



eCOMMONS

Loyola University Chicago  
Loyola eCommons

---

Master's Theses

Theses and Dissertations

---

2014

## High-Stakes Testing: The Student Voice

Julianna Marie Cechowski  
*Loyola University Chicago*

Follow this and additional works at: [https://ecommons.luc.edu/luc\\_theses](https://ecommons.luc.edu/luc_theses)



Part of the [Education Commons](#)

---

### Recommended Citation

Cechowski, Julianna Marie, "High-Stakes Testing: The Student Voice" (2014). *Master's Theses*. 2617.  
[https://ecommons.luc.edu/luc\\_theses/2617](https://ecommons.luc.edu/luc_theses/2617)

This Thesis is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Master's Theses by an authorized administrator of Loyola eCommons. For more information, please contact [ecommons@luc.edu](mailto:ecommons@luc.edu).



This work is licensed under a [Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 License](#).  
Copyright © 2014 Julianna Marie Cechowski

LOYOLA UNIVERSITY CHICAGO

HIGH-STAKES TESTING: THE STUDENT VOICE

A THESIS SUBMITTED TO  
THE FACULTY OF THE GRADUATE SCHOOL  
IN CANDIDACY FOR THE DEGREE OF  
MASTER OF ARTS

PROGRAM IN CULTURAL AND EDUCATIONAL POLICY STUDIES

BY

JULIANNA CECHOWSKI

CHICAGO, ILLINOIS

AUGUST 2014

Copyright by Julianna Cechowski, 2014  
All rights reserved

## TABLE OF CONTENTS

ABSTRACT	v
CHAPTER ONE: HISTORY OF TESTING	1
Introduction	1
Schools Before Testing	2
Tests Will Measure....	2
Shift From Traditional to Testing	4
Test Bias	5
Statistics	7
Conclusion	9
CHAPTER TWO: THE CLEARING HOUSE: 1950S & 1960S	10
Introduction	10
The Start of Testing	10
How to Administer a Standardized Test	13
Validity of Tests	14
Effects of Testing	16
Yield the Best Results	18
Testing Data	21
New Pressures	22
Testing Advocates	23
Testing Take-Over	25
What Should Assessments Looks Like?	26
Conclusion	27
CHAPTER THREE: PURPOSE OF THE STUDY	29
Introduction	29
Significance of the Study	30
Variable and Data Sources	31
Potential Limitations	31
Research Question	32
CHAPTER FOUR: ACCOUNTABILITY	33
Strengths	33
Data Manipulation	34
Retention and Graduation	36
Curriculum and Student Focus	39
Other Issues	41
How do Students Feel?	44
CHAPTER FIVE: I LIKE TO READ, BUT I'M NOT GOOD AT IT	46

CHAPTER SIX: METHODS	48
Participants	48
Instruments	49
Procedure	50
Analysis	52
CHAPTER SEVEN: RESULTS	53
Introduction	53
Meet the Interviewees	54
Interviews	55
Negatives in Testing	56
Feeling Pressured	57
Nervous and Anxious	59
Scared and Stressed	60
I Like High-Stakes Tests	63
Attitude Towards Testing vs. Performance	64
School and Curriculum	67
Teacher Performance	68
Promoting Testing	71
Online Test Prep	74
Test Results	76
Accurate Assessments	78
Conclusion	81
CHAPTER EIGHT: CONCLUSION	83
REFERENCE LIST	86
VITA	89

## ABSTRACT

There has been an intense debate about standardized testing since they were first introduced into public schools in the nineteenth century. In this research, it is important to know why standardized tests were created and how they have been debated. Roy C. Owens said, “We cannot know where we’re going if we don’t know where we’ve come from.” Past information is important in this research because it will show how far we have progressed since the start of the standardized testing movement.

What started out as standardized testing later changed into high-stakes testing. The high-stakes were created by tying student test scores to promotions within Chicago Public Schools (CPS). If a student did not meet a predetermined test score, they attended summer school to obtain the successful score. If the student did not meet the score during summer school, the student was required to repeat the grade level.

Through the debates of high-stakes testing, the students’ perspective was repeatedly over looked. The goal of this research is to determine how students feel about high-stakes testing and if they feel that the tests have changed their feelings toward their school, themselves, or the learning experience. This research is important because it will help U.S. policy makers see if standardized tests are an accurate portrayal of student performance.

CHAPTER ONE  
HISTORY OF TESTING

**Introduction**

Examining the history of standardized testing allows for many perspectives and interpretations. This text will examine: Robert L. Williams “The War Against Testing: A Current Status Report”, Nicholas Lemann’s *The Big Test: The Secret History of the American Meritocracy*, Ryan, Ann Marie Ryan’s “From child study to efficiency: district administrators and the use of testing in the Chicago public schools, 1899 to 1928”, and William Reese’s *Testing Wars in the Public Schools: A Forgotten History*. A deeper understanding will emphasize the importance of the different perspectives which could be influenced by personal bias or the time period the piece was written.

Looking into the history of any topic is important because it helps to understand how we got to where we are today. Reforms in any capacity usually have those who agree with the reform and those who criticize it. IQ testing which later turned into standardized testing (with or without high-stakes) is no different. In this analysis of four secondary sources from three different eras, the sources are consistent in discussing the following: schools before testing, what testing was created to measure, the shift from traditional run schools to schools influenced by reforms, testing bias, and statistics that produce quantitative data.

### **Schools Before Testing**

Lemann discussed the struggle to establish elementary schools in the nineteenth century and the continued struggle to establish high schools in the nineteenth century (Lemann, 1999, p. 8). In the 1800's students were not expected to compete in standardized written exams that were timed. Instead he pointed out that students performed in recitals and exhibitions. These exhibitions were the assessments for the student and the school (Reese, 2013, p. 158). Other community members were also invited to view the exhibitions and assessed the quality of the teachers and schools (Reese, 2013, p. 14). This type of display allowed for parents and community members to enjoy the children's performance while seeing what they learned in school. However, Reese also discussed that oral exams failed to document beyond impressions (Reese, 2013, p. 137). This is an interesting fact because none of the other authors analyzed in this text brought up this aspect of early educational assessments.

Both Lemann and Reese discuss the early years of education. The creation of schools and the school's first practices were important because it showed the timeframe of when testing was implemented in schools (both K-12 and higher education). In the beginning, education of children was more for the community and less for the state. In later years, tests were created to help improve educational performance and results were used by the states to determine success.

### **Tests Will Measure....**

Throughout *The Big Test*, Lemann discussed the founders of testing, James Bryant Conant, Henry Chauncey, and Carl Brigham. However, Lemann only shortly discussed Alfred Binet who created the first IQ test in 1905, and E.F. Lindquist who



established the Iowa Academic Meet in 1929. Binet was interested in understanding the mental age of the student and providing help to slower learners. IQ test promoters, Lewis Terman and Edward Thorndike wanted to have widespread use of IQ tests to sort students according to their capacity (Lemann, 1999, pp. 17-18). In 1936, Lindquist did not want the tests to select gifted students or change curriculum because they only benefited a particular type of student, and were damaging to all other students (Lemann, 1999, p. 25).

Lemann discussed Binet and Lindquist's perspective about identifying the lower performing students, but did not touch on their perspective of ability grouping. Ryan discussed the use of "classification" testing in 1925 which was used to place high school juniors in ability groups (Ryan, 2011, p. 352). Similarly, Williams discussed the use of ability testing and how they falsely labeled students during 1969. The mislabeling of Black children dehumanized them and created an intellectual genocide (Williams, 1980, p. 266). Did Ryan and Williams have a different perspective of ability grouping because they were further away than Chauncey?

Conant, Chauncey, and Brigham believed they were trying to use the tests to measure people's abilities which included developed ability and academic ability. After Brigham recanted his earlier views on testing, in 1929 he stated, "I feel we should all stop naming tests and saying what they measure" (Lemann, 1999, p. 33). Later in 1950, Conant started to distance himself from the testing in which he helped to create because he did not approve of what developed (Lemann, 1999, pp. 78-79). Even though both Brigham and Conant saw the mistakes in testing, Chauncey loved everything about the tests and even enjoyed taking them himself (Lemann, 1999, p. 81).

Did tests label students and create ability grouping? Williams discussed T.E. Newland's perspective in 1970; Newland believed that ability tests did not label students but the adults who used the tests labeled the students. Newland's defense was that it was the people in charge who decided what tests were used and how they used the data collected from the tests. The test publishers were not to blame if a child felt like a failure because of the result of their product but instead the adults administering them in the school (Williams, 1980, pp. 268-269).

Binet, Brigham, and Lindquist saw standardized testing as a way to improve education. Utilizing the test results helped identify slow learners; the tests were not intended to identify students as gifted or change the curriculum. The use of testing to track students and label them into different groups was an unforeseen circumstance of testing. Although, some of the testing founders realized the problems later, testing advocates continued to push for standardized testing for all students.

### **Shift From Traditional to Testing**

Reese's extensive research on exhibitions was vital when compared to other research because it showed how public schools transformed. Reese discussed that exhibitions were popular but reformers persuaded the public that written examinations could measure reliability through statistics (Reese, 2013, pp. 28-29). Reese also brought up that students were divided into age-graded classrooms in 1837 (Reese, 2013, p. 51).

Both exhibitions and age-graded classrooms are important in the history of standardized testing because it showed the transition from previous practices to current practices. By the 1870's all of the cities within the United States were intrigued with

standardized tests and this phase was acknowledged by John Swett in 1911 (Reese, 2013, p. 161).

Written examinations were seen to improve the way that students and teachers were evaluated. Exhibitions were thought to provide bias because the students were not being judged just on knowledge, but also on presentation. The advocates of testing felt that written exams were more reliable in evaluating students and providing accurate data to compare students in various schools. The traditional way of evaluating students through exhibitions disappeared and testing became the new way of calculating student performance.

### **Test Bias**

IQ testing was similar to standardized testing of today because it was believed that the tests contained biased questions. The biased questions allowed whites to be successful while other races failed to meet set expectations. Eugenics, a belief that genetics within the human species could be improved by discouraging reproduction of less desirable traits and encouraging reproduction of those with desirable traits. Eugenics was believed to be proved by standardized testing.

Lemann pointed out Brigham's book, *A Study of American Intelligence*. In Brigham's book, he claimed that the IQ tests confirm the race order within the United States. The use of IQ tests showed that there was proof of race and class order. "Officers scored higher than enlisted men, the native-man scored higher than the foreign-born, less recent immigrants scored higher than more recent immigrants, and whites scored higher than the Negroes" (Lemann, 1999, p. 30). It was believed that test results created proof that the American culture was in order according to intelligence.

In 1924, Ryan observed the impact of IQ tests in Chicago Public Schools (CPS). The tests were seen as class-biased because they tracked students and were seen to promote administrative progress by shaping the educational system into a corporate model. In addition, teachers tried to resist IQ testing in Chicago but were unsuccessful (Ryan, 2011, p. 342). Ryan also pointed out Frank G. Bruner who criticized the relationship between the testing movement and the eugenics movement. After a 1917-1918 survey, Bruner unmistakably ruled out heredity as the reason for subnormal children (Ryan, 2011, pp. 350-351).

Williams brought up similar concerns about bias in testing when Black children postponed taking tests in 1968 until tests became more culturally sensitive. Williams also expressed that standardized test publishers believed that their tests were neutral for all races and socioeconomic statuses. Williams insisted that the tests were biased in their nature because the test makers were subjective and determined what the tests measured (Williams, 1980, p. 268).

Reese also told the white elite perspective. He pointed out that William Lloyd Garrison believed that African-Americans had a “natural capacity” and their education could never equal white achievement because blacks had smaller brains (Reese, 2013, p. 84).

Another way to prove bias in standardized testing was through social inequalities. Lemann acknowledged Chauncey’s article in the *Scientific Monthly* and a section of his diary during 1948. Chauncey felt it was inappropriate to have different tests for different socio-economic groups. In his personal diary, he stated that higher income groups had

more ability than lower income groups and if heredity was involved ability would continue to future generations (Lemann, 1999, pp. 66-67).

Conant originally believed that testing could allow students from a lower socio-economic status but a high academic level to be selected to go to universities in which they would have normally been overlooked. Testing bias was not considered until it was apparent that race order and socio-economic order were being increased because of testing. Some worked to fight against testing bias while others were oblivious that the problem existed.

### **Statistics**

In 1918, Samuel B. Allison, the head of the Department of Standards and Statistics, pushed to increase the value of quantitative data in the whole Chicago system. He believed that before the data could be measured, there had to be a standard and scale of measurement. The quantitative data could ultimately measure the progress of both teacher and student achievement (Ryan, 2011, p. 250).

During 1989, Winton Manning was invited to a conference in Oklahoma because the University of Oklahoma was interested in increasing their admission of minority students. Manning knew that state universities were short-staffed and the administrators wanted to vindicate their admissions decisions. Manning figured out “the statistical relationship among class, scores, and grades for an individual test-taker” (Lemann, 1999, p. 271).

Manning created the Measure of Academic Talent to allow the states to use quantitative data to select students through SAT scores which accounted for background factors. Later, Manning’s research was refused funding from Episcopal Theological

Seminary during Nancy Cole's time as executive vice president. Unable to fund his research, Manning switched sides and believed that tests should not be used to select students for admission to college (Lemann, 1999, p. 276).

While Lemann discussed the creation and desire for quantitative data, Reese discussed the flaws. "As early as 1849, an observer in the *Massachusetts's Teacher* concluded that local school reports, filled with tables on academic performance, 'are often even more deceptive than those of our Railroad Corporations.' Like railroad reports, school records were often poorly assembled" (Reese, 2013, p. 212). Many administrators also recognized that statistics were untrustworthy because of data that was incomplete and the ununiformed way in which the data was collected.

Lemann discussed Brigham's excitement of validity of intelligence tests in 1924. Brigham defended his low experimental results by saying he believed that his subjects from Yale and Princeton were not working to their full capacity because of distractions (Lemann, 1999, p. 32). This was a legitimate concern because just like today, distractions can cause a decrease in student achievement.

Overall, the authors touched on situations in which someone tried to utilize quantitative data; the data could help to identify slow learning students, increase university admission to minority students, and helped measure academic progress for teachers and students. Unfortunately, in some cases, the individual working towards creating the data started to see the critics' perspective. Reverend Joshua Leavitt believed and stated that numbers do not lie, but when assembled correctly had the ability to deceive (Reese, 2013, p. 137).

## **Conclusion**

Looking into the history of standardized testing brought an interesting view on our current situation in the United States. In many situations reforms were implemented to improve education, but in the end created a wider educational gap between minorities or lower socio-economic students and the elite. Standardized testing is an example of the best intentions gone wrong. In this analysis, issues that were visible in the beginning of testing are still struggles today. Why document history if it is not used to learn from our mistakes? After a century, we should have eliminated the educational gap by decreasing the use of quantitative data that can be manipulated to create bias and reinforce social stereotypes to benefit the elite.

## CHAPTER TWO

### THE CLEARING HOUSE: 1950S & 1960S

#### **Introduction**

In recent years there have been many education debates in regards to standardized testing. Still today (2014) issues such as excessive testing, teaching to the test, weakened teacher control, validity of the tests, testing bias, narrowed curriculum, and many more topics are on personal blogs and in social media. Standardized tests were never my favorite thing when I was in school but I became more passionate about standardized testing when I started studying policy studies and saw the effects policies had on our culture and community.

Curiosity has brought me to evaluate the beginning of widely used standardized testing in the United States. How did parents and educators feel about standardized testing? What were the discussions that occurred in support of or against standardized testing? This is a brief look at the discussions that surrounded standardized testing published in The Clearing House journal during the 1950's and the 1960's.

#### **The Start of Testing**

To see the perspective solely through The Clearing House during the 1950's and 1960's the history of testing will only be discussed as referenced in the journal articles.

Standardized tests began being widely used in schools during World War I (1914-1918) because of the tests success classifying military personnel. In 1958, the National



Defense Act gave standardized tests a permanent place in the U.S (Daniels, 1964).

Internal tests were used for the schools to assess themselves while “external tests [were] those offered to the school by national, regional, or state agencies for use by the agency, not the school” (Tompkins, 1961, p. 515). The use of tests to evaluate students became very common.

As with many policy changes there were some in support of the change while others opposed the modifications. Letters were sent to “congressmen asking for the special rider on the extended National Defense Education Act to prohibit the use of Federal funds for tests and inventories that [elicited] information invading the privacy of the home” (Fredrickson & Marchie, 1966, p. 357)

In the beginning of testing, the most discussed issues were what was appraised and what the instruments of measurement were. Traxler discussed, “for a good many years, leaders in evaluation have advocated this approach to measurement and appraisal” (Traxler, 1953, p. 3). However, critics “[were] able to go back as far as 1936 to quote the chairman of this committee to the effect that “In order to make a list of major objectives usable in building examinations, each objective must be defined in terms which clarify the kind of behavior that the course should help to determine among the students” (Beymer, 1966, p. 540). Unsure of what should and could be tested exactly, the discussion continued with what the assessments should have looked like.

One of the discussions was to closely coordinate testing programs to geographical areas. Schools should not take identical tests, but a common core of tests, and tests that meet the individual needs of schools, classes, and students (Traxler, 1953). There were many mixed reviews in regards to how and who evaluated testing data. George Lucht

believed that the teacher should design classroom measurements. Teachers should determine what should be measured, they should design the measuring instrument, analyze the results to inform the students of their academic standing and finally correct the process to create better results (Lucht, 1960).

Educators were anxious to measure student performance in areas that improved educational performance. However, “in most cases, the classroom teacher [had to] rely on his own training and experience in analyzing test results” (Fredrickson & Marchie, 1966, p. 357).

Soon after, Beymer pointed out criticisms that, “regardless of the title, "National Assessment Program," is in fact a national achievement testing program” (Beymer, 1966, p. 540). The National Assessment Program made funds available to establish and maintain aptitude and ability testing for secondary public and private schools (Scales, 1964).

Daniels suggested that the transition into standardized testing was secretly being done by individuals that should not have the right to exercise power (Daniels, 1964). Many of the educators believed that those with a financed opinion were pushing to change the curriculum. Supporters of the National Assessment Program argued that Project TALENT was a form of assessment and had not corrupted anybody at that time. Critics should not have over-reacted about changes, since previous attempts to change curriculum had failed. (Beymer, 1966). Looking forward to today, some critics also believe that No Child Left Behind has a similar allusion and financial supporter.

### **How to Administer a Standardized Test**

Herman Peters, professor of education at the Ohio State University, went into detail about the importance of properly administering a standardized test. He believed that tests should be “administered with expert preciseness” (Peters H. , 1959, p. 164).

The proper test administration processes, according to Peters (1959), are the following:

- Approach – Testing errors occur if too much emphasis is on the negative aspects. In elementary schools, the standardized test should be referred to as a game that will be played.
- Test Skill – There is a question of validity if students are not familiar with the standardized test procedure. Teachers cannot assume that student’s know how to take a test and should make time to assure students know testing procedures.
- Directions – Test administrators should take the directions seriously because modifying the directions could invalidate the test results. Teachers should practice reading the directions aloud to assure they are precise in the directions.
- Timing – Timing accuracy is essential in comparing local and national results on standardized tests. Teachers should use a stop watch to guarantee the length of the test is aligned with the allowed time limit.
- Group Size – Elementary level students should be tested in groups of ten to fifteen, while older students can be tested in larger groups. The group size helps to monitor student progress on the exam.
- Guessing – When a student is not 100% sure of the answer but has a hunch, it is considered a positive guess. A positive guess is different from a negative guess because the student is marking answers with little attention to the question. The

tester cannot ethically tell a student if they should guess, but instead tell the student that a blank answer will be wrong but a positive guess may be right.

- Follow-through – Scoring the test is vital to making sure that student's score is correct. Scoring the test incorrectly could change the students score and the student could be labeled incorrectly.

Peters' list is important because it points out issues that could invalidate a standardized test.

When looking at errors in test administration, Peters (1959) also discusses

Downie's five principles of psychological evaluation:

1. Readiness can be enhanced when the student understands, values, and accepts the objectives of evaluation.
2. The student should be engaged in learning to emphasis the evaluative instruments.
3. Research proves that individuals learn better when evaluated.
4. Tests are one source of motivation for students.
5. Testing allows the students to be evaluated in a familiar process.

Looking back at childhood standardized tests; many of his recommendations seemed to be well-executed. These guidelines appear to have been streamlined in schools along with the test culture that appears to be effortless today.

### **Validity of Tests**

As Peters warned, without proper execution, standardized testing data can become invalid and ultimately useless. But properly administering a test is not the only way standardized test data can lose its validity. Testing data can be invalid without the proper

assessment tools, students not being tested on the correct academic level, and testing data can be used improperly after the data has been collected and analyzed. A teacher of English at William Dean Howells Junior High School, points out that “Intelligence tests are no more needed for the recognition of genius than for the recognition of idiocy” (Allen, 1958, p. 133). Allen believed that using a test to determine someone’s intelligence could not be scientifically proved. He felt that observing a student could more accurately determine intelligence than a test score.

As mentioned previously, one of the discussions about testing was having a teacher develop, administer, and grade the tests. If this was the case, teachers would have been completely responsible for all aspects of testing. If the teacher was to find an error in the measuring instrument, then they were to blame (Lucht, 1960). But this is not the case. Test developers were in charge of creating and scoring tests.

The American desire to compare products has now entered our school system. Standardized test data allowed for students to be compared to other students and it did not stop there. Data could be used to compare classrooms, schools, districts, and states to determine if they were above-average (Kelly, 1965). Test makers were aware that standardized tests brought a great deal of responsibility; with the development and score of the tests, they were also responsible for test’s validity. Test developers and publishers cautioned educators not rely on the tests’ validity. Despite their warnings, counselors who were often untrained still used the data as if it had been validated (Peters H. J., 1966).

Students are also a big factor when it comes to validity of standardized tests. Wedeen pointed out that in order for a student to accurately take a paper-and-pencil test, the student should be able to read at the expected level (Wedeen, 1960). Therefore, if a

student was not placed in the correct grade level, they were required to test at a level which was outside their reach.

When evaluating a student, the individual should look at many aspects of learning in conjunction with one another. It is almost impossible to look at individual factors without taking other areas into account. As Beymer nicely explained how external factors influence a student's academic achievement. "Standardized achievement tests measure much more than the individual's level of attainment; they also measure the adequacy of his learning environments, home, school, and community, as well as the competence of those responsible for those environments -parents and teachers" (Beymer, 1966, p. 541). Although tests are important, they should not be the only evaluation and are not the most important way to measure progress (Tompkins, 1961).

### **Effects of Testing**

Tests originally started without any stakes tied to the student's test score. In time, state school officials and state legislatures demanded by law that standardized testing determined the educational quality (Kelly, 1965). "The standardized test [had] invaded the classroom" (Kelly, 1965, p. 546) and so did the effects of them. Allen believed, "Schools [required] pupils to pass standard achievement tests as conditions of promotion or graduation [were] less susceptible to charges of watering down courses or of failing to try to instill fundamentals" (Allen, 1958, p. 134). The standardized test started out so innocent, but began to be a big part of education success.

Just as it is almost impossible to isolate one aspect that affects learning, we should also remember that a test score does not only impact a student's academics. The use of

standardized tests, whether they had validity or not, impacted the student academically and emotionally.

When test scores came back, students were academically grouped. In most cases, standardized testing could not identify if the student could apply the knowledge to solve a problem or measure the ability to reason. The brightest students may not have been identified through a standardized test (Kelly, 1965). Students with higher culture capital had higher test scores than students with less cultural capital (Stafford & Shafer, 1961). Advocates for testing believed that testing did not create inequalities, but instead point them out (Beymer, 1966). Higher test scores may have not judged students on true knowledge and could falsely classify students.

Students that scored high on standardized tests were filled with a sense of false knowledge and began to develop a superiority complex; while a student in the slow group soon began to feel inferior and developed a defeatist attitude and accepted the secondary role in school (Urevick, 1965). Students in the fast group focused on grades as end all and did anything (even cheat) to accomplish their goals. On the other hand, the slow group's only goal was to get by, knowing that if they did not, it was alright because they were not smart enough (Urevick, 1965).

Standardized testing allowed for schools to group students of similar academic levels in the same classroom and allowed for teachers to tailor their lessons more efficiently to the group, allowing students to learn more (Urevick, 1965). Students that were categorized according to ability grouping also showed similar traits in behavior. The higher the group, the fewer problems; while the lower, the group the more discipline issues (Urevick, 1965). Ability grouping was unfair and detrimental to democratic ideals

(Urevick, 1965). It was later concluded that students worked best in a heterogeneous classrooms because both fast and slow learners could learn from each other and benefit from being around one another to improve society (Urevick, 1965).

Shortly after, teaching to the test began in an attempt to improve the appearance of public schools by improving test scores. David Goslin, author of *The Search for Ability: Standardization Testing in Social Perspective* was referenced by Beymer stating that teachers, parents, and administrators should improve test preparation to assure high performance of tests. To avoid criticism, they listened! (Beymer, 1966). Teaching to the test is still an issue today.

### **Yield the Best Results**

Standardized tests were used to determine where a child was academically and if improvements were made over time. Traxler states,

Years ago when schools began to use tests in considerable numbers, there was a tendency for them to give the tests at irregular intervals and to choose the tests each time with little attention to what had been given before. Consequently, they obtained scattered test results which were difficult to interpret and which gave little information concerning the growth of pupils because the tests used at various times were different in content and were standardized on different populations (Traxler, 1953, p. 4).

His solution was to use testing programs systematically so that they could be compared year to years and grade to grade, and to record student test scores in their cumulative record. The best use of tests were to administer them around the same time each year and stay consistent with the test or similar tests as long as they could be modified to advance progress (Traxler, 1953).

Traxler also pointed out, "A quarter of a century ago... most schools depended mainly on tests which provided overall, omnibus measurement; that is, tests which



yielded total scores only” (Traxler, 1953, p. 4). His solution to the clumped results was to separate the scores to measure different aspects of achievement and ability (Traxler, 1953).

Traxler determined that we need to be able to compare the student’s score to the norms of each test (compared with others) and their standard score (compared with themselves). Test results should have been put in a profile form so the student’s strengths and weaknesses could be easily compared (Traxler, 1953). Beymer states, supporters believed that critics confused norms (what is) and standards (what it ought to be). The National Assessment Project was set up to collect data and later applied standards.

Stafford and Shafer were skeptical on the use of norms to compare students. They believed national norms may not have been accurate because it was unknown if they had a true sample because if it was true, than half of students in the fifth grade did not meet national standards. Scales emphasized the importance of the normative group to have a true sample (Scales, 1964). On the other hand, if students consistently surpassed the norms, then the school could not take credit for student success (Stafford & Shafer, 1961). Scales revealed limitations of norms; they were created with the majority in mind and did not consider the minority (Scales, 1964).

National norms may not have been accurate in determining students’ achievement because it was difficult to obtain and supply for every meaningful subgroup. No matter how difficult, the testing authors and publishers had to accept the challenge and combine the subgroups in a way that made sense (Scales, 1964). Some test publishers legitimately worked to create true norms while others had illegitimate norms by manipulating the data. Scales believed that testing could be enhanced if test publishers had accurate norms

but until then educators should reject all assessments that did not have accurate subgroups (Scales, 1964).

Traxler did not want students to be judged exclusively on testing. He believed that students should supplement the common test with other appraisal devices for individuals and groups (Traxler, 1953). Many factors should have been considered when evaluating the students. “The supplementary devices include[d] especially interest inventories (often a part of the common testing program), inventories of personal qualities, anecdotal records, projective techniques, and sociometric devices” (Traxler, 1953, p. 5).

Tompkins stated similar issues that aligned with today’s issues. Schools were skeptical of large-scale external tests because students were taking too many tests, tests were scheduled during class time and instruction suffered, large amounts of testing effected the school’s curriculum, testing spread to multiple grade levels, tests undercut the importance of teaching, teaching to the test became necessary, comparing schools solely on test scores did not benefit the majority, test makers were making millions of dollars, and test makers created more work for schools and staff (Tompkins, 1961).

In order to yield the best test results, Traxler believed locally constructed tests were the best solution. Traxler stated, “For reasons of convenience and saving of expense, standardized tests [were] preferred to locally constructed ones wherever they [did] the appraisal job the schools need” (Traxler, 1953, p. 6). The use of locally constructed tests were encouraged because it was geared more closely to the school’s curriculum and objectives than mass produced tests (Traxler, 1953)

In addition, the use of standardized test scores became a time saver because counselors and psychologists did not have to meet with teachers to discuss individual student performance (Fredrickson & Marchie, 1966). Traxler believed teachers and counselors should have access to general tests and other instruments of appraisal. The best use of the results was by teachers and counselors to improve guidance and instruction for the individual student (Traxler, 1953). However, Scales pointed out that in order to evaluate students based on national norms, the norm must be as accurate as possible (Scales, 1964).

Critics of testing did not believe that tests were absolutely valid and they worried that too much faith was placed in the results (Daniels, 1964). Standardized test data should have been used to benefit the individual learner rather than the groups. After correcting the student's low performance, follow-up data should have been used to determine if the changes improved the student's performance (Stafford & Shafer, 1961).

“The only defensible position with respect to norms [was] that they should have been interpreted as aids in the evaluation of relative aspects of the school's program, not as standards to be attained” (Stafford & Shafer, 1961, p. 271). The overall argument was that standardized tests should be used as the means and not the end (Stafford & Shafer, 1961). Contrary to the belief that test scores should have been used as a means and not the end, today we use them as an end because of current educational policies.

### **Testing Data**

When schools were criticized for student achievement, administrators and the board of education used standardized test scores to defend their progress (Stafford & Shafer, 1961). When schools were negatively criticized because of tests scores it was

important to remember Peters' sources of error. He mentioned that testing errors could occur if the approach towards the tests was portrayed as negative (Peters, 1959)

Testing data should not have been used alone to prove a school's success when being compared to national averages. Achievement test data was vulnerable, therefore it allowed the school's claims to be refuted (Stafford & Shafer, 1961). Test makers agreed that tests were not completely valid. They admitted there was a margin of error and tests were an estimate of future success (Daniels, 1964). Schools misleadingly used the mean test score to signify the entire school's academic performance. The mean could be inaccurate because of overlapping test scores or because of a shift in the student population (Stafford & Shafer, 1961). This also occurs today to determine school success.

### **New Pressures**

In 1965, Kelly indicated how parents and students judged teachers more than they had ever done in the past because standardized tests put pressure on the students, parents, school board, and administrators. The school board evaluated teachers on the basis of standardized testing data. Kelly noted, "After a closer look, it [was] apparent that pressure from a variety of groups and individuals [had] forced the teacher into objective [measurements] rather than a comprehensive evaluation program to determine student grades" (Kelly, 1965, p. 546). The pressures that arose because of standardized testing were new and created the appearance that schools were failing.

Critics believed that the data from standardized tests was being used to unjustly attack schools and forced them to teach to the test. Critics also said that the National Assessment Program created pressures that altered curriculum to improve test scores

(Beymer, 1966). More regularly, the test results were being used as a tool to determine if a student progressed to the next academic level (Peters H. J., 1966). “Some testing [was] desirable; too much [was] bad for students, teachers, and school programs. We [should either] cut down the number of external tests or [reduce] their interference with curriculum” (Tompkins, 1961, p. 516).

With the new pressures, one might have assumed that teachers rejected the use of testing. On the contrary, a limited survey was done during the 1960’s to determine the percentage of teachers’ acceptance of testing records. Fredrickson and Marchie concluded, “The accumulated data appeared to support the finding that a majority of the teachers in the study failed to question a test score as fixed value or mention the test’s possible limitations in use” (Fredrickson & Marchie, 1966, p. 358). Although this was a limited survey it showed that teachers did not openly object to testing and Beymer agreed that more educators supported standardized testing (Beymer, 1966).

### **Testing Advocates**

“The purpose of testing is to understand someone better”

– Scales, 1964, p. 202

Tompkins exclaimed, “Almost every informed person [agreed] that a balanced testing program in secondary schools [was] both good and desirable” (Tompkins, 1961, p. 515). Lucht discussed that teachers spent little time on perfecting their tests because teaching duties filled most of their time. Lucht continued by stating, “Testing is teaching and teaching is testing”, therefore testing should be a part of the regular student learning routine (Lucht, 1960). Educators believed that education was becoming too big and expensive without systems to determine progress (Beymer, 1966).

Due to the increased advertisement of tests, parents believed test scores were the end-all. “Extravagant claims for the efficiency of testing programs [had] been made over a long period by those who [prepared] and [marketed] them, with the result that both parents and the public generally [had] come to believe that tests [were] infallible” (Tompkins, 1961, p. 515). When standardized testing data was analyzed the information was used to create an improved education program rather than to prove that a school was successful (Stafford & Shafer, 1961).

Critics made claims to end standardized testing in schools, but advocates responded by saying that before eliminating standardized tests, anti-testers should provide an alternative to evaluate community members (Daniels, 1964).

Until the objectors can advance a better method for evaluating individual potential and economic efficiency and personal satisfaction, standardized tests were the best available means we have for coping with a complex and fluid and impersonal society. I consider them a necessary, albeit imperfect, by-product of progress, and I accept them in the same spirit that my father accepted Federal income taxes (Daniels, 1964, p. 14).

As Daniels put it, if critics did not approve of standardized tests, they needed to come up with a better assessment.

Before testing, students were assessed on a person-to-person basis. The student was judged on the length of time and location of their current residence and the performance of other family members. Although the previous process was more personal, it was filled with bias and proved less valid than standardized tests (Daniels, 1964). Testing data should be used to help evaluate where the school system needed to make improvements and the school could show that they improved with later assessments (Stafford & Shafer, 1961). The use of testing helped remove personal school level bias and could possibly improve education if used properly.

Testing allowed for students to be evaluated by outsiders without an extensive interaction. Testing created a way for colleges and universities to evaluate a larger amount of students in a shorter time. Daniels stated, “Without tests, the Dean of Admissions will be reduced to rejecting everyone with blue eyes. Or all the Baptists” (Daniels, 1964, p. 13). He believed that without testing, not only would organizations suffer, but so would individuals.

### **Testing Take-Over**

Stated in the editor’s note of *Unintelligence Tests*; “No one can seriously say that standardized testing is not a big business” (Allen, 1958, p. 131). “In 1961 Buros' published a rather complete bibliography of tests which included 2,126 separate, different tests or test batteries” (Scales, 1964, p. 196). Out of the 2,126, 46.4% of the tests were classified as achievement tests and the remaining were categorized as personality or I.Q. tests (Scales, 1964).

Scales believed that testing was here to stay. “There can be little disagreement that tests often play an important part in many decisions which sometimes critically affect the lives and careers of young people in our society” (Scales, 1964, p. 197). Urevick concluded that the use of standardized testing to categorize students into an ability group was the worst use of testing. If a student was placed into the slower group, they received a lower quality of education and standardized test scores reflected the consistent low test scores (Urevick, 1965).

“While it [was] extremely difficult to obtain the proportion of kinds of tests sold, a conservative estimate based on this report [indicated] that approximately two million commercially, produced ability tests were administered in 1961” (Scales, 1964, p. 196)

Scales was unable to determine how many tests were administered annually, but he believed that more ability tests were administered than people in the United States (Scales, 1964). College freshmen were so accustomed to taking tests, they thought that aptitude tests could help them determine what they should do as a career. However, aptitude test results may not measure a student's true potential (Wedeen, 1960).

“Our mind is an erratic, dangerous and  
unthinking tool; it is difficult to reconcile it with  
order and moderation”

– Peters H. J., 1966, p.266.

### **What Should Assessments Look Like?**

Lucht believed that assessments should not be limited to a paper-and-pencil device (Lucht, 1960). Peters' elaborated, “Too often emphasis [was] given to the statistical impact on the individual, when it would [have been] better to spend the time in the counseling interview assisting the individual to think through his singularly unique concerns with their group overtones as revealed by the seeming magic of mathematics.” (Peters H. J., 1966, p. 224).

Too much faith was being put into testing to determine academic performance and future success. Vocational aptitudes were unable to assess the student's personality and therefore could place a student in a profession that did not make them happy or guarantee job availability after school or training (Wedeen, 1960). To assure true knowledge, follow-up data should have been gathered after graduation to evaluate ultimate objectives instead of immediate objectives (Stafford & Shafer, 1961). Ureic proclaimed, “The results of our teaching will not be measured by the test scores our students make on a



standardized T test, but by the impact our students make on our society” (Urevick, 1965, p. 530).

Lucht cried, “One might suppose that the perfect measuring instrument will never be made” (Lucht, 1960, p. 76). The editor wrote, “The author admits that standardized tests [were] not perfect but suggests that they should be used within reason until something better comes along” (Daniels, 1964, p. 12). However, it was necessary to only utilize standardized tests to improve the instructional program (Tompkins, 1961). Teachers recognized that standardized tests did not always identify the brightest students. However, the testing trend would not change until teachers were recognized as being able to give subjective evaluations (Kelly, 1965). Every child should be provided with quality education, we cannot sacrifice the slower students to benefit the faster students (Urevick, 1965).

### **Conclusion**

Looking at standardized testing in the 1950’s and 60’s brought an interesting perspective to current standardized testing issues. In 1961, Tompkins, *External Tests and the School* was mostly aligned with critics of today’s standardized testing. As I was reading it, I had to remind myself that the piece was written in the 1960’s. Tompkins touched upon excessive testing, loss of instruction time, changes to the curriculum, wider use of testing, weakened teaching and teachers, teaching to the test, use of test scores to solely evaluate students and schools, and testing publishers’ financial success from testing (Tompkins, 1961). Critics of today would praise Tompkins’ perspective.

However, the advocates’ perspective was not consolidated into a single historical piece. Just as today there are teachers that are at various levels when it comes to testing

perspectives there were similar perspectives in the fifties and sixties. Used correctly, standardized tests could be used to improve education and help individual students succeed. Testing could help point out imbalances in the system between schools and districts, and to remove in-person bias. Although standardized testing is not perfect, it can be used to help improve educational quality when used correctly.

I'll conclude with a quote from Urevick, "Our nation's future will always be measured by the caliber of citizens we produce: by the way they act, feel, cooperate, and contribute to each others' well-being" (Urevick, 1965, p. 530). We as a society have to use our past to improve our future. If history shows irrefutable proof that something does not work than we have to work, to improve the process to improve our society.

## CHAPTER THREE

### PURPOSE OF THE STUDY

#### **Introduction**

The debate about high-stakes testing and accountability is not a new debate. Students and states have been debating the pros and cons for decades. In 1981, Debra P. V. Turlington involved African American students who failed tests and challenged the Florida Student State Assessment after they failed to receive a passing score on a high-stakes exam (Holme & Heilig, 2012). The controversy of biased exams was one of the major issues then as it is now. Today, “Students and parents across the country have been taking the lead in organizing test boycotts as well as other forms of protest” (Kohn, 2005, p. 326). Students understood that these evaluations may not have been fair and they were taking a stand for what they believed in, but did their voice matter to educational policy makers?

The objective of this research study is to have a better understanding of how students feel about high-stakes testing and how it affects them. The student perspective is important to get them engaged and on track to become successful long-term learners. Bader explains the importance, “give voice to students in order to better understand how they define their identity” (Bader, Horman, & Lapointe, 2010, p. 26). The curriculum should be aligned with the student population’s culture and identity to allow the students to connect learning with actual life or community events.

Contrary to many beliefs, teachers and policies are not the main contributor of student academic growth. A student's socioeconomic status and experiences have a lot to do with a student's success in school. Sandy explains, "gains in student achievement over time are a consequence of changes in the socioeconomic characteristics of families" (Sandy & Duncan, 2010, p. 297). This is important to this research because it will allow for students to vocalize if they feel their test results are influenced by more than school curriculum and environment. Students' input on high-stakes testing could help policy makers better understand the students and what works best for them.

### **Significance of the Study**

Much of the research that has been done on high-stakes standardized testing had little information that included how students feel towards the tests. This research will add a much needed perspective on high-stakes testing. The research was conducted through community organizations in Chicago, as well as through parent organizations on social media sites. The use of multiple community organizations and online media allowed for a wider variety of students and school types.

Chicago Public School students are required to complete the Illinois Standards Achievement Test (ISAT) and NWEA with high stakes attached. Students in the benchmark grades (third, sixth and eighth grade) are reviewed to be retained if they do not meet Achievement Level 2 District Wide Assessment (DWA) scores<sup>1</sup>. Collecting data from a variety of students allowed for a deeper understanding of the influences on the curriculum and overall school climate. Will students at a school on probation have more pressures and anxieties when testing? Does the school's status affect the students' feelings towards themselves or their school?

1. See Chicago Public Schools, 2007, p. 2. Students in the benchmark grades have Achievement Level 2 status: (a) when their DWA reading and mathematics scores are both below the 24th National Percentile Ranking, or (b) when one of their DWA scores (either reading or mathematics) is below the 24th National Percentile Ranking, or (c) when they have no DWA scores in either reading or mathematics, except for those students who are not required to take the ISAT.

This research examined the student perspective and feelings through an online survey and interviews. The data evaluated if student's testing scores were dependent on how students felt towards high-stakes testing. The research was completed using mixed methods but emphasizes qualitative data to assure the students' voices are portrayed. It was anticipated that the student's feeling towards high-stakes testing would be dependent on their performance on the exams.

### **Variable and Data Sources**

The purpose of the study was to understand if a student's perspective towards high-stakes testing was dependent on their test score. The data that has been collected on accountability standards has had little documents that included the student's perspective or feelings towards testing. "The U.S. Department of Education acknowledges that the urban achievement gap is due, in part, to the high concentration of minority and low-income students who, on average, do not perform well on these tests" (Sandy & Duncan, 2010, p. 302). This is important because it shows that research data has yet to change the policies that may have negatively influenced students' self-esteem or overall assessment of themselves.

### **Potential Limitations**

The data will be collected from 5<sup>th</sup> and 6<sup>th</sup> graders who completed the online survey. Sandy and Duncan (2010) explained many issues that could affect student achievement on high stakes standardized tests:

Urban children are more likely to face limitations to their educational experience such as parents who are more likely to be divorced and unemployed, difficulty with English, or emotional and learning disabilities... higher rates of unemployment and divorce among parents in poor urban neighborhoods weaken the kinds of community ties that are associated with parental involvement in schools... youth in poor urban areas are exposed to attendant risks such as risks to

safety and health, as well as greater participation in risk-taking behaviors that are associated with lower educational outcomes (p. 298)

Sandy and Duncan showed that collecting accountability results from an urban school could have many aspects that influence student achievement. The multiple aspects from students could result in a lack of themes or trends in student responses.

### **Research Question**

Do high-stakes accountability and/or curriculum changes (due to testing) influence students' perspective of themselves and/or their perspective about their school or community?

## CHAPTER FOUR

### ACCOUNTABILITY

#### **Strengths**

In 2002, President Bush passed the No Child Left Behind Act (NCLB). NCLB increased education accountability and the movement has continued to increase in momentum for nearly a decade (Jacob, 2002). NCLB allowed for implemented sanctions that Hong and Youngs (2008) discuss:

Severe sanctions for schools and districts that fail to make adequate yearly progress (AYP), supporters of NCLB argue that the law is pressuring schools to improve instruction for minority and low-income students, two groups that have traditionally been poorly served by public schools in the U.S. (p.2)

NCLB required teachers to be highly qualified in the subject area in which they were teaching. Teachers in low-income schools were more likely to be teaching a subject in which they were not certified. Therefore, low qualified teachers were potentially adding to the educational gap.

High-stakes testing required teachers and students to be more accountable for learning. Hong explained, “advocates of high-stakes testing have contended that NCLB and similar state policies are necessary to ensure that teachers and schools maintain high standards for low-SES and minority students and help them achieve at high levels” (Hong & Youngs, 2008, p. 4). Jacob stated, “there were dramatic increases in math and reading achievement under high-stakes testing” (Jacob, 2002, p. 15). Low-achieving schools had

the largest gains; it appeared that these schools focused more on test prep after the accountability policy was started. (Jacob, 2002).

The purpose of accountability policies was to decrease the achievement gap, and early studies show that test results of low-income and minorities did begin to increase (Diamond & Spillane, 2004). “Indeed, after the inception of the new accountability program, Chicago students’ test scores rose to some extent and the proportion of students in the gate grades with test scores below the minimum standard for promotion fell significantly” (Hong & Youngs, 2008, p. 10). In addition, Jacob states, “Iowa Test Basic Skills (ITBS) increased sharply following the introduction of the accountability policy (Jacob, 2002, p. 5). The use of test scores did allow for U.S. public schools to appear that accountability policies were working.

### **Data Manipulation**

Data manipulation can occur if a person or organization wants outsiders to believe that research proves a certain outcome. “Any aspect of learning (or life) that appears in numerical form seems reassuringly scientific; if the numbers are getting larger over time, we must be making progress... is easier to measure efficiency than effectiveness” (Kohn, 2005, p. 316). Therefore, if data is collected and results go up, onlookers believe things are getting better and if numbers go down, onlookers believe things are getting worse. Testing quantitative data, can then be guided to support a variety of perspectives.

In 1982, a Nation at Risk created a false belief that public schools were failing, allowing policy makers to insist on tougher graduation requirements (Holme & Heilig, 2012). Today, a similar situation is occurring with high-stakes standardized testing data. Kohn explains, “no matter how difficult the questions are, the pattern of results is



guaranteed to be the same: Exactly 10% of test takers will score in the top 10% and half will always fall below the median” (Kohn, 2005, p. 318). This statistical calculation allows for a bell shaped curve with a majority of students being close to the median and a small percent being both in the top and bottom percentage.

States and school districts are allowing for tests to make the public believe that the public school system is failing. “It may reflect a desire to cast public schools in the worst possible light as a way of justifying a privatization campaign (any one, after all, can invent a test that many students—or teachers— will fail)” (Kohn, 2005, p. 316). Jacob (2002) points out those schools, in an effort to improve test scores, are negatively impacted by low-ability students; place students in special education programs or limited English proficient classes which exclude the students from testing. Hong states, “Texas had excluded higher percentages of students from taking the NAEP tests than most other states” (Hong & Youngs, 2008, p. 8). What are the long term effects of low-ability students placing in special education programs?

Test data is also influenced by the test makers themselves. In the U.S. society, students that are not categorized as upper socioeconomic status (SES) and whites are considered the minority. This allows test manufactures to justify or deny that tests are created for individuals with high social capital. “A test item on which African-Americans do particularly well but whites do not is likely to be discarded because of the interaction of two factors: African- Americans are a minority, and African-Americans tend to score low” (Kohn, 2005, p. 324). The choice to eliminate questions that African Americans answered correctly, allowed for biases to influence the results of students from various demographics and SES levels.

When schools do not meet standards they are put on probation, and are penalized if test scores are not improved. This fear results in schools changing their curriculum to raise test scores. Jacob states, “test preparation associated with high-stakes testing may artificially inflate achievement” (Jacob, 2002). When test scores are improved because of test prep, artificial scores do not allow for the tests to be accurately compared to other exams.

Teaching to the test allows for schools to increase their test scores, but may not help to improve actual student learning. Jacob (2002) explains that students did improve in basic skills twice as much as complex skills. Trying to manipulate test scores did change results, but the results did not accurately evaluate students. Kohn explains, “they overestimate what some students know... On the other hand, standardized tests underestimate what others know because, as any teacher can confirm, very talented students often get low scores” (Kohn, 2005, p. 317). Will low test scores stop students from becoming life-long learners?

The total amount of misclassified students was broken down by Kohn, “41% had memorized the process without really understanding the idea, whereas 11% understood the concept but made minor errors that resulted in getting the wrong answers. A standardized test therefore misclassified more than half these students” (Kohn, 2005, p. 317). When schools work to stay off of probation, legalized manipulation occurs and students’ testing results are not accurately portrayed.

### **Retention and Graduation**

When high-stakes accountability was established students were subjected to being retained or not being allowed to graduate if they did not meet testing standards. Jacob

(2002) explains that 25 states allow for high-stakes testing to determine if a student is allowed to be promoted or graduate. When students do not meet accountability they are required to attend summer school. If the students do not successfully meet standards at the end of the summer they are retained. Jacob states, “Roughly 20 percent of third grade students and 10 to 15 percent of sixth and eighth grade students were ultimately held back in the Fall” (Jacob, 2002, p. 9). Retention rates increased by 33 percent from 1993-95 to 1997 for students in the first grade; 100 percent for students in the second grade and grades four, five and seven increased to 150-200 percent (Jacob, 2002, p. 37).

In addition to an increase in retention of students that do not meet accountability standards, there was also an increase in drop-outs (Diamond & Spillane, 2004). Hong explains, “while the state tests provided opportunities for some low-income and minority students to acquire cultural capital, they led many others to leave school without their diplomas—a key form of institutionalized cultural capital.” (Hong & Youngs, 2008, p. 9). Retention and drop-out rates are important because they may influence how students feel about high-stakes accountability. What are students’ feelings about being retained or not being able to graduate if they do not meet standards? Hong describes, “findings were consistent with other research findings that retention had more harmful effects on matured students who were more sensitive to their social reputation” (Hong & Youngs, 2008, p. 12). High-stakes testing is focused on short-term goals, but we as a society should focus on long-term development.

Unfortunately, “Students may be punished for low scores by being held back a year—or, in the case of exit exams, prevented from graduating on the basis of a single

test regardless of their academic records” (Kohn, 2005, p. 319). Retention policies are explained by Hong and Young (2008):

Proponents of retention policies argued that establishing cutoff standards, making clear that achievement matters, and imposing negative consequences would lead students to work harder and teachers to pay attention to the needs of the lower-performing students. They also contended that if students have not mastered basic skills, they would be better served by repeating the same grade and gaining those skills than by struggling with more advanced materials (p. 10)

The data from my study will help understand how students feel about retention and graduation. Do students drop-out because they are discouraged by accountability standards?

Kohn believes, “there is little question about which students will be disproportionately denied diplomas as a consequence of failing an exit exam or which will simply give up and dropout in anticipation of such an outcome” (Kohn, 2005, p. 325). The drop-out rate affects urban students at a higher rate. Hong and Youngs state, “To the extent that dropout rates among low-income and racial minority students remain the same or increase under high-stakes testing policies, it seems likely that such policies would not help such students acquire institutionalized cultural capital” (Hong & Youngs, 2008, p. 5) In addition, “This higher rate of grade retention among African American students seemed likely to eventually result in higher dropout rates” (Hong & Youngs, 2008, p. 12). Is the increase in the drop-out rate directly correlated to economic status?

Holmes and Heilig feel that leaders in low-income schools need to be accountable for the increased retention and drop-out rate. “The largest responsibility falls upon leaders of schools serving low-income students, students of color, and ELLs. These student populations are not only more likely to fail the tests but also more likely to drop out as a result of failing” (Holme & Heilig, 2012, p. 1178). Accountability policies

showed research that was not expected by policy makers as Hong and Youngs (2008)

explain:

There were indications that students' performance on state or district tests increased following the implementation of high-stakes testing and accountability policies... research from Texas and Chicago revealed that these accountability policies seemed to have had deleterious effects on curriculum, instruction, the percentage of students excluded from the tests, and student dropout rates (p.13)

The negative issues should have been addressed to help improve learning and student self-esteem. But instead, curriculum and academic changes were focused more on staying off probation.

### **Curriculum and Student Focus**

High-stakes testing is done yearly to determine student academic growth. This standard sets the one size fits all learning mentality that every student learns the same information at the same pace (Kohn, 2005). Diamond & Spillane (2004), Jacob (2002) and Kohn (2005) show that resources in curriculum have shifted to teach to the test and ultimately ignore the lowest performing students. When schools are coming close to getting on or off probation, they put additional resources into students that are close to meeting standards while ignoring the students that have the lowest scores.

The main goal of accountability was to decrease the achievement gap, but "low SES and racial minority students lost opportunities to acquire embodied cultural capital in the form of higher-order thinking, analytical writing, and problem-solving skills" (Hong & Youngs, 2008, p. 8). The focus on test preparation minimizes the opportunity for more complex thinking because the curriculum is more focused on testing drills. "Lipman also contended that while more affluent students in Chicago were engaged in intellectually challenging curriculum, low-income and minority students had to memorize fragmented

facts and information and master simple test-taking techniques” (Hong & Youngs, 2008, p. 11). As discussed by Bader (2003), for students to acquire real learning, they should be involved in the projects themselves.

Students are losing their desire to learn due to test preparation. Kohn explains that it becomes harder to get students involved, “The more prominent and relevant the tests become, the more difficult it is for teachers to invite students on an intellectual adventure, to help them acquire the ability and desire to solve realistic problems in a thoughtful way” (Kohn, 2005, p. 321). Higher test results do not confirm student learning. “Other studies also corroborate the contention that improved test scores do not mean that students have learned higher-order thinking and academic skills... test preparation for students had replaced potentially rich educational experiences in many urban schools” (Youngs, 2008, p. 11). Jacob (2002) believes that high-stakes testing has potential but should be approached with caution.

What remains important in education is the students and how they feel about learning. Program planning should include students’ perspectives and needs. Jean explains, “These efforts are increasingly important if early care and education programs are to fulfill their promise to prepare children from diverse backgrounds to meet the challenges of formal schooling at the beginning of the primary grades” (Jeon, Langill, Peterson, Luze, Carta, & Atwater, 2010, p. 934). Looking at the student perspective in early education also helps to confirm that the student perspective is important throughout the student’s educational career.

Teachers are at the front line of the educational fight to improve student learning. Athanases discusses the importance of teacher teamwork and linking subject areas to

improve student learning. “Professional development is most successful through sustained and intensive efforts, and when it includes subject matter focus, teachers' active learning, and links to daily school life, it is more likely to produce enhanced knowledge and skills” (Athanases, 2003, p. 109). It is important for teachers to encourage rich thinking, discussions and writing. This will help to increase learning by connecting lessons to student life experiences.

Athanases continues, “middle school students, including those academically underprepared, can benefit enormously from a thematic approach to literature that broadens potentially narrow perspectives and that provides context for learning” (Athanases, 2003, p. 118). “When a theme appeared to fit developmentally and socially, it provided context and motivation, giving students a reason to be interested” (Athanases, 2003, p. 113). When a teacher is excited and motivated about learning, so are the students. To accomplish a deeper understanding of material, teachers need and want support to improve student knowledge (Athanases, 2003).

### **Other Issues**

Variances in high-stakes testing are most effected by socioeconomic issues. Improvement in achievements and gains can be due to improvements in family characteristics (Kohn, 2005; Sandy & Duncan, 2010). Hong and Youngs explain, “researchers have shown that the lack of cultural capital among low-income and minority students can result in reduced access to school resources and academic and social supports from teachers” (Hong & Youngs, 2008, p. 3). Students in a higher socioeconomic status had the culture capital that allowed them to be successful in school. “Bourdieu argues that schools contribute to reproducing existing social hierarchies”

(Hong & Youngs, 2008, p. 4). Do the current systems in place create an invisible caste system within the U.S.?

There are other ways to improve student learning without forced accountability on teachers and students. Jeon (2010) explains:

Lower teacher–child ratios... and smaller group sizes... are associated with higher scores on measures of global classroom quality ... Higher global quality scores are related to more extensive teacher–child interactions... less restrictive and controlling teacher behavior, and more complex language and play (Jeon, Langill, Peterson, Luze, Carta, & Atwater, 2010, pp. 914-915)

In addition, “Children’s relationships with their teachers are important because they predict later academic and behavioral adjustment” (Jeon, Langill, Peterson, Luze, Carta, & Atwater, 2010, p. 931). Sandy and Duncan state, “class size seems to be the most promising tool for improvement” (Sandy & Duncan, 2010, p. 311). However, regardless of curriculum, every student will have a different experience even when they are in the same classroom.

A strong curriculum keeps students engaged in learning. What is important in curriculum, is keeping students to stay engaged in learning. Curriculum should incorporate the student’s culture and experiences. Athanases states, “selected texts at times were problematic because of a lack of attention to cultural diversity of characters and authors, or due to inclusion of works depicting people of color in disempowered states, written from white perspectives” (Athanases, 2003, p. 118). Kohn says, “The more a test is made to “count” in terms of being the basis for promoting or retaining students or for funding or closing down schools, the more that anxiety is likely to rise, and the less valid the scores become.” (Kohn, 2005, p. 316). Increased anxiety could result in students’ withdrawal from wanting to learn.



The transformational leadership project in Quebec allowed for the students to form a single community (Bader, Horman, & Lapointe, 2010); while in the United States, society encourages students to compete against each other.

Students in Quebec took part in a transformational leadership program which required members of the community to make changes to improve their communities (Bader, Horman, & Lapointe, 2010). In the United States, policy makers are making decisions on how to improve student achievement without the help of students or community members. Giroux is quoted “Hope must be tempered by the complex reality of the times and viewed as a project and a condition for providing a sense of collective agency, opposition, political imagination, and engaged participation” (Bader, Horman, & Lapointe, 2010, p. 36). Does student involvement increase the learning environment?

Students are required to have similar outcomes on high-stakes standardized tests but students in different SES schools do not receive the same resources. Students in lower SES schools have less social capital, and therefore do not perform as well on high-stakes tests because of lower quality learning environments (Hong & Youngs, 2008; Sandy & Duncan, 2010; Jeon, Langill, Peterson, Luze, Carta, & Atwater, 2010). Sandy and Duncan (2010) explain:

U.S. Department of Education (2002–2007) reports that urban schools, compared to the rest of the nation, have significantly more students testing below the basic level in reading, math, science, and writing on the National Assessment of Educational Progress (NAEP) test... In addition to lower achievement, urban students are also more likely to be poor (p. 297)

Sandy and Duncan (2010) believe that the difference in achievement is because of family backgrounds and race (Diamond & Spillane, 2004); “family background characteristics

explain no more than 25% of the gap, whereas differences in school quality explain from 50% to 69%” (Sandy & Duncan, 2010, p. 310).

Money is an issue with current accountability practices. Kohn (2005) believes, that efficient tests do not allow important areas of learning to be taught or tested. Are high-stakes tests more about maximizing profit instead of maximizing student learning? “The process of assigning children to percentiles helps ensure that schooling is more about triumphing over everyone else than it is about learning... every distribution of scores will contain a bottom, it will always appear that some students are doing badly” (Kohn, 2005, p. 319). Jacob (2002) explains that accountability policies are not usually implemented alone, which makes it hard to determine which policy is achieving or diminishing growth.

### **How do Students Feel?**

The current research on accountability and high-stakes testing includes little data that discusses the student perspective. Jacob explains, “Interview and survey data provide evidence that students, particularly students in the sixth and eighth grades, were acutely aware of and worried about the accountability mandates” (Jacob, 2002, p. 29). Kohn states, “students who perform well on tests are often those who are least interested in learning and least likely to learn deeply” (Kohn, 2005, p. 317). Is this true? How do student’s test scores compare to their passion for learning?

In Quebec the student perspective was taken into account when they felt that their school reputation defined them as individuals; Bader explains:

The project emerged out of the need to respond to an event that threw into question not only the way these young people defined themselves, but also the type of relationship that they maintained toward their school, which they

esteemed and appreciated, but which enjoyed no public recognition (Bader, Horman, & Lapointe, 2010, p. 28).

The reputation of the school or community influenced how students felt about themselves and what they could accomplish. Do students taking high-stakes tests identify themselves with the test scores? The students felt that just because something bad happened in the school or community does not mean that is who they are (Bader, Horman, & Lapointe, 2010). “He also stressed the importance of seeing culture and media as forms of pedagogy which shape the way young people define themselves” (Bader, Horman, & Lapointe, 2010, p. 26). Ultimately, “teachers and students did not regard test scores as an expression of academic abilities” (Hong & Youngs, 2008, p. 11). Do students truly believe that test scores are not a reflection of their intelligence?

## CHAPTER FIVE

### I LIKE TO READ, BUT I'M NOT GOOD AT IT

Dutro and Selland's research was most closely related to the research that was conducted. The results showed that students knew that high-stakes standardized tests were important and that there were heavy consequences associated with them. However, they were unsure of who was collecting the data. "Children expressed understanding of high-stakes tests as being used to judge their own learning and performance" (Dutro & Selland, 2012, p. 353). The students were most concerned with the thought of retention if they did not do well on the tests. When describing who collects the data, the students used pronouns such as "they" and "them" and described someone behind the scenes. Dutro reported, "19 of the 33 children indicated that the adults invested in testing were located outside of their school" (Dutro & Selland, 2012, p. 352). The students thought that schools whose scores did not meet scores meant that the teachers were not good. They were aware that low performing schools could be closed.

As for the students personal perspectives, "children in our study did not express strong negative emotions in relation to tests" (Dutro & Selland, 2012, p. 356). Students that scored proficient on the tests had positive feelings toward high stakes testing. The test results influenced how students felt in regards to their reading level. If the students performed poorly on high-stakes standardized tests, they felt that they were bad readers. Dutro used Molly as an example "She had told me that she loved to read and was happy that she could finally read well enough to have chapter books. But, she said, 'I know I'm

not good at it. I do bad on those tests. When we take them, I just know it will be another low points, so the books I like, like I know they are too low for those tests.’ Molly, like all students, negotiates high-stakes testing as part of her school experience” (Dutro & Selland, 2012, p. 341).

Students that score proficient receive praise while students who have a limited score do not. This is important because this affects the student’s perspective of themselves and their environment. Dutro and Selland bring up important aspects of high-stakes testing and how students feel about their reading proficiency. Dutro spent over two years at Davis Elementary School observing, collecting and analyzing quantitative data, but had a small sample size. This study is a great start to future research that has a larger population at schools with comparable and contrasting demographics.

## CHAPTER SIX

### METHODS

#### **Participants**

The study consisted of subjects that were in 5th and 6th grader, and who had completed a benchmark grade or were currently in a benchmark grade. The students participated in the study by voluntarily completing an online survey. The online survey collected data from students from a variety of locations and ages. To keep a consistent research sample, interview participants were selected based on the following: the student resided within Chicago or the Chicagoland area, was currently enrolled in 5<sup>th</sup> or 6<sup>th</sup> grade, was between the ages of nine and twelve, and had completed a high-stakes test.

The age range was selected because the students were at a good age to provide assent, and they had been subjected to high-stakes and standardized tests for several years. Additionally, the researcher believed they were more detailed in regards to their emotions than younger students. Finally, the age group had been tested in a third grade benchmark and was expected to be tested in another sixth grade benchmark shortly. The scores on a high-stakes test in a benchmark grade determine if a student is promoted, retained or required to attend summer school.

A variety of perspectives was included in this research because the sample included students that were not labeled and others that were labeled as gifted. These students passed the test and moved to the next grade without being exposed to the

accountability measures. The sample also included a student that possibly failed the high-stakes test and was subjected to the accountability process. Although the sample included a wide variety of academic and socioeconomic levels, the sample size was too small to allow for a theory to be determined.

### **Instruments**

In the research design, qualitative data was the most appropriate for this study because it publicized the student voice. The research evaluated how high-stakes standardized tests affect students in variety of socioeconomic areas. In this study, an online survey was used to evaluate the student's feelings and current opinions on high-stakes standardized testing. The survey consisted of both quantitative and qualitative questions which allowed for both quantitative and qualitative data to be collected from every subject; however, the quantitative data was not used in the final research because of the small sample size.

Utilizing the online survey, five students were selected to complete a one-on-one interview. The interviews allowed for a more thorough understanding of the students' responses. The students provided demographic information and additional quantitative and qualitative data with the survey. The interviews consisted of only qualitative data which was converted later. The thematic categories were not predetermined, but instead flowed from the coding process. The qualitative data that was collected from the surveys was then compared with the qualitative data and any inconsistencies were addressed as they arose.

Using this method helped to reduce bias and emphasized the students' stories and opinions. Students were selected randomly; therefore, the students ranged in academic

levels and high-stakes testing experiences. Random sampling was important because previous research showed that lower performing students had more negative feelings towards high-stakes standardized tests. Including students from many academic levels allowed for the research to have a balanced student perspective. The validity was higher because of the use of random sampling which included students from many academic and socioeconomic levels.

The interviews consisted of follow-up questions that emerged from the individual's responses on the survey. If the student stated that they liked high-stakes testing, the interviewer tried to find out why. If the student stated that they did not like high-stakes testing, the interviewer tried to understand the reasoning behind the negative perspective. Students were also asked to recollect on a previous testing experience that they could remember. The researcher also worked to determine if an authority figure within the school contributed to the student's feelings towards the accountability standards. The interviewer's goal was to understand the reasoning behind why the students feel the way they do.

### **Procedure**

The first step in the research was to obtain consent from Loyola's Institutional Review Board (IRB). The IRB allowed the research to be exempt because the information being collected was not predicted to be harmful for the students. After the IRB approval was received from Loyola University Chicago, flyers for the online survey were distributed. The flyers were distributed at local parent and teacher events, local community organizations, as well as on parent and community online media.



The survey was available online for four weeks to allow for a wider and larger sample. The survey allowed for the students perspective to be vocalized without stresses from the researcher or unknown variables. The sample consisted of students that had taken a high-stakes test, as well as students that will be taking another high-stakes test in their current grade. The survey contained both opened ended as well as multiple choice questions.

After receiving the surveys, the subjects were selected. Ideally, the research hoped to have student participants from schools that were and were not on probation together with low-income and middle/high-income. Using perspectives from a variety of students allowed for a comparison of different socioeconomic statuses and performance ratings. The data was then analyzed before interviews and used to help structure the interview script.

Next, consent from all of the student's legal guardians and student assent (for 6th grade/age 12+ students as appropriate) was required. To assure consent from the start of the research, the guardian's year of birth was required to enter the survey. Additionally, the survey required electronic parent consent and student assent. It was understood that every participant's information and their responses were confidential. It was also suggested that their responses would possibly be used to determine the use of high-stakes standardized testing in future generations. All information about the study was provided to the participants and their guardians before the start of the survey. Providing the participants with the expected results did not alter the results because the research was to evaluate the student's feelings and perspectives of high-stakes standardized tests.

The random sample included students from high, low, and mid-ranged academic levels, and who held a variety of viewpoints about high-stakes testing. The random sample minimized the bias in the final research results. After the interviews were completed, a transcript for each interview was created. The transcripts were evaluated for repetitive themes, shared testing experiences, and similar trends in the schools execution and promotion of high-stakes testing.

### **Analysis**

The data was evaluated to see if the student's perspective of high-stakes testing was dependent on their performance on the tests as well as their school's performance. When the research was completed, qualitative data was largely incorporated to minimize altering the student voice. Humanizing the student's perspectives was the main goal of the research and student quotes were extensively used in the paper. Unchanged quotes illustrate patterns within the student responses and unique individual interpretations of the accountability standards and personal feelings toward the tests.

## CHAPTER SEVEN

### RESULTS

#### **Introduction**

The research consisted of an online survey and an in person interview. The survey was advertised to 5<sup>th</sup> and 6<sup>th</sup> grade students in Chicago and the Chicagoland areas. With the survey being conducted online, there were students that did not meet the research sample criteria that completed the survey. After the survey portion was closed, students that met the advertised criteria were selected to participate in the interview.

The survey section of the study was completed in April of 2014 and the interviews were conducted in May 2014. At the time of the interviews, Chicago Public School students had recently completed the ISAT and the NWEA, while students in the Chicago suburbs had only completed the ISAT. A total of five students were interviewed who ranged from borderline special education to gifted. Students were selected from several of their survey answers.

Out of the five students interviewed, two were from a Chicago suburb and three attended a Chicago Public School. Four out of five of the students were female and all classified themselves as Caucasian. Every student was asked a series of similar questions and, based on their responses, a variety of follow-up questions. Two out of three of the students stated on the survey that they liked the ISAT or NWEA MAP.

Overall, each student had an interesting perspective on testing. The majority of the students liked to know where they stood in terms of their learning, but did not like the high-stakes tied to testing. The students each had a different experience with their school environment with testing, receiving their results, and their perspective of how things are taught at their current and previous school.

### **Meet the Interviewees**

Dawn is a ten year old Caucasian female. At the time of the survey and interview, she was enrolled in the 5<sup>th</sup> grade at a Chicago Public School (CPS) classified as a regional gifted center. The school enrolled students from preschool to eighth grade and had less than 700 total students with Dawn's class having 20-29 students.

Dawn's school was a level 1 (level 1 being the highest) according to CPS's school rating and was above average on both the ISAT (CPS average) and NWEA (National average both Reading and Math). Her school had a mixed student demographic with 3% Asian, 37.9% Black, 30% Hispanic, 23.9% White and 5.2% Other. In addition, 55.7% of the students were low income, 10.6% were classified as special education, and 7.2% were labeled as limited English.

Lisa is a ten year old Caucasian female. At the time of the survey and interview, she was enrolled in the 5<sup>th</sup> grade at a Chicago Public School (CPS) classified as a neighborhood school. The school enrolled students from kindergarten to eighth grade and had less than 750 total students, with Lisa's class having 30-39 students.

Lisa's school was a level 1 according to CPS's school rating and was above average on both the ISAT (CPS average) and NWEA (National average both Reading and Math). The student demographic was .4% Asian, 52.5% Black, 5.8% Hispanic,

39.3% White and 2.1% Other. In addition, 28% of the students were low income, 13.5% were classified as special education, and .3% were labeled as limited English.

Mike is an eleven year old Caucasian male. At the time of the survey and interview, he was enrolled in the 5<sup>th</sup> grade at a Chicago Public School (CPS) classified as a neighborhood school. The school enrolled 850 students from kindergarten to eighth grade, with Mike's class having 30-39 students.

Mike's school was a level 2 according to CPS's school rating and was above average on both the ISAT (CPS average) and NWEA (National average both Reading and Math). The student demographic was 0% Asian, .7% Black, 98.9% Hispanic, and .4% White. In addition, 95.5% of the students were low income, 10% were classified as special education, and 29.6% were labeled as limited English.

Both Melissa and Sarah are twelve year old Caucasian females. At the time of the survey and interview, they both were enrolled in the 6<sup>th</sup> grade in an Illinois public school. The school district enrolled 1,050 students from sixth to eighth grade with Melissa and Sarah's class having 20-29 students.

Their school was 9 out of 10 (Above average) according to GreatSchools rating and was above the state average on the ISAT (both Reading and Math). The student demographic was 2% Asian or Asian Pacific Islander, 7% Black, 25% Hispanic, 63% White and 3% Two or more races. Information regarding income level, percentage of special education, and limited English students was not available.

### **Interviews**

At the time of the interviews (May 2014), all of the students had recently completed a standardized test. The Illinois Standard Achievement Test (ISAT),

completed with paper score cards, was one of the tests students reported recently completing. The ISAT was used as a high-stakes test in previous years in Chicago. However, the last year it was administered (2013--2014 school year), the test was low-stakes. In the interviews, some of the students were aware that the stakes had changed, while others were not.

Another standardized test that students had recently taken was the Northwest Education Association (NWEA), a computerized test. Just as the ISAT was in transition, so was the NWEA. The NWEA was becoming the new high-stakes test in Chicago. Students were aware that the transition of high-stakes from the ISAT to the NWEA was being implemented during the 2014-2015 school year. During the interviews, students voiced concerns of technical issues during the administration of the NWEA. During the interviews the ISAT and NWEA referred to as a high-stakes test, standardized test, "Big test" or discussed by test name (ISAT or NWEA).

The interviewees were selected at random and the interviews were conducted in public locations and in private residence. Only the researcher and student were present for the interviews. Interview lengths ranged from twenty minutes to an hour. The student's voice is essential to the research; therefore, some quotes may be elongated. During the interviews, many themes became constant with each student's perspective on high-stakes testing. The results below explored the student's perspective and organized them according to theme.

### **Negatives in Testing**

In the survey and during the interviews four out of five students mentioned negative aspects of high-stakes testing. The themes included, but are not limited to,

feeling: stressed, scared, anxious, and pressured. In addition, some students mentioned that they dislike the testing experience.

### **Feeling Pressured**

Lisa was very persistent about the pressures that are put on students when it comes to high-stakes testing. She first mentioned pressure on her survey. On an open-ended question, she wrote, “I do not like having these tests b/c teachers take periods away from their regularly scheduled school day to practice for them. Teachers put lots of pressure on students.” We continued the issue of pressure during the interview.

In the interview, Lisa stated, “[The teachers] usually put a lot of pressure on the kids and my teacher, you can tell she really thinks testing is stupid.” She added, “Teachers or counselors, they say things like, ‘Everything you’ve learned goes into this test. It matters about everything. Everything depends on you getting to the next grade.’ That’s pressure”!

She continued the discussion, “I think that, they do the MAP test to see how much you’ve grown and if you are ready to move on to the next grade. But, teachers give you a report card every quarter and they see how well you’ve done. That’s basically the same as these tests, the ISAT and the MAP or NWEA test. Only they apply so much more pressure”! Lisa disliked the pressure that high-stakes testing puts on her and her responses were consistent throughout the survey and interview processes.

Lisa used the word “pressured” several times when discussing her testing experience. No other student discussed the pressures at such length, but the theme appeared in different terminology.

Dawn stated on her survey that she liked high-stakes testing. Without the one-on-one interview, Dawn would have been categorized as not being pressured on high-stakes testing. Her comments during the interview proved differently:

I guess I get weirded out because, there are people all over tugging you in different directions. And they're saying, 'You should really work on this.' Then another person, 'You should really work on this.' Then another person, 'You should really work on this.' Some of these people, I don't even know. There was a random person on the intercom today in our school; came up saying, 'Do well on the MAP kids.' A random voice. I don't know who they are. I think he's part of the staff, but if he is, I never heard him. It kinda weirded me out.

Dawn stated that she receives high scores when she completes high-stakes tests, but she still feels pressured before the test.

Dawn also expressed that she feels pressured during testing due to distractions. She said, "I was worried that I was getting some of the questions wrong because I was distracted." Mike also discussed distractions and the pressure that occurs. Mike stated,

If you're tired while taking a test, well it wasn't a problem this year because they gave us mints to wake us up. If you're tired while taking a test, umm, you use 10% of your brain power every day. Not sleeping could bring that to a minimum of 1%. Which utterly sucks when you're taking a test, I mean. That one question, that one question you're not sure about.

Mike subtly discussed the pressures of testing as he explained his fabricated statistic.

Although he did not describe his concern for lack of sleep as "pressured", it shows that the slightest distraction can impact the student's thought process.

This was also apparent in Melissa's concerns about distractions during testing. She describes,

When you finish ISAT, [teachers] have you read. And some kids will flip their pages really loudly when you're still taking the test or they will tap their pencil. And I'm like really? Some teachers will like open and close the doors and it gets distracting 'cause they will make noises.



Why does a page turning frustrate Melissa so much? Melissa believes she is distracted “probably ‘cause I don’t want to take it, I just don’t want to do it.”

The pressure appears in Melissa’s and Sarah’s statements. Children usually want some sort of noise around them at all times. Sarah discusses similar distracts, “Some people tap their pencils, I don’t like that. Some people blow their nose half the time, I don’t like that.” She concluded, “I don’t like the extra noises.” Both Melissa and Sarah appear to feel a large amount of pressure during testing; the smallest distractions appear to create huge problems for otherwise active students.

### **Nervous and Anxious**

During the test, several of the students mentioned being nervous or anxious about the test and/or the testing experience. Sarah and Melissa attended the same school and had similar perspectives when it came to the tone of testing.

Melissa stated, “I was nervous, ‘cause I didn’t know how I was gonna do, or if I would forget some stuff. I was kinda just wanting to get it over with.” Because of Melissa’s comment about wanting to get the test over with, the researcher inquired as to whether or not she took the ISAT seriously. Melissa responded that she did take the ISAT seriously.

Sarah referred to the ISAT as the “Big test” and felt “Kinda nervous because [the teachers] didn’t teach us enough to like, answer the questions.” She added, “They didn’t go over it well enough.” Sarah displayed concerns about being prepared enough for the tests. Sarah also discussed her anxiety when the test begins, “right when I get the test, I don’t know what the questions are like. So I don’t know if they are hard or easy.” Both Sarah and Melissa were concerned about their performance on the tests, which therefore created an undesirable testing experience.

Dawn shared similar concerns about her performance on the tests because of the schools and programs that use the results for admittance. Dawn stated she gets “A little anxious, because that ISAT or MAP, whichever standardized test we’re talking about, kind of holds the key to me going to the next grade, slash going to my choice of high school, slash academic centers, slash going to the places I want to be.” She pointed out that she was “still a little nervous ‘cause it decides if I get into Lincoln Tech or Whitney Young, getting into academic centers.”

She expressed, “I like it when I get a good score and I wanna get a good score, but I’m scared that I won’t get a good score sometimes.” It is understandable that Sarah, Melissa, and Dawn are concerned about their performance on high-stakes test because a great deal is tied to the scores they receive.

### **Scared and Stressed**

Similarly to being nervous and anxious, being scared and stressed was also discussed during the interviews. As previously discussed Lisa feels unwanted pressure when it comes to high-stakes testing. Reflecting on her first experience taking the ISAT, she remembered, “In third grade, I guess, I think third grade is the first time they give you an ISAT test. That was most stressful for me. I guess that would be the worst ISAT experience.” The unknown is never easy, but the unknown tied to high-stakes is even worse.

Lisa also discussed the stress that the high-stakes tests bring into the classroom. On the survey, Lisa stated that she disliked high-stakes testing. During the interview, she discussed, “I don’t like the tests because there are all these things that my teacher is going

over that you've never heard before." In addition, "We ran out of time to go over material that I got wrong on the test in the Fall." Lisa's stress continued.

When discussing distractions during testing, Lisa indicated that she does not allow herself to get distracted during a high-stakes test. Lisa stated,

Sometimes it gets pretty stressful, so you just keep your mind on the test. But, sometimes you get really stressed out. In the lower grades when I was in there, I use to get really stressed out about tests. In ISATs especially because [the teachers are continuously] saying 'You're gonna need this on the ISAT, you better pay attention.' Or the technology teacher, she said, 'If you don't do so many hours on this website that were doing for the school, then you're going to fail it.'

The pressures from outsiders created a huge stress for her and she did not allow herself to get distracted during a test.

Lisa believed high-stakes testing created stressful learning and is not an accurate measurement of her performance. She stated, "It creates stress on the subjects you're being tested on and the subjects you're not being tested on. Because they matter too." Her honesty was shown when she revealed, "[when] you are just so stressed out, then you're like 'I don't care anymore, I'm just going to put down any answer that I want.'" So, it doesn't really judge your growth accurately, I think, because you can just use your letter grades, right?" Lisa's concerns about evaluations are important and will be discussed later in further detail.

In addition to Lisa, Melissa mentioned she gets "Kinda scared 'cause I don't want to mess up or anything on it. I know everyone says you shouldn't get worried about it, but kids get worried about it 'cause it's like a big test." Melissa expressed that she disliked high-stakes testing because "I guess it's like you have a test in each class and it's kinda overwhelming." She elaborated, "It's kinda overwhelming how you have it in

Math and Reading because those classes are next to each other. So, you don't really get a break or anything. We have sections that we do; we'll do session one with the teacher in Math and then session one with Reading. Then the next day we do session two" Melissa does not like the ISAT because the school "could use the time to go over things that kids didn't understand in the year." Sarah also mentioned in passing that she gets "bored because I don't like sitting in a chair for a long time." Their concerns over lengthy testing sessions and use of classroom time are reasonable for active children.

Dawn takes it one step further than just mentioning that stress is an issue in testing. She stated on the survey that she would perform better in a little to no stress environment. During the interview she brought up that "There are all these people saying 'Oh my God, the ISAT, they are so IMPORTANT.' Guys, guys, guys, you're stressing me out, you're stressing me out, I won't be able to concentrate." As the students mentioned, stress is being incorporated in many ways which increase the fear that they will not perform successfully on the tests.

Mike is on the other end of the spectrum. He stated he was neutral about testing on his survey and confirmed his selection during the interview. He stated, "I'm just like that one normal kid, that one normal kid that doesn't feel weird about the test or anything." When asked to define "normal" Mike responded, "Like someone who does get emotions, but not too far. Not like, kids who are just super smart aleck that are so scared of touching other people. Which could be very weird, it's just stupid." When asked if he knew someone like his description of a smart aleck, he stated he did not know anyone like that. Mike appeared to want to be described as "normal" but still intelligent student who was not affected by high-stakes testing.

### **I Like High-Stakes Tests**

Although most of the interviewees mentioned negative issues with high-stakes testing, confidence during testing was also discussed. Both Mike and Dawn stated on their survey that they liked the ISAT or NWEA. During the interview, they both brought up current and past testing experiences where they were confident or performed very well. Mike reflected on his most recent testing experience, “I felt pretty well, I felt confident and that confidence helped me in my favor. It basically gave me an extreme rise in my score.”

Dawn also discussed the confidence in her testing ability. She stated, “This Reading MAP test, I came in confident, I took the test confident, and I finished confident. Because when you see your scores, I got a good one. I’m happy; it was in like 250 or something.” Both Mike and Dawn are confident when they are completing a high-stakes test and they believe that their confidence influences their overall performance and score.

Interestingly, Mike indicated that he does not get nervous or stressed about high-stakes testing. He believes that confidence is important in taking a test and he stated, “I think of it as just a normal test (normal test = classroom test), that’s what I think of it as.” But he unknowingly discussed his thoughts about high-stakes testing (without additional information from researcher) and how high-stakes testing influenced him even though he scores high on the tests. Mike rambled,

Some kids forget to eat at home and they don’t like the school breakfast, so they don’t take it. Or they just got into the routine, were aww, they don’t take one of the breakfasts. Like me for example, I sometimes forget to eat [at home] and don’t take a breakfast. And that brings your brain power to an about, I say, 6 or 7%.

When Mike was confronted about him not talking about himself, he stated he was concerned about other kids,

Because they are, most of my kids in my class are my friends and if they fail on the ISAT, it will mean that they will probably don't go up to the next grade or go to summer school. Making their summer not fun and not being able to do stuff with me when they have the time. So, it's just boring over the summer and I'm just sitting at home.

Although Mike was not stressed when he took a high-stakes test, the stress of high-stakes testing still affected him because if one of his friends failed the test, then his summer would not be as enjoyable.

Dawn also changed her opinion about high-stakes testing. Dawn originally stated, "I like the ISAT and the NWEA because it is a nice test just to see what you've learned." The researcher explained the difference between a standardized test and a high-stakes test. When Dawn realized that high-stakes testing determines if a child passes a grade, she stated, "I think that that's not good. It's better to see what you've learned and go off that than just, if you fail this test you don't get to go to the next grade. What if it is that the questions just [aren't] right for the kid not that the kid isn't right for the test." Overall, Mike and Dawn were in favor of standardized testing but did not like that high-stakes could negatively affect other students' progress.

### **Attitude Towards Testing vs. Performance**

Out of the five students interviewed, Melissa was the only student that said she did not perform as well as she would have liked to on the ISAT. She was also the only one who had attended summer school, although she was unsure why. She pondered, "I didn't know the reason why I went, I just know I had bad grades. I think it was because of Math. I think." Due to the benchmark requirements within Chicago, it is likely that

she had to attend summer school because of her performance on the ISAT in the third grade.

Melissa recalled her summer school experience. “We had to watch the kids, when we went to recess; we got to watch the kids swim. It was horrible.” She continued, “I just really hope I don’t go this year,” or any year because it was kinda horrible.” The researcher asked if she believes that the experience affects her today? She responded, “I don’t think it still affects me.” However, Melissa currently is involved in an Enrichment Program. Melissa stated, “I just go for E period because of ISATs last year. I think they were kinda low in Math.” She explains how she is involved in E period, “It’s like that last 25 minutes of the class you go with the teacher. You’re in a class of like fifteen kids and she’ll come around and you get help with like homework.” Melissa is unaware that her low performance on the ISAT has resulted in her school receiving more individual attention.

Melissa openly confessed that she disliked the tests. The predicted outcome of the research was students who performed poorly on high-stakes tests, disliked tests; while students who received high scores liked the tests. If the surveys were the only information evaluated in this research, the results would have been untruthful. However, the interviews revealed a deeper understanding of the student’s views on high-stakes tests.

All other interviewees stated that they received average or above average scores on the ISAT and NWEA. They all stated that they liked receiving high scores on the tests. Lisa declared, “I usually score really high on the tests.” She elaborated by saying, “I like when I get done with a section and I feel like I did really good. Like half the

questions, I answered correctly.” Overall, when she feels like she did good, she feels good about herself.

Lisa also discussed her dislikes about high-stakes testing. She believes, Sometimes they just try to confuse you. Like, when they just ask you a simple math problem, or multiplication problem. This is what they did to me last year, instead of just using the multiplication sign; they did the dot in the middle. Just to make sure you knew what that meant. And teachers always say they’ll give you questions you don’t know about like, they say, ‘using your Latin knowledge, what does this mean?’ I don’t have any Latin knowledge. Hahaha

Lisa dislikes the testing experience due to the pressure bad stress associated with the high-stakes. She admits there are times when “I know I only did it just a few weeks ago and this was just my own brain fart, but I just forgot how to do it. And I was like, oh no, what do I do next?” The researcher inquired, “How does that make you feel when you aren’t sure what you are doing and you are being timed?” Lisa responded, “You know what, I feel like being timed is a little unfair because I am usually one of the last people to finish.” The confusion and tension added to her distaste of testing.

Similarly, Sarah has received average scores on high-stakes tests but dislikes them. Sarah believes the tests are used to determine “what [students] know for the next grade.” Her scores are mid-level and she has mixed feelings toward the material on the tests. “The good [questions] are like the multiple-choice and how they explain it. I don’t like the charts and stuff because they are hard.” She expressed she gets sad “when I don’t really get a lot of the questions right.” She provided examples of feeling discouraged: “When I do an equation and I don’t get an answer that are in the multiple choice”; “I don’t like reading ‘cause you have to memorize a lot in the passage, but you can go back, but it’s like really long paragraphs so it’s really hard.” All in all, Sarah shows the



characteristics of the research prediction. When she is discouraged, she dislikes testing. The testing is tolerable when she is confident in her performance.

On Mike's survey, he stated that he liked high stakes testing. When thinking about distractions during tests, he wondered into the thought that his friends may not pass the high-stakes tests. When the researcher informed him that on his survey he responded as liking high-stakes testing, he stressed, "That was at the time, and jeez, the NWEA, they push you too much on it and I didn't realize that until I actually took the test. Can I change my answer there? They put too much emphasis [on testing]." The researcher approved his choice to change his response.

Melissa had an open hatred for high-stakes testing. She did not state it was because of her experience in summer school or being singled out because of her low ISAT scores. She explained her dislike towards testing was because of the lengthy and boring process. Although, Lisa receives high scores on high-stakes testing, she also dislikes the process and added stress. Sarah has varied feelings during the test but ultimately dislikes the experience. As discussed earlier, Dawn and Mike stated that they like high-stakes tests. However, during the interview it was discovered that they do not like how the tests are used as an end all for promotions. They should be used as an evaluation. After analyzing the data, the student's performance does not reflect their opinion towards high-stakes testing.

### **School and Curriculum**

One of the arguments against testing is that schools are so obsessed with high scores that they are emphasizing test prep over rich curriculum. This section will discuss

the different issues that were suspected to impact the classroom and curriculum, as well as issues presented by the students.

### **Teacher Performance**

The research was expected to find that administrators and teachers have altered curriculum to teach to the test. Although test prep was discussed, it was not as dominant as expected.

Lisa explains the extensive review and introduction of new material that is discussed weeks before a high-stakes test. She explains her teachers deviate from standard lessons before a high-stakes test,

She'll go on about how they're gonna ask you questions about this, so in a couple weeks we get ready and we learn about different types of poetry. She goes over propaganda and types of questions, you know important stuff you've never heard about before. They are gonna ask you questions about that.

Lisa also brings up a discussion she had with her mom about the change in the curriculum and the increased test prep.

Because of the pressure and the way that they take away what you're learning in class. I think it is true a lot. I believe that and my mom believes it too. And then in Math, she takes time away from the chapter we're in in the book, just to go over how to carry out an extended response. So basically in the last weeks they go over everything that you've learned. They make sure that you remember what we went over in the beginning of the year.

According to Lisa, her mom had discussed the changes in the curriculum with her and she can see that test prep has changed the way she learns in school.

Mike brings up a wide variety of ways that the curriculum is influenced and taught at his school. He discusses his Math class,

There are a lot of challenges that my Math teacher puts on the board. Like we had to find out what an exponent of one was, which geez, that's actually a little bit hard. 'Cause, I don't think we're able to use the internet abusively. Umm... because then the internet would just make it too easy.

He then discusses the recent changes in his Reading class. “Recently, we just did something like a book club. I think doing that on my own could show my growth in reading.” Mike appeared to enjoy the current set of courses and did not want to change or add anything to his school’s curriculum.

Melissa brought up a more disheartening perspective of the curriculum and her teachers. She believed, “[Teachers] don’t do anything fun with [the curriculum], they lose the kids. They lose the kids when they teach ‘cause it’s not fun learning. Kids don’t like going to school, so they should make it more funner with learning.” Troubled by her views toward learning, the researcher asked, “Was there a time when learning was fun?” Melissa answered, “Sometimes in Math class we do like stuff with hands on activities that are fun.” The Math class example showed that Melissa did have a small amount of curiosity to learn.

The researcher followed up with, “What is your favorite subject?” Melissa replied, “Probably Social Studies but that’s it. They don’t do ISATs or anything for social studies. But we have A day and B day. A day I have Social Studies and B day I have Science.” She mentioned that she would like to have Social Studies every day.

Continuing the conversation on the curriculum, Melissa said,

Sometimes I think she goes a little too fast with some of the things she teaches us. Like, we’ll get packets in school. And she’ll teach us, like when we’re in Math, she’ll teach us some stuff but, then she goes into a different topic and it’s like we wanna learn all about that topic. But then she’ll go into kinda a different. I don’t really get it ‘cause she teaches two different parts. But, [if] I don’t understand the first part, how can I understand the second part?

Melissa's concerns about the pace are justifiable. However, it is unknown if the teacher is quickly moving through the material because she thinks the students understand it, or because she is cramming information for a standardized test.

The conversation continued. The researcher asked Melissa about her relationship with her teachers. Melissa unenthusiastically stated, "I don't really like my teachers, they're nice, I just don't really like them." When encouraged to provide a reason for disliking her teachers, she stated, "Sometimes out of the classroom, you wish you could ask a question, sometimes they're walking." She felt that the teachers did not have time for her. In the end, it appeared that Melissa did have a joy for learning, but the current structure within her school discouraged her from wanting to learn.

At the same school, Sarah also disclosed concerns about the curriculum. She believed the curriculum was confusing "because of the way they put it. Some people follow along ok but some don't understand the vocabulary either." She also stated, "In some classes like Social Studies and Science, they make us read from a book and like in Math they do it with us, they teach it." She vocalized, "I like to do stuff on paper," which displays that she is a visual learner, and some of her classes do not use a variety of teaching styles to meet the students' needs.

She also revealed her concern about the use of vocabulary. "They can sometimes improve because they sometimes put it in like really big words." Sarah also brought up the issues of behavior and the teacher's zero tolerance outlook. She said that her teachers yell at her for no reason. When asked for an example, she stated, "I would ask someone for a pencil, and she would start yelling." From Sarah's description, it appears that her

school still uses the traditional style (silence and lecture) of teaching and has not incorporated a progressive style (student directed and unrestricted movement).

The students' view towards their teachers is important to the research because it was considered to be a contributor to the student's desire to learn. It was thought that if the student had an active and fun learning environment that they would enjoy learning. On the other hand, if they had a negative classroom culture they would dislike learning. Out of the five students that were interviewed, three of the students praised their teachers, one disliked how her teacher disciplined, and the last student did not like her teachers. Teacher performance did not have a substantial effect on the students' perspective toward tests.

### **Promoting Testing**

As previously discussed, Lisa mentioned the faculty involvement in promoting tests at her school.

Usually they have teachers come and talk about it. Like last year, what they did was for the ISAT, they had you take an index card and write something that made you really happy. So when you're really stressed, you can just look at it and be like, okay, I can do this. Then they had teachers give you a pep talk. Usually teachers you never seen before or the counselor comes into your class. Last year for the ISATs, the counselor she came in and gave us a talk about the ISATs and someone else came in from the school; and she doesn't work there as a teacher. She just came in and made us do the index card thing and draw a picture about something you really like or your happy place.

The pep talks and "happy place" index cards are tactics used to convince the students that high-stakes testing is not something to stress about. Lisa lingers on the conversations and promotions in school and recalls,

I think after the test what [the principal] does is makes an announcement 'we've had so many good test scores on the MAP test or ISAT test and it's finally over.' That's what they did. At the end of the ISATs, usually the principal does make an announcement over the speakers, 'The whole school has done a fabulous job

with the ISATs' or something like that. It makes you feel good because the principal knows that you did good.

As a result of the principal making a general announcement on the intercom, Lisa believed that the principal knew her individual ISAT score.

Lisa discussed websites that provide test prep activities outside of school. She stated, "They give you a website, like Study Island and you have to do ten hours every quarter. Which nobody really does. But [the Computer teacher] says, 'if you don't do it then you aren't going to make it into the other grade.' That's pretty bad." When asked how she felt about that teacher's statement, she emphasized,

It makes my mom really mad. I was kinda like okay then. And it was, it probably got the kids really nervous. Like 'Oh my gosh I have to do it.' I think that's what she was going for, trying to get more kids to actually do it instead of just being like 'oh, well it doesn't matter, it's just a stupid website. This is serious.' It's really boring, so kids don't like it.

However, the emphasis on creating a no stress environment with the pep talks and "happy place" index cards. Lisa expressed the attempt was diminished with a single comment from the computer teacher.

Dawn discussed, what she thought, was both a predicted and surprising test promotion at her Chicago school. The school displays "posters all over telling you. 'Get a good sleep, eat well, study hard.' They were posters, they were student made." In addition,

They have an entire billboard up with how to prepare for a test and it's kinda motivational. But all the things telling you to do stuff and eventually after a while, it's kind of creepy. I was tossing and turning for a good part of the night. The signs, they're pressuring me!

While the use of posters might be enough at some schools, Dawn's school goes a step further,

Well at [school name] the people who have the most growth shown in their MAP test are invited to a pizza party with the principal. I personally think that, no offense, it's idiotic. People in my class are intentionally getting bad scores on their beginning Math test so that they can get the best growth score and get into the pizza party.

Dawn has never been to the pizza party, and does not want to go because her dad makes better pizza at home. "I don't need your evil artificial pizza, I have my own pizza."

Dawn is not fooled by the use of pizza to promote testing and dislikes that her classmates purposely do badly to go to a pizza party.

Mike also introduced an unexpected situation. He explained that teachers set goals for the students, but the students are not consulted when the goals for test scores are determined. He stated, "I missed both my goals this year by one point." Mike's teacher sets his goals in both Reading and Math.

In my opinion it's not that fair. Because the fact that they set up a goal for you and you don't think you're able to do that goal, making a little bit less confident, is, confidence is just a good thing on tests and stuff so by them giving you a high score to reach and you not thinking that you can reach that score, lowering confidence, which would not be a good thing for a test.

Student's opinions are undervalued when they are not able to have input on goals that they are expected to obtain.

He continued by saying, "But I still didn't reach my goal. Mostly I didn't reach my goal because in the Math NWEA because they just changed fraction number and operation, and number and operations they combined those into one category." The unexpected change on the test influences Mike's expected performance on the exam.

Just as Lisa's school tried to create a positive attitude towards the ISAT, so does Mike's school. He discussed his school's way of promoting the ISAT. He explains, "Prep rally for the ISAT mostly. Sometimes they do parody of songs. In my memory of

last year, they did a parody of Black Eye Peas.” Melissa shared a similar experience. “In my old school [an Illinois public school], we made songs about it, and they would give us packets. But this year, they didn’t really do anything.” Similar to Mike’s current school, Sarah’s old school also made kid songs to promote high-stakes testing. Throughout Illinois, schools are promoting high-stakes testing with songs and pep rallies to create a “fun” testing environment.

### **Online Test Prep**

A completely unexpected area of student concern arose with both Dawn and Lisa. Lisa discussed in preparation for the ISAT, students were required to complete a set number of hours on certain websites. Dawn also discussed her concern with kid friendly websites advertising test preparation websites.

Unexpectedly, when discussing if the faculty have changed or influenced her perspective about testing, Dawn stated,

It’s not the teachers and the faculty, it’s websites. There was an ad I saw once Tests are important, study, learn, succeed.’ Everything was adorable and then this blue gray message. Very solemn! In a batman voice it was like ‘study, learn, succeed.’

Dawn’s perspective was interesting because she noticed such a depressing difference from the joyful website to the testing advertisement.

Lisa also brought up a concern about the use of websites to prepare for testing.

Lisa explains,

Seeing commercials about the ABC Mouse thing, that’s basically getting kids ready for standardized testing, isn’t it? It is a commercial I see all the time and it is for little kindergarteners and I think it’s really funny that they use websites like that for kindergarteners. It’s basically just Study Island for really little kids. Study Island is basically the MAP test only it doesn’t really matter a lot because they give you the same questions only you have the option to choose game mode. So if you get the question right, it’s like, ok now I get to play a little game now. I



guess it's just trying to get kids used to testing. Invented about a year ago and it has parents on the commercial saying how it got kids higher grades in school. Testing with games for little kids and they think it's so much fun and it gets them higher grades. So when they actually take the test, they're like, 'oh it's just like that website.'

Lisa said that her mother brought up the idea that the website was used for standardized testing I thought that was really funny, what they were using it for. It sounded like the K-12 thing at the time.

Lisa and her mother had discussed the test prep websites being disguised as "fun." Her concern continued and started reflecting about her own experience with websites in school. She recalled,

When I was in preschool and it was really fun. 'Cause it was just games, one game was a mouse he was writing a letter and you got to choose who you were writing the letter to. It had like four actions. Who are you writing to? 'Grandma, grandpa, mom, your best' It was mostly games, but there was this separate thing when you got to choose a letter of the alphabet and it gave you a few words with the same letter. It was mostly games but they through a few things in like letter writing and you got to choose what you were writing. At the end they showed you what you wrote.

When asked to provide the name of the website, Lisa was only able to recall the activities she completed.

As mentioned in the section *Promoting Testing*, Lisa's computer teacher had scared the students by telling them that if they did not complete the required hours on the website, it was possible to not excel to the next grade level. The students' believed, the use of "fun" websites to promote student learning is a great use of technology. However, using websites to increase standardized testing is a concern for the researcher and the students.

## Test Results

From previous research, it was determined that a student's view towards high-stakes tests was dependent on their performance. Some of the students divulged that they do not receive their scores from standardized tests while others consistently received their scores.

During Dawn's interview, she declared that she consistently receives her test scores she has to wait for the results.

At the end of every test, every standardized test, when we get our test scores back, there's always, always, I don't know if it is on accidental or purpose, boop, a little lag. There's lag, there's lag, I don't know if it's gonna be good or bad. I get really, really, really, really, really, really, really, really time five hundred nervous"

The researcher questioned, "How do you think you'd feel about yourself if you got a low score?" Dawn explained,

I don't think it will change how I feel different about myself. I think it would change about how I feel about the subject. 'Cause sometimes [she thinks], Aww I got this down, low score, well geez, I need to practice more. I would study more and hope to become better in the subject. So no, it wouldn't change how I feel about myself, it would change how I feel about what I need to do.

Dawn's anxiety when waiting for her test scores seems to be a common emotion.

However, would Dawn's perspective of herself be untouched if she scored low on a standardized test? She believes that her strategy to study will leave her self-esteem unharmed. This is important because it shows that even with a low test score; Dawn will continue to have a positive attitude towards herself, her school, the tests, and learning.

Just like Dawn, Lisa is a CPS student. But Lisa's experiences with receiving her test results were different.

I have never received my ISAT scores. You just assume you did well because usually you try to stay pretty confident during it. The ISAT really isn't a hard

test. The MAP test gets harder because they ask you about things that you never heard about and you're timed on it.

She continues, "You do receive your MAP score back, 'cause when you finish, you get that chart that shows your score." In addition to past experiences, when Lisa received her scores, she brought up a horrible foresight. She stated,

You don't know what the high or the low score is. I heard you can't get anything higher than a 300, but I don't even know if that's right because I just heard it from some boy. The schools do not communicate what the scores really mean, but if you get over a 200, you know you've done a really good job.

How does she know that a 200 is a really good score if the school has not provided her with the testing rubric?

Melissa and Sarah, who attend a suburban school near Chicago, also remarked that they did not receive their high-stakes test scores. Melissa mentioned that, at her current school, she has not received her scores back for the ISAT or any other test. However, when she attended a CPS school, she thinks her mom received the scores. Correspondingly, Sarah does not remember getting her scores back at her school. She commented that without the scores, the tests do not really mean anything to her.

The above views provided by Lisa, Melissa, and Sarah are unfortunate. If what the students said is true; then students are expected to complete high-stakes tests without being provided vital testing information. Students should be able to know and understand their standardized test scores. The schools are to blame for this oversight and they need to work to improve the student's understanding of the test scoring system.

Mike, a CPS student, stated,

Last year my test scores were pretty good. I felt really confident about this year. Though my test scores may have just raised slightly. Or they could have gone down. I think am still going to be confident about them.

Mike will most likely receive his result at the end of the school year. When asked if he likes high-stakes testing, he responded,

Yeah! I mean, a little bit I didn't pay attention and just a little bit and it really did not help me that much. My score is just about the same as last year, it is a little bit higher, I hope, I don't know, I don't think I've received the results yet.

It appears that, although Mike has not received his test scores back, he has mixed emotions about his results. Mike's emotions emphasize that Lisa's, Melissa's, and Sarah's schools should not withhold test information. Although Mike stated he thinks his scores went up or stayed the same, he will never know unless he is given his score report.

### **Accurate Assessments**

In both the survey and the interview, students were asked to indicate if they felt the high-stakes tests accurately assessed their academic ability. Those selected to be interviewed were asked to provide an alternative way of assessing academic growth. The responses varied from students who did not believe any changes should be made, to the opposite spectrum of a complete overhaul of the current system.

As previously discussed, Dawn disliked that high-stakes testing was designed without the student's individualized academics being considered. Dawn discussed her ideas on a how to improve the current academic assessments, as well as what she feels is wrong with the current process.

I think that it would be cool if this could happen. I'm not sure if they have something like this, it would be like, you plug in what [the student is] doing good on and what [the student is] not doing so good on. And it gives you a test that could help with what you're not doing good on and show you how good you are at on what you're doing good on.

Dawn expressed that she would like a more individualized assessment system.

The follow-up question to her response was, “what if a test indicated that you made an incorrect response, and then showed how to correctly answer the question later in the test?” Thus increasing the student’s chance of getting the answer right if a similar question appeared on the test? She responded,

I think that would help if the questions were different. ‘Cause on some of these tests, I remember getting the exact same question sequence on the last test. I’m like, this question, I know this question. And I think it’s good if the questions change you’d be able to see what you got wrong and be able to get problems like that right with a different problem. But, if it’s just the same problem in different orders, it’s a little like cheating. The people with the best memory get the best scores.

Her feelings about the students with the best memories getting the best scores, was derived from her experiences with her peers. “I went to [the other students] after the test and asked how it went. They said, ‘ah, I got everything right, I remembered these things from the last time.’ She added, “A lot of the questions are the same, and it just doesn’t make sense.”

Dawn shared her thoughts on the current and potential academic assessments, I feel that [the ISAT is] a nice test of skill, it shows what level you’re on. But, it’s not that good for actually getting people into the next grade or something. If you don’t know this, then you don’t get into the next grade. Again, if a test had been personalized for each student in the class, it would probably make the response a lot better to the test in general. It would give, umm you could see what you’re good at. You can see what you needed to work on, and then you could work on those things without having to repeat the grade.

She did not believe that repeating a grade because of a test score was the answer to evaluating students. Dawn believes that high-stakes tests should not consist of similar questions, and should be individualized for the student’s academic level to include questions that student is familiar with, as well as questions they had difficulty with in the past.

Consistently, Lisa was unhappy with the current use of high-stakes testing and the additional stress and pressure created by the testing environment. She contemplated,

I thought if you get A's in Reading and you get A's in Math, then that shows how well you're doing. Since you are getting A's and you really good at it, shouldn't you be able to also be ready for next year? Because I'm sure that your scores, whether you get a really high grade in school, like an A+ in Reading then you are probably going to get a high grade on the test too. If you get an A, then your teacher knows that you are really good at Reading. Then you are ready for the 6<sup>th</sup> grade. You're probably really ready because, they just give you a lot of questions on the test that you've never heard before. And that's why my Homeroom teacher just goes over things just in the weeks before the test [her teacher says,] 'they're gonna ask you questions like this. You've never heard this before, so let's go over it now.'

Ultimately, Lisa believes, "they should really just rely on letter grades because I think if you're getting bad grades, then you might want to work on something more. You might want to go to summer school or get a tutor. If you get good grades, you can handle the next grade." Lisa struggled with why high-stakes tests were more significant in determining if a student passed a grade than the letter grades from the teacher.

As stated previously, Mike believes high-stakes test judge him accurately. He stated on the survey that he liked high-stakes tests "because it gives me a chance to learn from my mistakes and other things." He decided during the interview that he wanted to scratch his open-ended comment along with change his multiple choice answer. Mike did not believe that changes were needed to the current academic assessments, even though he had briefly discussed that the tests did not evaluate students on all subjects learned in school.

Later in the interview, he stated, "I like learning from my mistakes, I just think that it could be done differently." When asked again about changes to the evaluation system, he believed that hands-on learning and class participation could be used to

evaluate students. He explained, “The little things I mentioned for Science, Reading and Math. They could, as the grades go on, they’ll evolve it a little bit more.” Unsure on his standpoint, the researcher asked, “Do you think the teacher should determine if you pass a grade?” He responded, “Yes.”

The researcher then gave an example of the use of portfolios in other schools and asked if they were to be used in his school, who would evaluate them (current teacher, another teacher within the school, or the principal)? He answered, “Well, maybe the teacher that has the most experience would grade it. Well, I should say teachers because that would be a lot of work for one teacher, jeez.” In the end, Mike appeared to believe that students should be evaluated by their performance outside of testing by an experienced teacher.

Both Melissa’s and Sarah’s responses were more direct and to the point. Melissa stated, “Tests, teacher tests.” When asked why, said believed teacher tests were better she said because there was less material to being covered in a single test. Sarah believed students should be evaluated through presentations and the student’s personality. Sarah’s perspective was interesting because the history of education discusses that early assessments consisted of student performances, which encompassed the student’s personality.

### **Conclusion**

All of the students brought up areas of concern when it came to high stakes. Dawn felt that the testing environment was stressful. Although she likes showing others how well she performs on tests. She does not believe that they are successful when high-stakes are tied to the results. Lisa did not like that the tests created such a stressful

environment, and did not like that so many websites were trying to promote standardized testing. Although she performs well on the tests, she does not believe that the testing environment is healthy for learning.

In addition, Mike feels confident in his testing ability and feels confidence is an important factor in the testing environment. He originally stated on his survey that he liked high-stakes tests, but during the interview, he reconsidered the high-stakes that are tied to test results, he retracted his original perspective. Melissa, on the other hand and though did not feel that standardized testing was productive. She did not like the tests and felt school was not fun. Since she was one of the two students in the research who was not labeled as gifted, and had attended summer school possibly due to low grades, it is understandable that she had a negative attitude toward, as she believes, “boring” curriculum and the tests that go along with them. Finally, Sarah who was labeled gifted had mixed feelings about school and high-stakes testing. Her concerns seemed to lean more towards the meaning of learning and the overall quality of the school experience.



## CHAPTER EIGHT

### CONCLUSION

Do high-stakes accountability and/or curriculum changes (due to testing) influence students' perspective of themselves and/or their perspective about their school or community?

As originally hypothesized the students' perspective about high-stakes tests would have been aligned with "I Like to Read, but I Know I'm Not Good at It". Dutro and Selland concluded that students who performed well on high-stakes reading tests thought that they were good readers. On the other hand, students who performed poorly did not believe that they were good readers.

Using Dutro and Selland's (2012) research the following was expected:

- Students who scored "proficient" or above on high-stakes tests, would express positive feelings toward the tests.
- Students with poor scores would be more likely to have negative feelings toward the tests and their learning ability.
- Students would believe that test scores determined the quality of their school and their teachers.
- Educational policy would drive classroom instruction, and teachers would teach to the test against their better judgments.

- Test prep would constrain teachers and lower student's opportunities to improve learning.
- Students would describe the testing environment as anxious and boring.
- Students would know that their test scores impacted them and their school.
- Concerns about retention and summer school would arise.
- School staff would praise and reward high scores of students.

Dutro and Selland's sample contained 33 third graders from an urban elementary school.

This research contained five students from both urban and suburban schools. The subjects in this research also varied in age from ten to eleven.

The results of this research are as follows:

- Students who stated they received "proficient" or above average had mixed emotions; they were both confident and anxious about tests and the tests and the environment.
- The one student that expressed having trouble with high-stakes tests did not like the tests or school.
- Four of the students believed they went to a good school and that their school's overall performance on assessments was good. One student had negative concerns about his school. He thought the scores he saw online meant his school was underperforming. Therefore, the scores of his school proved his teachers were underperforming and, as a consequence, his school would close.
- Test preparation was mentioned, but in the student's opinions, was not seen as a main concern for students. Three of the students believed that they needed more test preparation before the ISAT or NWEA.

- Students did describe the testing environment as anxious and boring. In addition, two students described themselves as excited and confident before and during the tests; but were anxious while waiting for their results.
- Concerns about retention and summer school appeared in a majority of the student responses. Interestingly, concerns for other students appeared more often than concerns about themselves.
- Test promotions in the school did result in students with the highest scores or highest growth being rewarded.

This research had similar trends, however did not mirror the results of "I Like to Read, but I Know I'm Not Good at It" in full. The similarities that arose did influence the overall group's perspectives of high-stakes tests. Unlike Dutro and Selland's results, all of the students in this research disliked the high-stakes tied to tests.

After interviewing students from different schools and different academic levels, it is determined that the opinions of the five students are alike. They do not like the high-stakes attached to standardized tests, or the stressful testing environment. A larger sample size is needed to have an enhanced understanding of the student perspective. Due to the small sample size, the research was unsuccessful in providing a more concrete student perspective. It is suggested that more research is completed to better understand how students feel about high-stakes tests to develop a more student oriented curriculum.

## REFERENCE LIST

- Allen, T. E. (1958). Unintelligence Tests. *The Clearing House*, 33(3), 131-135.
- Athanases, S. Z. (2003). Thematic Study of Literature: Middle School Teachers, Professional Development, and Educational Reform. *English Education*, 35(2), 107-121.
- Bader, B., Horman, J., & Lapointe, C. (2010). Fostering Community and Civic Engagement in Low-Income Multicultural Schools Through Transformative Leadership. *Exceptionality Education International*, 20(2), 25-37.
- Beymer, L. (1966). The Pros and Cons of the National Assessment Project. *The Clearing House*, 40(9), 540-543.
- Chicago Public Schools. (2007, September 26). *Chicago Public Schools Policy Manual*. Retrieved January 11, 2014, from Chicago Public Schools: <http://cps.edu/Spotlight/Documents/Board%20policy.pdf>
- Daniels, V. (1964). Concerning the Validity of Standardized Tests. *The Clearing House*, 39(1), 12-14.
- Diamond, J. B., & Spillane, J. P. (2004). High-Stakes Accountability in Urban Elementary Schools: Challenging or Reproducing Inequality? *Teachers College Record*, 106(6), 1145-1176.
- Dutro, E., & Selland, M. (2012). "I Like to Read, but I Know I'm Not Good at It": Children's Perspectives on High-Stakes Testing in a High-Poverty School. *Curriculum Inquiry*, 340-367.
- Fredrickson, R. H., & Marchie, H. E. (1966). Teachers View Test Results. *The Clearing House*, 40(6), 357-358.
- Freire, P. (1970). *Pedagogy of the Oppressed*. New York: Continuum International Publishing Group.
- Holme, J. J., & Heilig, J. V. (2012). High-Stakes Decisions: The Legal Landscape of High School Exit Exams and the Implications for Schools and Leaders. *Journal of School Leadership*, 22, 1177-1197.

- Hong, W.-P., & Youngs, P. (2008). Does High-Stakes Testing Increase Cultural Capital among Low-Income and Racial Minority Students? *Education Policy Analysis Archives*, 16(6), 1-18.
- Jacob, B. A. (2002). Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools. *NBER Working Paper Series*, 1-71.
- Jeon, H.-J., Langill, C. C., Peterson, C. A., Luze, G. J., Carta, J. J., & Atwater, J. B. (2010). Children's Individual Experiences in Early Care and Education: Relations With Overall Classroom Quality and Children's School Readiness. *Early Education & Development*, 21(6), 912-939.
- Kelly, E. T. (1965). Measurement Masters Evaluation. *The Clearing House*, 39(9), 546.
- Kohn, A. (2005). Burnt at the High Stakes. *Journal of Teacher Education*, 51(4), 315-327.
- Lemann, N. (1999). *The Big Test: The Secret History of the American Meritocracy*. New York: Farrar, Straus and Giroux.
- Lucht, G. (1960). Testing - A 4-Stage Rocket. *The Clearing House*, 35(2), 74-76.
- Peters, H. (1959). Some Key Sources of Error in Test Administration. *The Clearing House*, 34(3), 161-164.
- Peters, H. J. (1966). Testing or Testomania. *The Clearing House*, 41(4), 223-226.
- Reese, W. J. (2013). *Testing wars in the public schools : a forgotten history*. Cambridge, Mass: Harvard University Press.
- Ryan, A. M. (2011, June). From child study to efficiency: district administrators and the use of testing in the Chicago public schools, 1899 to 1928. *Routledge Taylor & Francis Group*, 47(3), 341-354.
- Sandy, J., & Duncan, K. (2010). Examining the achievement test score gap between urban and suburban students. *Education Economics*, 18(3), 297-315.
- Scales, E. (1964). Measured: What Is the Standard? *The Clearing House*, 39(4), 195-202.
- Stafford, C., & Shafer, D. (1961). Public Relations and Test Results. *The Clearing House*, 35(5), 270-272.
- Tompkins, E. (1961). External Tests and the Schools. *The Clearing House*, 35(9), 515-516.

- Traxler, A. (1953). 12 Current Trends in Testing: How Schools are Making Better Use of Tests. *The Clearing House*, 28(1), 3-7.
- Urevick, S. J. (1965). Ability Grouping: Why Is It Undemocratic? *The Clearing House*, 39(9), 530-532.
- Wedeen, S. U. (1960). Uses and Misuses of Aptitude Tests. *he Clearing House*, 35(1), 11-12.
- Williams, R. L. (1980). The War Against Testing: A Current Status Report. *Journal of Negro Education*, 49(3), 263-273.
- Youngs, W.-P. H. (2008). Does High-Stakes Testing Increase Cultural Capital among Low-Income and Racial Minority Students? *Education Policy Analysis Archives*, 16(6), 1-18.

## VITA

Julianna M. Cechowski was born on the Southside of Chicago. She attended Chicago Public Schools for her K-12 education. During her Junior year of high school, Miss Cechowski was offered a position as a student worker in the Main Office of her high school. After Senior year of high school, Miss Cechowski was unsure if she wanted to continue her educational career.

While still working at her high school, her Freshmen year homeroom teacher convinced her to enroll at a community college. That summer she enrolled at Moraine Valley Community College in Palos Hills, Illinois. After receiving her Associates degree in Psychology, Miss Cechowski waited a year before entering Saint Xavier University Chicago. She soon completed her Bachelor's Degree in Business Administration.

During this time, Miss Cechowski had continued her career at Chicago Public Schools. Not knowing where she wanted to go, Miss Cechowski decided to enroll at Loyola University Chicago and obtain her Masters of Arts in Cultural and Educational Policy Studies. During her time at Loyola, Miss Cechowski developed a better understanding of the educational environment and decided she wanted to stay with Chicago Public Schools.

