, F. (1988). *Beyond standardized testing: assessing authentic academic achievement in the secondary school*. Washington, DC: Office of Educational Research and Improvement.
- Austin, Z., & Sutton, J. (2014). Qualitative research: getting started. *The Canadian Journal of Hospital Pharmacy*, 67(6), 436–440. <https://doi.org/10.4212/cjhp.v67i6.1406>
- Ayers, W., (1993). *To teach: the journey of a teacher*. New York: Teachers College Press.
- Banta, T.W., & Palomba, C.A. (2015). *Assessment essentials: Planning, implementing and improving assessment in higher education*. San Francisco, CA: John Wiley & Sons, Inc.

- Barton, P. E. (2002). *Staying on course in education reform*. Princeton, NJ: Statistics & Research Division, Policy Information Center, Educational Testing Service.
- Barnes, S. (1993). A model for implementing outcome based education. *The Center Review*, in press.
- Beck, C. L., & Gargiulo, R. M. (1983). Burnout in teachers of retarded and nonretarded children. *The Journal of Educational Research*, 76(3), 169-173.  
<https://doi.org/10.1080/00220671.1983.10885444>
- Belland, B. & Fee, S. (2012). The role of criticism in understanding problem solving.  
[https://doi.org/10.1007/978-1-4614-3540-2\\_16](https://doi.org/10.1007/978-1-4614-3540-2_16).
- Berger, P. L., & Luckmann, T. (1991). *The social construction of reality: A treatise in the sociology of knowledge*. Penguin Uk.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 8–21.
- Black, P & Williams, D. (1998) Assessment and classroom learning, assessment in education: principles, *Policy & Practice*, 5(1), 74, <https://doi.org/101080/0969595980050102>
- Bloom, B.S. (1956). Taxonomy of educational objectives: The classification of educational goals: *Handbook I*, Cognitive Domain. New York-
- Bobbitt, F. (1956) Social efficiency movement: Bobbitt’s contribution.  
<http://education.stateuniversity.com/pages1794/Bobbitt-Franklin-18761956.html>
- Bogdan, R., & Biklen, S. K. (2003). *Qualitative research for education: An introduction to theory and methods*. Boston: Allyn and Bacon.
- Boser, U. (2019). Teaching to the Test? - Education Week.  
<https://www.edweek.org/ew/articles/2000/06/07/39alignh19.html>

- Boss, S. (2011). Comprehension assessment: a short story – Edutopia. Retrieved from <http://www.edutopia.org/comprehensive-assessment-history>
- Britt, A. (1997). Perceptions of beginning teachers: Novice teachers reflect upon their beginning experiences. (ERIC Document Reproduction Service No. ED415218).
- Britannica, T. Editors of Encyclopedia (2020). *Chinese civil service*. *Encyclopedia Britannica*. Retrieved from <https://www.britannica.com/topic/Chinese-civil-service>
- Brookhart, S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education*, 10, 161–180.
- Burns, R. C., & Drake, S. M., (2004). Meeting standards through integrated curriculum. Alexandria, Va: Association for Supervision and Curriculum Development.
- Carter, P., & Welner, K. (2013). Closing the opportunity gap: What America must do to give every child an even chance. Oxford University Press. <https://doi.org/82981.001.0001/acprof-9780199982981>
- Check, J. W., & Schutt, R. K. (2012). *Research methods in education*. Thousand Oaks, Calif: Sage Publications.
- Cherry K. (2019). How Lewis Terman influenced the field of psychology. *Very Well*. Retrieved from <https://www.verywellmind.com/lewis-terman-biography-2795523>
- Church, M. A., Elliot, A. J. and Gable, S. L. 2001. Perceptions of classroom environment, achievement goals, and achievement outcomes. *Journal of Educational Psychology*, 93: 43–54
- Coleman, J. S. (1987). “Norms as social capital.” In *Economic imperialism: The economic approach applied outside the field of economics*, [Eds.]

- G. Radnitzky, & P. Bernholz. New York: Paragon House Publishers
- Collins, R. (2013). Authentic assessment: Assessment for learning. *Curriculum and Leadership Journal* 7(11). <http://www.curriculum.edu.au>
- Concepts, L. (2015). The glossary of education reform. The Glossary of Education Reform. Retrieved from <http://edglossary.org/>
- Cope, S. (2016), Teaching Creative Nonfiction: The transformative nature of the workshop method. *Teaching and Learning*, 67- 73. <https://doi.org/10.1002/tl.20200>.
- Corbett, H. D. & Wilson, B. L. (1991). Two state minimum competency testing programs and their effects on curriculum and instruction. In R. E. Stake (Ed.), *Advances in program evaluation: Vol. 1. Effects of mandated assessment on teaching* (pp. 7-40). Greenwich, CT: JAI Press Ltd.
- Corbin, J., & Strauss, A. (2015). *Basics of Qualitative Research*. Thousand Oaks, CA: Sage.
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Sage Publications, Inc.  
<https://doi.org/10.4135/9781452230153>
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, Calif: Sage Publications
- Creswell, J. W. (2007). *Qualitative inquiry & research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.
- Creswell, J.W. (2009) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Los Angeles: Sage
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson

- Creswell, J. W. (2013). *Qualitative Inquiry & Research Design: Choosing among Five Approaches* (3rd ed.). Thousand Oaks, CA: SAGE.
- Crossman, J., (2007). The role of relationships and emotions in student perceptions of learning and assessment, *Higher Education Research & Development*, 26:3, 313-327, <https://doi.org/10.1080/07294360701494328>
- Crossman, Ashley. (2020). *Understanding Critical Theory*. Retrieved from <https://www.thoughtco.com/critical-theory-3026623>
- Curtis, B., & Curtis, C. (2011). *Social research*. SAGE Publications, Inc. Retrieved from <https://www.doi.org/10.4135/9781526435415>
- Darling-Hammond, L. (2004). Standards, Accountability, and School Reform. *Teachers College Record - TEACH COLL REC*. 106. 1047-1085. <https://doi.org/10.1111/j.1467-9620.2004.00372>.
- Darling-Hammond, L. (2011). Wanted: a national teacher supply policy for education: The right way to meet the “highly qualified teacher” challenge. *Education Policy Analysis Archives*, 11(33), 1–55. <https://doi.org/10.14507/epaa.v11n33.2003>
- Davidson, C. (2011). *Now you see it: How the Brain Science of Attention will transform the way we live, work, and learn*. New York: Viking Retrieved from <http://cps.edu/SchoolData/Pages/Assessment.aspx>
- Definition.net (2019). *Educational assessment*. Retrieved from <https://www.definitions.net/definition/educational+assessment>.
- Dillion, S. (2007). Battle grows over renewing landmark education law. *The New York Times*, p. 25a.

- Donnor, J.K., & Shockley, K.G. (2010). Leaving us behind: a political economic interpretation on nclb and the miseducation of African American males. *Educational Foundations*, 24(3-4), 43-54
- Dworkin, A. G., Sasha, L., Hill, A. (1987). Teacher burnout in the public schools. *International Education Journal* 2, 69-77.
- Earl, L. M. (Lorna Maxine), 1948-. Assessment as learning: Using classroom assessment to maximize student learning. Thousand Oaks, Calif.: Corwin Press
- EdChoice (2019). The abc's of school choice: The comprehensive guide to every school. Boulder and Tempe: Education and the Public Interest Center & Education Policy Research Unit. <http://epicpolicy.org/thinktank/review-abcs-school-choice>
- Editorial Projects in Education Research Center. (2004, August 3). Issues A-Z accountability, education week. <http://www.edweek.org/ew/issues/accountability>
- Education Week (2017). Quality counts 2017: Reports and Rankings. Retrieved from <https://secure.edweek.org/media/quality-counts-2017-news-release.pdf>
- Education Government (1983). A nation at risk report. Retrieved from <https://www2.ed.gov/pubs/NatAtRisk/risk.html>
- Education Reform (2013) Achievement Gap Definition Retrieved from <https://www.edglossary.org/achievement-gap/#>.
- EduTopia (2008). Why is assessment important? Retrieved from <https://www.edutopia.org/assessment-guide-importance>
- Efron, S.E. & Ravid, R. (2013). Action Research in Education: A Practical Guide. New York, NY: Guilford Press. 268 pages. The Canadian Journal of Program Evaluation.

- Efron, S.E., & Ravid, R. (2019). *Writing the literature review: A Practical Guide*. New York, NY: The Guilford Press.
- Efron, S. E., & Ravid, R. (2020). *Action research in education: A practical guide* (2nd ed.). The Guilford Press.
- Eggers, J.P. & Kaul, A.,(2018): Motivation and ability? A behavioral perspective on the Pursuit of Radical Invention in Multi-Technology Incumbents. *AMJ*, 61, 67-93,  
<https://doi.org/10.5465/amj.2015.1123>
- Fairtest, (2012). How standardized testing damages education. The National Center for Fair and Open Testing. <https://fairtest.org/how-standardized-testing-damages-education-pdf>
- Fairtest, (2017). The limits of standardized tests for diagnosing and assisting student learning. *The National Center for Fair and Open Testing*. <https://fairtest.org/limits-standardized-tests-diagnosing-and-assisting>
- Farber, B. A., Miller, C. (1981). *Crisis in education: Stress and burnout in the American teacher*. San Francisco, CA: Jossey-Bass.
- Fletcher, D. (2009). Standardized testing. *TIME Magazine*.  
<http://content.time.com/time/nation/article/0,8599,1947019,00.html>
- Flick, U. (1998). *An introduction to qualitative research*. London: Sage Publications.
- Fosnot, C., (1996). Constructivism: A psychological theory of learning. In C. Fosnot (Ed.) *Constructivism: Theory, perspectives, and practice*, (pp. 8-33). New York: Teachers College Press.
- Franklin, C.A., & Snow-Gerono, J.L. (2007). Perceptions of teaching in an environment of standardized testing: Voices from the Field.
- Friesen, D., Prokop, C. M., & Sarros, J. C. (1988). Why teachers burn out. *Educational*

*Research Quarterly*, 12(3), 9–19

Friesen D., Sarros, J.C., (1989). Sources of burnout among educators. *Journal of Organizational Behavior* 10: 179-188

Fryer, R. & Leavitt, S. (2004). Falling behind: New evidence on the black-white achievement gap. *Education Next*. V4, N.4, pages 64-71

Gareis, C. R., & Grant, L. W. (2017). *Teacher-made assessments: How to connect curriculum, instruction, and student learning*. Larchmont, NY: Eye on Education.

Ghaicha, A., (2016). Theoretical framework for educational assessment: A Synoptic Review. *Journal of Education and Practice* www.iiste.org ISSN 2222-1735, Vol.7

Gold, Y. (1984). Burnout: A major problem for the teaching profession. *Education*, 104, 271-274.

Greene, M. (1977) Imagination and aesthetic literacy, *Art Education*, 30(6), 14-20, <https://doi.org/10.1080/00043125.1977.11651958>

Greene, M. (1993). The passions of pluralism: Multiculturalism and the expanding Community. *Educational Researcher*, 22(1), 13-18. <https://doi.org/10.2307/1177301>

Greene, M.(2001). *Variations on a blue guitar: The Lincoln Center Institute lectures on aesthetic education*. New York: Teachers College Press, Columbia University

## APPENDIX A

### Interview/Survey Questions

1. Which of the following ways do you assess students quarterly in your class?  
(check all that applies)

NWEA       Dibles/TRC    Teacher Created Summative       Teacher Created Formative

2. How much of your instructional time during a quarter do you use for mandated standardized assessments?  
(choose the one that applies)

More than a month    More than 2 weeks    More than a week

Less than a day       8 hours a day       24 hours a day

3. In your experience how aligned are the standardized assessments to CCSS (common core state standards).  
(choose the one that applies)

90% and above    75% – 89%       60% – 74%    73% and below

4. How do you use Dibles/TRC data in your classroom? (Check all that apply)

Small group instruction    Instructional Planning

Individual Learning Planning    Instructional Curriculum Support

Don't Use Data

5. What type of assessments benefit your classroom in Reading?  
(check all that apply)

TRC/Dibles

Teacher Created Formative Classroom Assessment

Teacher Created Summative Classroom Assessment

NWEA

Not Applicable

## Appendix B

### Informed Consent Observation Interview

My name is Cathy Wright, and I am a Doctoral student at National Louis University. I am asking you to participate in this study, "Are All These Assessments Necessary", occurring from 02-2020 to 04-2020. The purpose of this phenomenological study is to delve into and understand the effects of standardized assessments on elementary students. My goal is to find out how do standardized assessments impact students' educational, emotional, and social growth. This form outlines the purpose of the study and provides a description of your involvement and rights as a participant.

Your participation in this study is strictly voluntary and you can withdraw from it at any time with no penalty to you. The study does not have any known potential risks. You are invited to participate in two individual semi- structural qualitative interview that may last approximately 45 to 60 minutes long at a location of your choice. You will be able to respond to a series of questions pertaining to your educational background and the path to the position of leadership. The interview will be audio (pending your consent) and transcribed. The transcription of the interview will be presented to you for verification and accuracy.

There are not anticipated risks or benefits, no greater than that in daily life. Further, the information gained from this study could be useful to other schools and school districts looking to reduce their standardized assessments.

Upon request you may receive a summary results from this study and copies of any publications that may occur. Please email the researcher, \_\_\_\_\_ to request results from this study. In the event that you have questions or require any additional information, please contact the researcher at \_\_\_\_\_. To ensure confidentiality the researcher will secure recordings, transcripts, and field notes in a locked cabinet in her home office. Only Cathy Wright, the researcher will have access to data.

If you have any concerns or questions or during participation that has not been addressed by the researcher, you may contact Dr. Efron at \_\_\_\_\_, the co chairs of NLU's International Research Board: Dr. Shaunti Knauth; email: \_\_\_\_\_; phone \_\_\_\_\_ or Dr. Kathleen Cornett; email: \_\_\_\_\_

---

phone: . Co-chairs are located at National Louis  
University, 122 South Michigan Avenue, Chicago. IL.  
Thank you for your consideration.

Consent: I understand that by signing below, I am agreeing to participate in the study "Are All These Assessments Necessary". My participation will consist of the activities below during the Feb. – Apr. 2020:

- 2 interviews lasting 45 – 60 minutes each
- 1 60 minute observation of my classroom

\_\_\_\_\_  
Participants Signature

\_\_\_\_\_  
Date

\_\_\_\_\_  
Researcher's Signature

\_\_\_\_\_  
Date

You will be tendered a copy of your signed form. Please acknowledge with your signatures below your consent to participate in this study.

## APPENDIX C



**NATIONAL  
LOUIS  
UNIVERSITY**  
ACCESS. INNOVATION. EXCELLENCE.

Office of the Provost  
122 South Michigan Avenue  
Chicago, Illinois 60603-6162

www.nlu.edu  
P/F 312.261.3729

February 7, 2020

Cathy Wright

Dear Cathy Wright:

The Institutional Review Board (IRB) has received your application for your research study "Are All These Assessments Necessary?" IRB has noted that your application is complete and that your study has been approved by your primary advisor and an IRB representative. Your application has been filed as Expedited in the Office of the Provost.

IRB: ER00724

Please note that the approval for your study is for one year, from February 7, 2020 to February 7, 2021. As you carry out your research, you must report any adverse events or reactions to the IRB.

At the end of your approved year, please inform the IRB in writing of the status of the study (i.e. complete, continuing). During this time, if your study changes in ways that impact human participants differently or more significantly than indicated in the current application, please submit a Change of Research Study form to the IRB, which may be found on NLU's IRB website.

All good wishes for the successful completion of your research.

Sincerely,

**Shaunti Knauth, Ph.D.**  
**Chair, IRB**