

# **Standardized Testing: What is it Good For?**

## **-A Case Study in Connecticut**

The Honors Program  
Senior Capstone Project  
Student's Name: Megan Mapp  
Faculty Sponsor: Kristin Kennedy  
April 2015

**Table of Contents**

Abstract ..... 1

Introduction ..... 2

Literature Review ..... 2

    Definition and History of Standardized Testing ..... 2

    The Debate of Standardized Testing ..... 3

    Previous Studies ..... 4

Research Objective ..... 6

Research Question ..... 7

CAPT Background ..... 7

Methodology ..... 8

Results ..... 8

    Demographic Analysis ..... 8

    CAPT Math Score Analysis ..... 9

    Graduation Rate v. CAPT Score Analysis ..... 10

    Survey Analysis ..... 10

Future Implications ..... 12

Conclusion ..... 13

Appendices ..... 15

    Appendix A- School Demographic Information ..... 15

    Appendix B- CAPT Scoring Rubric ..... 15

    Appendix C- Graph of CAPT Mean Math Scores by School ..... 16

    Appendix D- One-Way ANOVA Test Results ..... 17

    Appendix E- T-Tests for Pre and Post 2006 Comparison ..... 18

    Appendix F- Graduation Rates ..... 19

    Appendix G- Map of Connecticut Teacher Survey Distribution ..... 20

    Appendix H- Teacher Survey ..... 21

    Appendix I- Results of Survey Questions ..... 24

..... 26

..... 27

..... 28

..... 29

..... 30

..... 32

    Appendix J- Teacher Comments ..... 33

References ..... 34

**Standardized Testing: What is it Good For?**  
*Senior Capstone Project for Megan Mapp*

---

**ABSTRACT**

The case study was developed in an attempt to shed more light on the debate of standardized testing. The goal of the study was to find evidence to support whether or not standardized testing is worth doing in public secondary schools. To investigate this question, the state standardized math test scores of three Connecticut public high schools were analyzed. The average math scores over thirteen years were observed and statistical analysis was performed to see if any significant differences existed between the three schools. Tests were performed before and after the change in standardized test. The graduation rates of the schools were observed and compared to the trend of the CAPT mean math scores over time. This analysis was then supplemented with responses from a survey distributed to Connecticut high school math teachers to take into consideration the educators' views of standardized testing. Both the quantitative and qualitative data had conflicting results. The standardized test scores appeared to improve over time, while the teachers found their teaching and student learning was interfered with the testing. Following the analyses, future implications of using standardized testing and how it may affect the transition to the Common Core Standards is discussed.

## **Standardized Testing: What is it Good For?**

### ***Senior Capstone Project for Megan Mapp***

---

#### **INTRODUCTION**

According to a report produced by the Brown Center of Education Policy at Brookings Institution, approximately 1.70 billion dollars is spent annually on standardized testing alone in the United States (Chingos 1). Why are we spending that amount of money and effort towards standardized testing? What are these tests used for and how effective are the tests in assessing student learning? Whom does it benefit and why spend time from the classroom to prepare for these assessments and take them? There has been a long debate over standardized testing in schools and whether or not they are advantageous to the education system. With the Common Core Standards based assessments being implemented in 46 of the 50 states, it would be wise to investigate how these standardized tests are really used. In the past and in recent news, varied claims have been made about standardized testing. Some say that too much time is spent focusing on standardized testing. Others say standardized testing is needed to hold schools accountable to improve student learning. In this study, a review of literature will provide a deeper look into these claims and what studies have been conducted to test these claims. Following this, I conduct a case study of three Connecticut high schools and look at standardized test scores to test the question of whether or not standardized testing is worthwhile. Responses to a teacher survey will illuminate the teachers' perspective on standardized testing. Interestingly, the teacher survey did not show that introducing the Common Core would be a positive move in the future. After an in-depth analysis of the data, I hope a discussion of future implications will encourage others to do more research on the topic of standardized testing.

#### **LITERATURE REVIEW**

##### **Definition and History of Standardized Testing**

As stated by the Gale Encyclopedia of Psychology, standardized testing is used in both psychology and education to assess personality, intelligence or achievement. These tests are "standardized" in that the same test is applied to a group of people under the same conditions and scored in the same manner. These tests are seen useful because they can eliminate bias by ensuring the same instructions and questions are administered. In addition, scores can be compared between those who take the test (Strickland 626-7). Standardized testing was first

## **Standardized Testing: What is it Good For?**

### ***Senior Capstone Project for Megan Mapp***

---

introduced in the 1900s to assess military intelligence (US Congress). Since then, it has become a more prominent tool in education. Specifically, it has been enforced by federal government through the Elementary and Secondary Education Act (ESEA) and more recently by the 2002 No Child Left Behind (NCLB) Act. The NCLB Act was established to require all states to implement their own set of standards to be tested annually at the primary and secondary level of education. These tests were to be administered in grades three through eight and once at the high school level. This was intended to hold schools accountable to strive to improve all student's achievement scores regardless of the student's background and/or learning abilities. Importantly, funding to schools for these assessments was used to motivate the use of these tests (Holmes 4).

According to the Glossary of Education Reform, standardized testing under the No Child Left Behind Act can be considered as "high-stakes testing". The main characteristic of these tests is that they are used to make decisions about education systems, students, and teachers. By associating certain rewards and consequences to these tests, it is believed to hold educators accountable and encourages improvement in overall test scores ("High-Stakes Test Definition").

#### The Debate of Standardized Testing

There has been a debate over the use of standardized testing and whether or not it should be used in education systems to evaluate school systems and assess student learning. Opponents of standardized testing such as the National Center for Fair and Open testing claims that standardized testing or "high stakes testing" is damaging to education for several reasons. They say that too much of teaching is focused on preparing for these exams. This causes the span of the curriculum to be narrowed and students do not receive information as in depth or other subjects are excluded from their learning. Other reasons against the use of high-stakes testing include: unfair to students, increased grade retention and dropout rates pushes away "good teachers", and misleads the public to conclusions that may not be true or lack sufficient evidence (FairTest).

Furthermore, the findings of a 2012 survey of over 20,000 public school teachers was published in *Education Week* and the overall conclusion was that most educators did not find standardized testing to be valuable in measuring and evaluating student learning. Teachers

## **Standardized Testing: What is it Good For?**

### ***Senior Capstone Project for Megan Mapp***

---

would agree that the tests are unfair and provide misleading results. These teachers say that these assessments are not an accurate portrayal of student learning because students tend to not take these tests seriously. Secondly, students may face test anxiety in this type of testing environment which could interfere with their focus and lead to lower, inaccurate scores (Rebora 14). Stephanie Schneider claims that the current use of standardized testing does not help to assess *all* students in their learning. Instead, she shows how its use discriminates against specific groups such as low income households or minorities. She believes there are better methods to assess students. She claims that the tests take time away from actual teaching and cannot be used effectively to evaluate teaching because teachers do not know what is asked on the test and do not see the exams (Schneider and Christison 30).

On the other hand, proponents of standardized testing like Matt Christison say that standardized testing can provide useful information about student learning if used correctly. Christison claims that tests that are developed well will show similar results between students who share similar characteristics and environments. He says that it is the extent of use and interpretation of the results that creates the problem rather than standardized testing alone. However, the author also suggests that other assessments, in addition to standardized testing, should be utilized in order to get a comprehensive evaluation and see academic achievement by all students (Schneider and Christison 30-2).

In an article from the Global Post, the author mentions that advocates of standardized testing find that these tests hold teachers and schools accountable in teaching and preparing their students for academic success. Standardized testing also makes sure there is equality in education and that everyone has access to education (Thompson).

#### Previous Studies

To further investigate these claims, several studies have been conducted to see the impact of standardized testing on students and other factors related to high stakes testing. For instance, a study from 1991 asks whether increases in standardized test scores means increases in learning. This study conducted by Joan Herman and Shari Golan composed of surveys/questionnaires that were distributed to elementary teachers from eleven school districts in nine states. A series of questions were asked to assess their views about the impact of standardized testing on education and how it has influenced their work. Based on teachers'

## **Standardized Testing: What is it Good For?**

### ***Senior Capstone Project for Megan Mapp***

---

perceptions, results showed that most teachers felt pressured to improve test scores and said that the standards influenced their instruction. (Herman and Golan).

In 2002, Martin Carnoy and Susanna Loeb asked “Does External Accountability Affect Student Outcomes?” A rating scale from zero to five was created to measure the strength of accountability through the use of high-stakes standardized testing amongst the 50 states. These ratings were then compared to National Assessment of Education Progress (NAEP) math scores from 1996 to 2000, retention rates, and high school completion rates. The results of this study showed that states classified as having strong accountability systems had higher increases in NAEP math scores for eight grade students than in low accountability states. In addition there was no significant difference in both retention and high school completion rates between states with strong accountability systems and those with low accountability systems (Carnoy and Loeb).

Another study conducted in 2005 by Amrein and Berliner looked at how high stakes testing implemented by the No Child Left Behind Act improved academic achievement scores of students. In order to assess this, NAEP scores of 27 states with the highest stakes were analyzed at the primary school level before and after the act was passed. They also looked at academic achievement at the high school level of these states where graduation exams were put into use. SAT, ACT, and AP test scores were used to evaluate this. The results of the study showed that high stakes testing did not show a significant enough impact on NAEP scores after the act was passed. In addition, academic achievement scores decreased after the use of high school graduation exams (Amrein and Berliner).

With many criticisms, about standardized testing, several articles and papers propose methods that can be used to assess education success and student academic achievement. Recently, talk about a national set of standards, the Common Core, is being considered

In Monty Neill’s article from Education Week, he believes that an effective assessment and evaluation process for education has three components: “limited large scale standardized testing, extensive school-based evidence of learning, and a school quality review process” (Neill). He finds it would be better to do less testing in schools, in which one assessment is given at each school level (elementary school, middle school, and high school). Also, it is suggested to use more class-based assessments as a way to evaluate student learning.

## **Standardized Testing: What is it Good For?**

### ***Senior Capstone Project for Megan Mapp***

---

However, these assessments would have to be well designed and consistent between classrooms. In general, the writer advocates the use of more “in the classroom” evaluation rather than standardized testing (Neill).

In addition, John Chubb discusses a set of new and better assessments that will be implemented as a result of the consideration of the Common Core Standards. The new assessments will be created by two groups of states, the Partnership for Assessment of Readiness for College and Careers (PARRC) and the Smarter Balanced Assessment Consortium (SBAC). These national tests will be used by each participating state: 23 under PARRC and 23 under SBAC. The tests will set higher standards than previous state standardized tests. And unlike previous standardized tests, these new tests and the new Common Core Standards will be used to assess “college and career readiness” to ensure students are prepared for postsecondary education. The author of this article advises to pay attention and consider this upcoming change (Chubb 9-11).

Similarly, a journal article by Frederick Hess and Michael McShane talks about the Common Core standards. However, the article suggests that the success of the new assessments to be used under the Common Core Standards will depend on how the standards are implemented in schools in each state. Specifically, the writers claim that these standardized scores will only have meaning if all teachers align their class instruction to these standards. Furthermore, the article expresses apprehension towards these new standards because drastic changes will need to be made in school development and materials. They stress that while a new set of standards may appear to help fix the NCLB act, there are still potential problems that can arise with this new solution (Hess and McShane).

As the focus of standardized testing continues to evolve over time, it is necessary to continue research in this area because the continuous changes to the state administered tests may not be beneficial to the educational environment, teachers, and most importantly, the students.

### **RESEARCH OBJECTIVE**

The objective of this research is to provide a recent study evaluating the impact of state specific standardized testing in secondary education. Based on the literature review, most studies on standardized testing were conducted five to ten years ago. Therefore, would a



## **Standardized Testing: What is it Good For?**

### ***Senior Capstone Project for Megan Mapp***

---

recent study show that similar results were occurring today? In addition, past studies have focused on overall standardized testing in both primary and secondary schools or only at the elementary and middle school level. This case study focuses on the secondary education level and attempts to analyze state standardized test scores in high schools. In addition, past studies did not have much analysis on how standardized testing impacts the state of Connecticut at an individual state level and in the aggregate of state comparisons. Therefore, this study is important in evaluating the impact of state-specific standardized testing in Connecticut high schools.

### **RESEARCH QUESTION**

The research question to be investigated is whether or not state-specific standardized testing is worth doing in the state of Connecticut.

### **CAPT BACKGROUND**

The implementation of the Connecticut Academic Performance Test (CAPT) began in 1994. This test assessed students in math, reading and writing at the tenth grade level. Specifically, the math portion of the CAPT tests students in concepts related to geometry, statistics, probability, and algebra. Over the years, the test has been updated and three “generations” or versions have been created. The first generation of the exam was distributed between 1994 and 2000, the second generation of the assessment began in 2001 to 2006, and the third generation of the test was taken from 2007 to 2013 (Sternberg, 2006). However, only public data is made available for the second and third generations of the test. There are no CAPT test score averages made available beyond 2013 because the state of Connecticut has decided to move to the Common Core Smarter Balanced Testing. While scores may no longer be relevant, a look at past test scores can reflect what happens when another test change is being made.

## **Standardized Testing: What is it Good For?**

### ***Senior Capstone Project for Megan Mapp***

---

#### **METHODOLOGY**

To test my research question, a case study of three Connecticut public high schools of similar demographics was conducted. These high schools were Berlin High School, Newington High School, and Wethersfield High School. All schools were located in the Hartford area of central Connecticut. Other demographic information about the high schools was collected through the state Department of Education in order to evaluate their similarities. To measure the effectiveness of state standardized testing, both quantitative and qualitative data was collected. For quantitative data, the Connecticut Academic Performance Test (CAPT) math score averages for each school were collected. Individual data for analysis could not be used due to federal privacy law and confidentiality issues. As a result, the summarized averages for the math portion of CAPT were obtained from the Connecticut State Department of Education. There were no summarized score averages after 2013 because the schools are no longer taking the CAPT and changing over to use a Common Core based assessment. While the CAPT is comprised of a reading, writing, and science portion in addition to the math section, the focus was to use the scores for the math section because the scoring is not subjective. Additionally, an online survey was created and distributed to Connecticut high school math teachers willing to participate. The survey consisted of approximately fifteen questions to collect demographic information, the teachers' perspectives of standardized testing today, as well as any personal comments the respondents felt comfortable sharing. This survey was completely voluntary and respondents could opt out of completing the survey at any time. Once all data was collected and organized, both the standardized test scores and survey responses were compared and analyzed using statistical tools.

#### **RESULTS**

##### **Demographic Analysis**

Before the CAPT mean math scores were collected, the demographics of each high school were analyzed to see if they shared similar characteristics and would be comparable. Data from the CT State Department of Education was obtained for each of the schools from 2001 to 2013 from their individual *Strategic School Profile Reports*. The information collected included: town population, median household income, average high school enrollment,

## **Standardized Testing: What is it Good For?**

### ***Senior Capstone Project for Megan Mapp***

---

average instructional time per year, average high school algebra class, average percentage of minority students enrolled in high school, and average number of students per computer.

With the exception of town population and median household income, all other demographic information was collected for each year and then averaged to get a summarized value for the 2001-2013 time period. As the table in Appendix A depicts, the three towns where the high schools are located are relatively similar. All three high schools have approximately 1,000 hours of instructional time per year, have between 1,000 and 1,500 students enrolled in high school on average, and have an average Algebra class size of approximately 20 students.

However, Berlin has a slightly larger reported median household income (\$85,735) and the lowest average percentage of minority students (7.7%) within the high school. This may have an influence on the CAPT score results and show significant differences. We can still compare the towns, especially since Connecticut's Department of Education identifies the three high schools to be within the same district reference group. This means that the schools are identified as having similar features and can be comparable.

#### CAPT Math Score Analysis

To analyze the data, a graph of the mean math scores for each school was created. The graph in Appendix C shows the trend of the scores generally increased over time. All scores are compared to the goal line, which is the rating assigned to a minimum score of 261. This is what the state considers to be the target for all students and indicates that the standards have been met. Yet, in 2005 to 2006, there was a significant dip in scores for Newington and Wethersfield high school. After 2006, the scores appear to increase, until 2012, when the scores fall again. This change may be a result of the introduction of the third generation of the CAPT, an updated version of the assessment. Test scores do not go beyond 2013 because the movement to the Common Core is in progress.

In addition, a one way ANOVA test (see Appendix D) was conducted to test if there were any significant differences between the CAPT mean math scores of the three high schools. An average of the 13 means for each school was calculated and compared. The test produced an F- value of 3.77 and a probability value of 0.033\*. This result is considered to be significant at the 5% significance level, indicating that there is a slight difference between the three mean scores. Berlin High School had the highest mean math score of 265.66, followed by

## **Standardized Testing: What is it Good For?**

### ***Senior Capstone Project for Megan Mapp***

---

Newington High School with a mean math score of 262.47. Wethersfield High School had the lowest mean math score of 261.14.

Next, a comparison of the average math scores before 2006 and after 2006 was conducted for each test. A two-sample t-test was conducted for each school. The scores from 2001-2006 were averaged while the scores from 2007 to 2013 were averaged and tested to see if there was a significant difference in the mean math scores between the two time periods. The results for these tests can be found in Appendix E. The tests indicated significant differences in mean scores for both Berlin High School (p-value = .006) and Newington High School (p-value = .001) at the 5% significance level, with mean math scores significantly higher after 2006. This may have occurred because students became more comfortable with the new version of CAPT over time and therefore the averages increase. On the other hand, Wethersfield High School did not show a significant difference in the mean math scores between the two time periods. There may be some underlying factors for why this occurred and it would be important to further investigate.

#### Graduation Rate v. CAPT Score Analysis

The available graduation rates from 2004 to 2008 were observed and compared to the trend in scores as a way to evaluate the assessment. If standardized tests were being used to help the students, we would assume that the graduation rates would remain steady or increase over time. The graph in Appendix F, shows that the graduation rates for Berlin and Wethersfield are bell shaped, with a slight incline and then a decrease after 2006. Newington graduation rates follow a similar trend but are more stable (show less movement). It is important to note that the corresponding test scores for these graduation rates are two years prior to the graduation year since students take the test in the tenth grade. In 2008, where the graduation rate was the lowest, corresponds to the dip in the CAPT mean math score in 2006. If the CAPT was used as a graduation requirement, those that didn't pass the CAPT or chose not to take the exam may not have graduated as well, resulting in the decrease in graduation rates.

#### Survey Analysis

The distribution of the survey to Connecticut high school math teachers resulted in 20 respondents, with six respondents representing both Newington and Wethersfield High School. While 20 is a reasonable amount of responses, it may not be large enough to have

## **Standardized Testing: What is it Good For?**

### ***Senior Capstone Project for Megan Mapp***

---

sufficient evidence to draw conclusions about the population. This may be due to the fact that respondents could opt out of taking the survey at any time. In addition, the target group for the survey was specific and narrowed the potential number of total responses. Fortunately, all respondents were from the central area of Connecticut. We could assume that the schools which the respondents teach at have some similar characteristics. A map of Connecticut displaying the distribution of responses can be found in Appendix G.

A majority of the participants were female, and most of the teachers had 10-19 years of total teaching experience and teaching experience at their current school. Approximately 60% of teachers indicated that a decent portion of classroom/personal preparation time was taken from standardized testing. Most of the teachers were unhappy and had low job satisfaction as 80% of the respondents rated this category as a 1 or 2 on a scale of 5. Furthermore, 90% of the teachers believed less time should be spent focusing on standardized testing. This may be due to the fact that approximately 70% of the teachers felt pressured by state requirements to teach a certain way and most teachers agreed or strongly agreed with the statement that standardized testing added more responsibilities to their job. There were mixed responses to the statement, "You have a say in what curriculum changes are made by the school in response to CAPT results." 30% disagreed, 30% agreed and more than 25% neither agreed nor disagreed. Most teachers think the time spent on standardized testing interferes with their learning, and of those, about 85% agreed or strongly agreed with this. Regardless of how teachers responded to the previous questions, all 20 of them believed the Common Core assessment will not be better and have similar issues arise. The results of each survey question asked can be found in Appendix H.

Of the 20 teachers who took the survey, 12 left responses in the open ended section for general comments. Overall, most of the participants shared similar thoughts in regards to standardized testing. They indicated that having more testing meant less time for instruction and student learning, they feared the performance on standardized tests would be part of their evaluation, and that switching to the Common Core based assessment would disrupt teaching because students would need to adjust to the new testing format. The specific comments can be found in Appendix J.

## **Standardized Testing: What is it Good For?**

### ***Senior Capstone Project for Megan Mapp***

---

#### **FUTURE IMPLICATIONS**

Based on these results, we can make some inferences on what will happen with continued use of standardized testing in secondary schools. However, we cannot make broad generalizations for all states and schools because only three high schools of one state was analyzed. In addition, only 20 teachers responded to the survey.

When the trend of scores were analyzed, we noticed there was a dip in scores prior to 2006 and then a significant increase after 2006, when CAPT was changed to a newer version. So what happens when the test is changed to a Common Core based assessment? Not only is it a completely new test format, but the focus of questions is now to assess college readiness of students. The CAPT was a paper and pencil based assessment while the Common Core assessment will be computer based and have two versions that can be taken. This could pose a problem for those schools whose students are not familiar with using a computer because they cannot afford one or the school doesn't have enough resources for their students. If we change the test again, will that glorify improvement? The scores will be low at first as students adjust to the test and then improve the following years after having practice and more familiarity with the test. So how can we say that the score assesses student learning if the change in score may be due to the fact the students have had more practice with the format and the type of questions being asked. Additionally, the CAPT was administered at the tenth grade level and will now be taken at the eleventh grade ("New Era Of Student Testing Set To Begin In Connecticut"). In this grade, students are already overwhelmed with test taking as they prepare for AP exams and the SATs. This could add more stress and lead to test anxiety, hindering a student's potential score. Therefore, would adding another standardized test to the students' agenda be beneficial?

Updating the test or changing the type of test every few years does not help in conducting longitudinal studies. It also makes the teacher's job more difficult because they too have to adjust their lesson plans to the new curriculum that goes along with the new test. There is no time for the teachers to adjust and truly see if there is progress in the scores. Data limitations also make it difficult to conduct research studies that are not produced from the state department of education.

## **Standardized Testing: What is it Good For?**

### ***Senior Capstone Project for Megan Mapp***

---

Another issue with switching the assessment to the Common Core is that it doesn't fix the problem of the inability to compare across states. Prior to the introduction of the Common Core, each state administered their own state specific assessment. One of the reasons the Common Core was built was to allow for state to state comparison. Unfortunately, there are two different assessments under the Common Core and not every state has agreed to make the switch. Some states are unable to afford the new assessment and many parents are opting out their child from taking the new test. Even if one test was given, can we actually compare the results between schools and states? Students have different learning backgrounds and teachers have different styles of teaching.

The results of this study and the future issues posed by the Common Core indicate that further research needs to be conducted. It would be interesting to poll administrators and parents about their thoughts on standardized testing as well as how the parents' views may influence their children's perception of standardized testing. I would also suggest more schools and different districts within Connecticut should be observed to see if there is a similar trend in scores across the state. I would also further investigate why Wethersfield High School's average math score did not change before and after 2006, especially since it had the lowest overall mean score in comparison to Berlin and Newington. Lastly, I would test to see if any gender or ethnicity differences in scores exist. This would provide evidence to show if the implementation of the NCLB act has helped to narrow the achievement gaps of students of various backgrounds.

### **CONCLUSION**

To answer whether or not standardized testing is worth having in the education system is a difficult question to answer because there is not a clear, straightforward reasoning. This is due to the fact that there are numerous variables that influence standardized testing scores and effectiveness which cannot all be evaluated at once. More specifically, data limitations and difficult with data accessibility prevents studies from obtaining all usable information of these variables.

While it is important to assess student learning, using standardized tests mandated by the NCLB act may not be reasonable because schools can be rewarded or penalized based on

## **Standardized Testing: What is it Good For?**

### ***Senior Capstone Project for Megan Mapp***

---

score results. The scores are used to determine school funding, and sometimes are used in teacher evaluation and as graduation exams. For example, implementing a new test when scores are falling may lead to an increase in scores over time because there has been more practice with that assessment and students have become more familiar with the new test material. Another reasoning for an improvement in scores could be that with more standardized tests, students may choose to opt out of taking a new test, leaving those students who are serious and willing to take the test. This can inflate the scores if we make the assumption that those students who better understand the concepts are taking the test, making the scores appear to increase. As the analysis of the CAPT math scores and teacher survey results show, what may appear to be an improving trend in scores may be hiding the true impact of standardized testing. Teachers are the ones who work so closely with students and know what is happening in the classroom. Unfortunately, based on the teacher survey results, the increased focus of standardized tests interferes with the teachers' ability and time to teach the students real world applications. Rather than focusing on evaluating the school and determining funding based on improvement in scores, student learning should be the main focus for evaluation. If schools are being rewarded for showing improvement and the schools failing to improve do not receive an increase in funds, then how can we expect those schools who lack resources and sufficient funding to provide low income/low achieving students with a fair and better education? Standardized testing prevents teachers from doing their job and deters "good teachers" from working at underperforming schools because their job evaluations may be tied to the results of the school testing. Overall, the "worth" of standardized testing is a debatable one and difficult to measure. Clearly, this study has brought more questions about standardized testing to be answered.



**Standardized Testing: What is it Good For?**  
*Senior Capstone Project for Megan Mapp*

---

**APPENDICES**

Appendix A- School Demographic Information

	<b>Berlin</b>	<b>Newington</b>	<b>Wethersfield</b>
<b>2012 Town Population</b>	19,974	30,520	26,670
<b>2012 Med. HH Income</b>	\$85,735	\$75,237	\$76,196
<b>Average High School Enrollment</b>	1,035	1,455	1,166
<b>Average Instructional Time/Year (in hours)</b>	1,036	1,004	972
<b>Average HS Algebra Class</b>	18.4	21.0	21.2
<b>Average % Minority</b>	7.7%	21.0%	18.6%
<b>Average # of Students/ Computer</b>	2.2	2.6	3.9

Sources: Connecticut Economic Resource Center. "Berlin Town Profile." Oct. 2014. Web. 5 Apr. 2015.

Connecticut Economic Resource Center. "Newington Town Profile." Oct. 2014. Web. 5 Apr. 2015.

Connecticut Economic Resource Center. "Wethersfield Town Profile." Oct. 2014. Web. 5 Apr. 2015.

Connecticut State Department of Education. "Strategic School Profiles." 2013. Web. 5 Apr. 2015.

Appendix B- CAPT Scoring Rubric

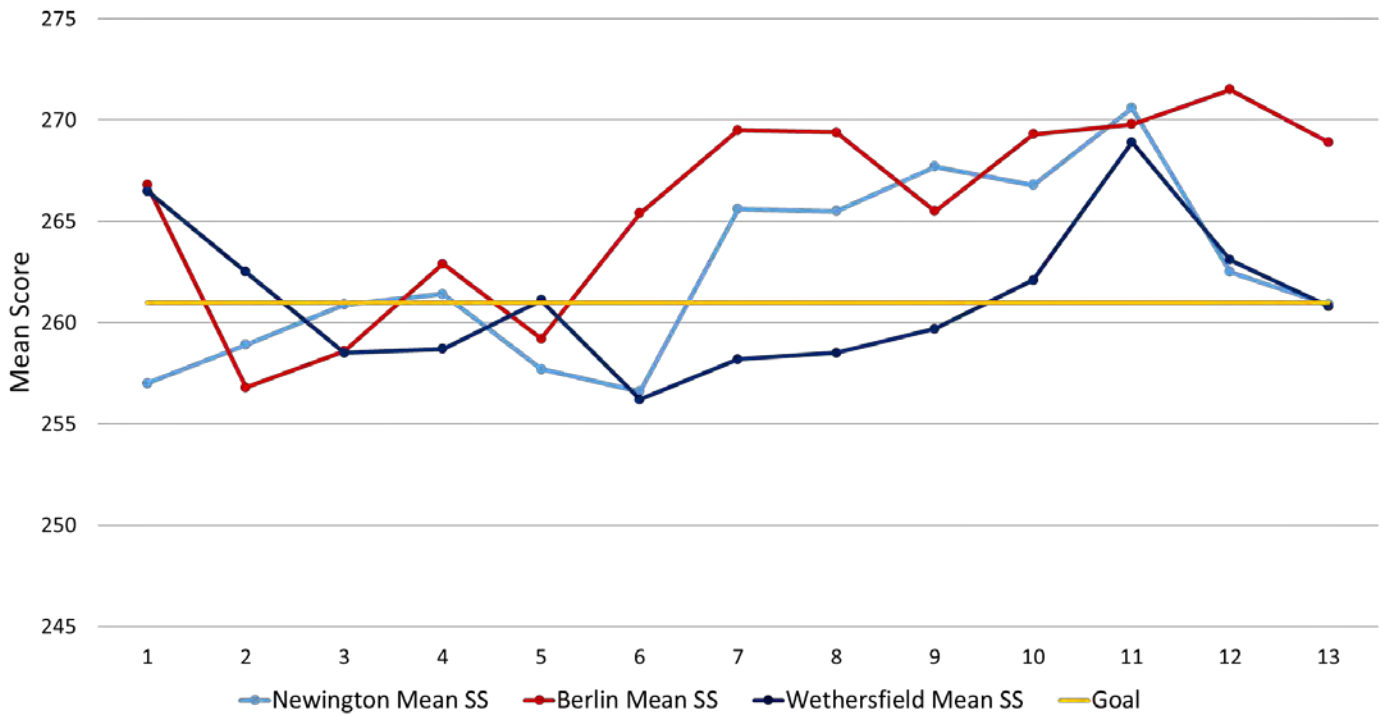
Total Score Range: 100-400

<b>Below Basic</b>	100 to 191
<b>Basic</b>	192 to 222
<b>Proficient</b>	223 to 260
<b>Goal</b>	261 to 286
<b>Advanced</b>	287 to 400

**Standardized Testing: What is it Good For?**  
**Senior Capstone Project for Megan Mapp**

Appendix C- Graph of CAPT Mean Math Scores by School

**CAPT Mean Math Score by School**



Year	Newington Mean SS	Berlin Mean SS	Wethersfield Mean SS
2001	257	266.8	266.5
2002	258.9	256.8	262.5
2003	260.9	258.6	258.5
2004	261.4	262.9	258.7
2005	257.7	259.2	261.1
2006	256.6	265.4	256.2
2007	265.6	269.5	258.2
2008	265.5	269.4	258.5
2009	267.7	265.5	259.7
2010	266.8	269.3	262.1
2011	270.6	269.8	268.9
2012	262.5	271.5	263.1
2013	260.9	268.9	260.8

Source: Connecticut State Department of Education. "CAPT Overall Summary Report." Web. 5

Apr. 2015.

**Standardized Testing: What is it Good For?**  
**Senior Capstone Project for Megan Mapp**

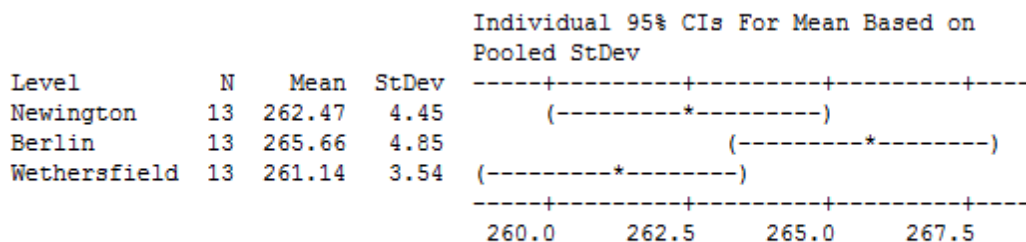
---

Appendix D- One-Way ANOVA Test Results

**One-way ANOVA: Newington, Berlin, Wethersfield**

Source	DF	SS	MS	F	P
Factor	2	140.5	70.2	3.77	0.033
Error	36	670.6	18.6		
Total	38	811.1			

S = 4.316    R-Sq = 17.32%    R-Sq(adj) = 12.73%



Pooled StDev = 4.32

## Standardized Testing: What is it Good For?

### Senior Capstone Project for Megan Mapp

---

#### Appendix E- T-Tests for Pre and Post 2006 Comparison

##### Two-Sample T-Test and CI: Berlin 1t6, Berlin 7t13

Two-sample T for Berlin 1t6 vs Berlin 7t13

	N	Mean	StDev	SE Mean
Berlin 1t6	6	261.62	4.02	1.6
Berlin 7t13	7	269.13	1.80	0.68

Difference = mu (Berlin 1t6) - mu (Berlin 7t13)

Estimate for difference: -7.51

95% CI for difference: (-11.86, -3.16)

T-Test of difference = 0 (vs not =): T-Value = -4.22 P-Value = 0.006 DF = 6

##### Two-Sample T-Test and CI: Newington 1t6, Newington 7t13

Two-sample T for Newington 1t6 vs Newington 7t13

	N	Mean	StDev	SE Mean
Newington 1t6	6	258.75	2.02	0.83
Newington 7t13	7	265.66	3.23	1.2

Difference = mu (Newington 1t6) - mu (Newington 7t13)

Estimate for difference: -6.91

95% CI for difference: (-10.19, -3.62)

T-Test of difference = 0 (vs not =): T-Value = -4.69 P-Value = 0.001 DF = 10

##### Two-Sample T-Test and CI: Wethersfield 1t6, Wethersfield 7t13

Two-sample T for Wethersfield 1t6 vs Wethersfield 7t13

	N	Mean	StDev	SE Mean
Wethersfield 1t6	6	260.58	3.63	1.5
Wethersfield 7t13	7	261.61	3.68	1.4

Difference = mu (Wethersfield 1t6) - mu (Wethersfield 7t13)

Estimate for difference: -1.03

95% CI for difference: (-5.56, 3.50)

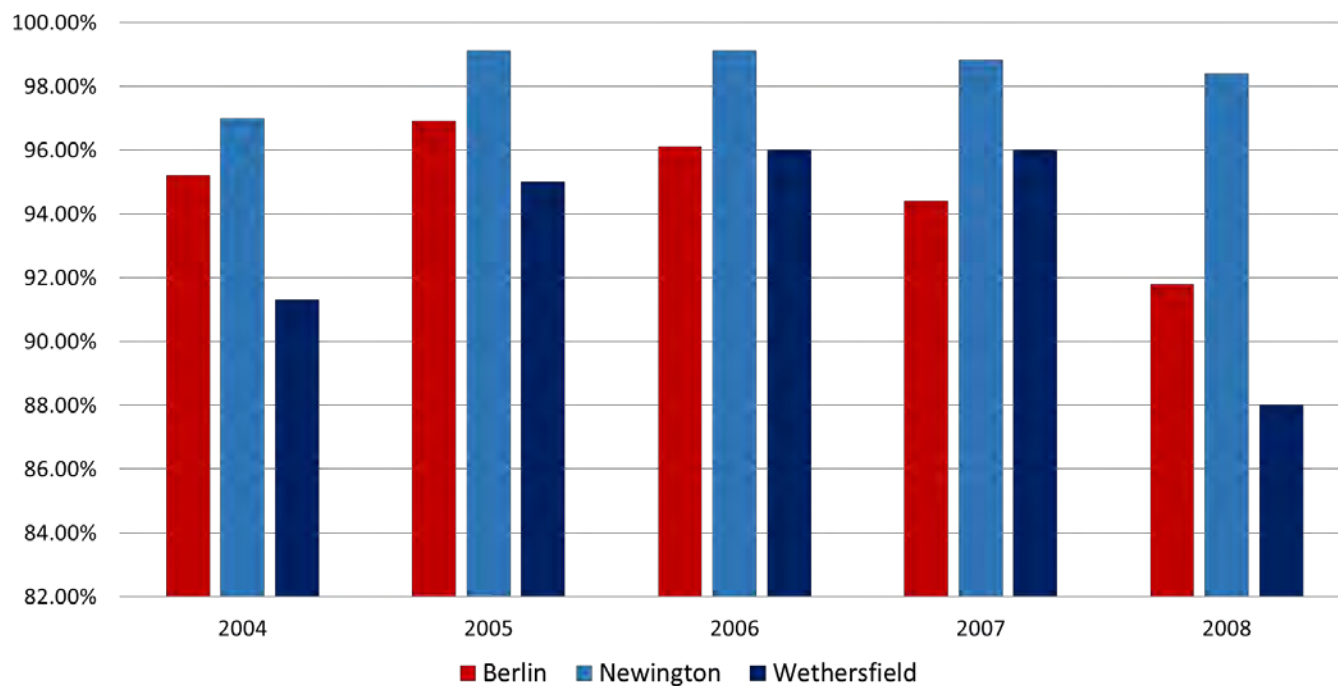
T-Test of difference = 0 (vs not =): T-Value = -0.51 P-Value = 0.623 DF = 10

**Standardized Testing: What is it Good For?**  
***Senior Capstone Project for Megan Mapp***

---

Appendix F- Graduation Rates

**School Graduation Rates by Year**



Source: Connecticut State Department of Education. "Graduation Rates Report." Web. 5 Apr.

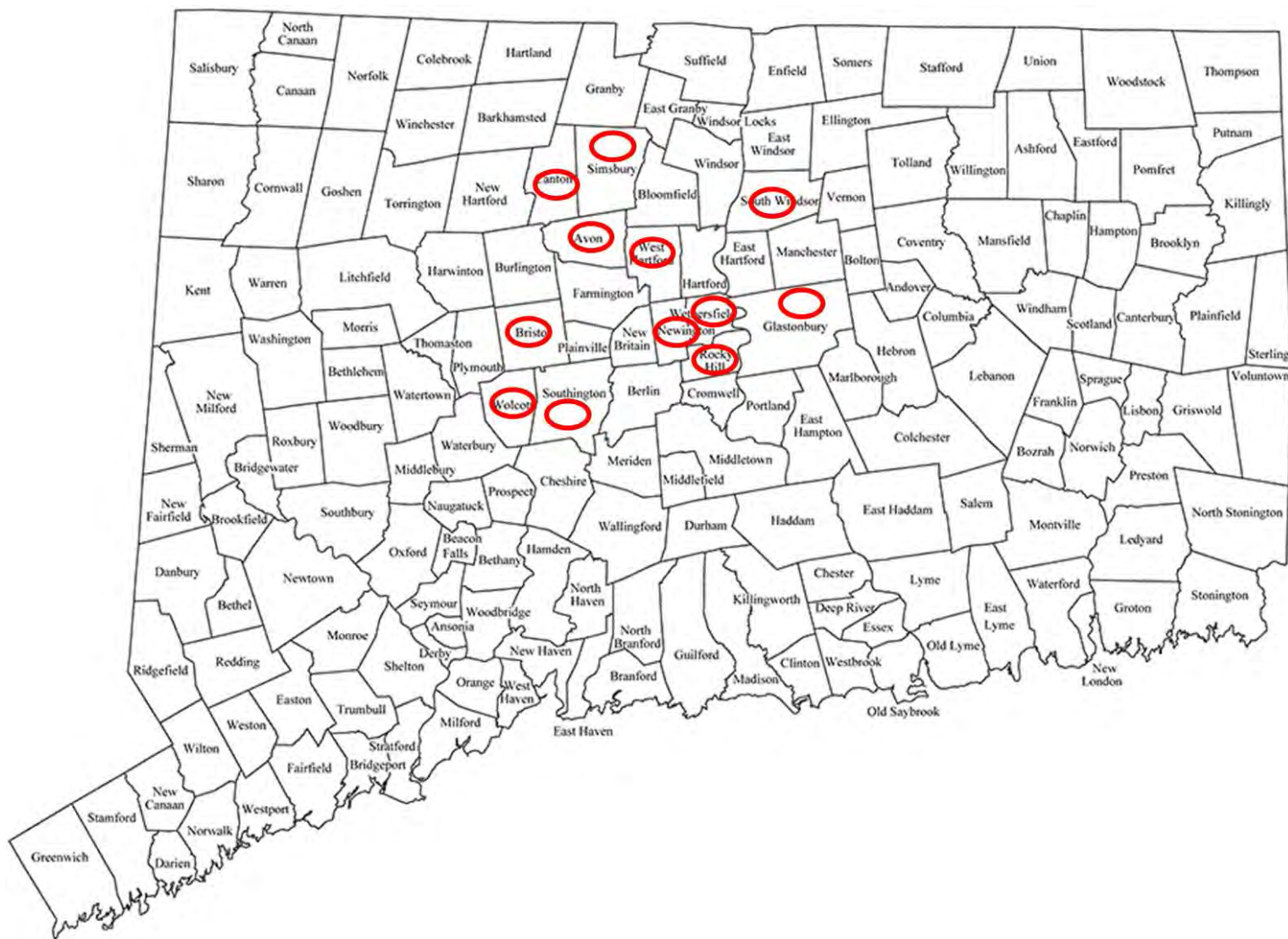
2015.

# Standardized Testing: What is it Good For?

Senior Capstone Project for Megan Mapp

---

## Appendix G- Map of Connecticut Teacher Survey Distribution



**Standardized Testing: What is it Good For?**  
***Senior Capstone Project for Megan Mapp***

---

Appendix H- Teacher Survey

Consent Form

Thank you for participating in this survey. By proceeding with this survey, you indicate your consent to take the survey. You are not required to complete the survey and your participation is voluntary. If at any time you do not feel comfortable with a question, you may choose to opt out of answering. If you wish to proceed, you may discontinue the survey at any time. Your responses will be completely anonymous and used for educational research purposes only. For any questions, please direct them to [mmapp@bryant.edu](mailto:mmapp@bryant.edu), [kkennedy@bryant.edu](mailto:kkennedy@bryant.edu), or [syoon@bryant.edu](mailto:syoon@bryant.edu) ).

Teacher Survey

I am conducting research in relation to the use of standardized testing in secondary schools. I am specifically investigating the impact it has on the educational environment and students and if these tests should continue to be administered. My honors capstone project will be focusing on CT state standardized tests distributed at the high school level and analyzing the CAPT math scores to assess performance. Your responses to this survey will be used to supplement the case study.

**School:**

**Gender:**

**Number of Years of Teaching Experience:**

**Years Teaching at School Named Above:**

**Math Classes You Teach:**

**Approximately how much classroom time is spent administering state standardized tests during the academic year?**

1- Not at all    2- Very Little    3- Somewhat    4- Quite a Bit    5- A Great Deal

**Approximately how much classroom time of teaching is spent on preparation for state standardized tests (including directions and practice)?**

1- Not at all    2- Very Little    3- Somewhat    4- Quite a Bit    5- A Great Deal

**Approximately how much personal preparation time is taken from the preparation for standardized testing?**

1- Not at all    2- Very Little    3- Somewhat    4- Quite a Bit    5- A Great Deal

**Approximately how much time is spent on paperwork for these tests?**

1- Not at all    2- Very Little    3- Somewhat    4- Quite a Bit    5- A Great Deal

**Standardized Testing: What is it Good For?**

*Senior Capstone Project for Megan Mapp*

---

**On a scale from 1 to 5, with 1= unhappy and 5= happy, how would you rate your happiness and overall job satisfaction as it relates to the use of standardized tests?**

1            2            3            4            5

**On a scale from 1 to 5, with 1= less and 5= more, how much time do you believe should be spent focusing on standardized testing?**

1            2            3            4            5

**On a scale from 1 to 5, with 1= no pressure and 5= significant pressure, do you feel pressured by state requirements to teach a certain way?**

1            2            3            4            5

*Please indicate how much you agree/disagree with the following statements:*

**You have changed your teaching curriculum in response to CAPT results.**

Strongly Disagree    Disagree    Neither    Agree    Strongly Agree

**There is a strong focus on CAPT/standardized testing in school.**

Strongly Disagree    Disagree    Neither    Agree    Strongly Agree

**You have a say in what curriculum changes are made by the school in response to CAPT results.**

Strongly Disagree    Disagree    Neither    Agree    Strongly Agree

**Standardized testing has added more responsibilities to your job.**

Strongly Disagree    Disagree    Neither    Agree    Strongly Agree



**Standardized Testing: What is it Good For?**

***Senior Capstone Project for Megan Mapp***

---

**Do you think the time spent and use of the CAPT standardized testing helps students or interferes with their learning?**

Helps

Interferes

**Based on your response to the previous question, to what extent do you support this statement?**

Strongly Disagree

Disagree

Neither

Agree

Strongly Agree

**With the movement toward the Common Core, do you believe this will be a better assessment than the CAPT or do you find similar issues will arise?**

Yes (will be better)

No (similar issues)

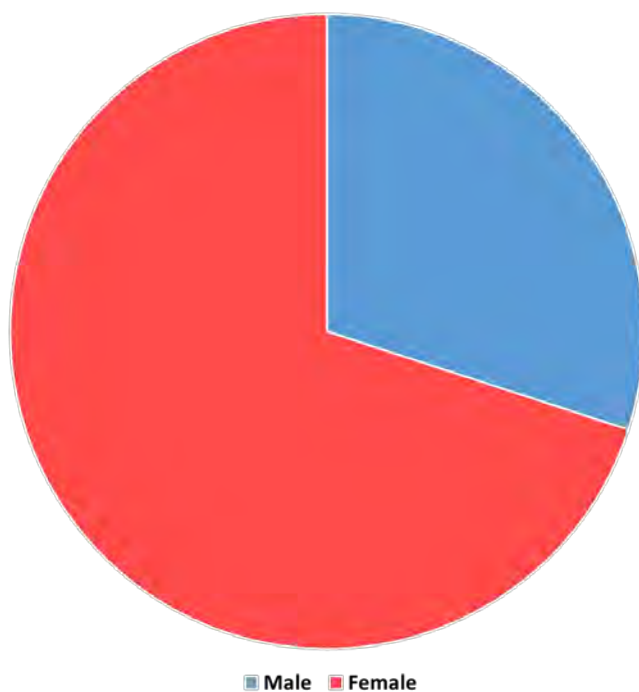
**Any other general comments about the assessments that you feel comfortable sharing?**

**Standardized Testing: What is it Good For?**  
*Senior Capstone Project for Megan Mapp*

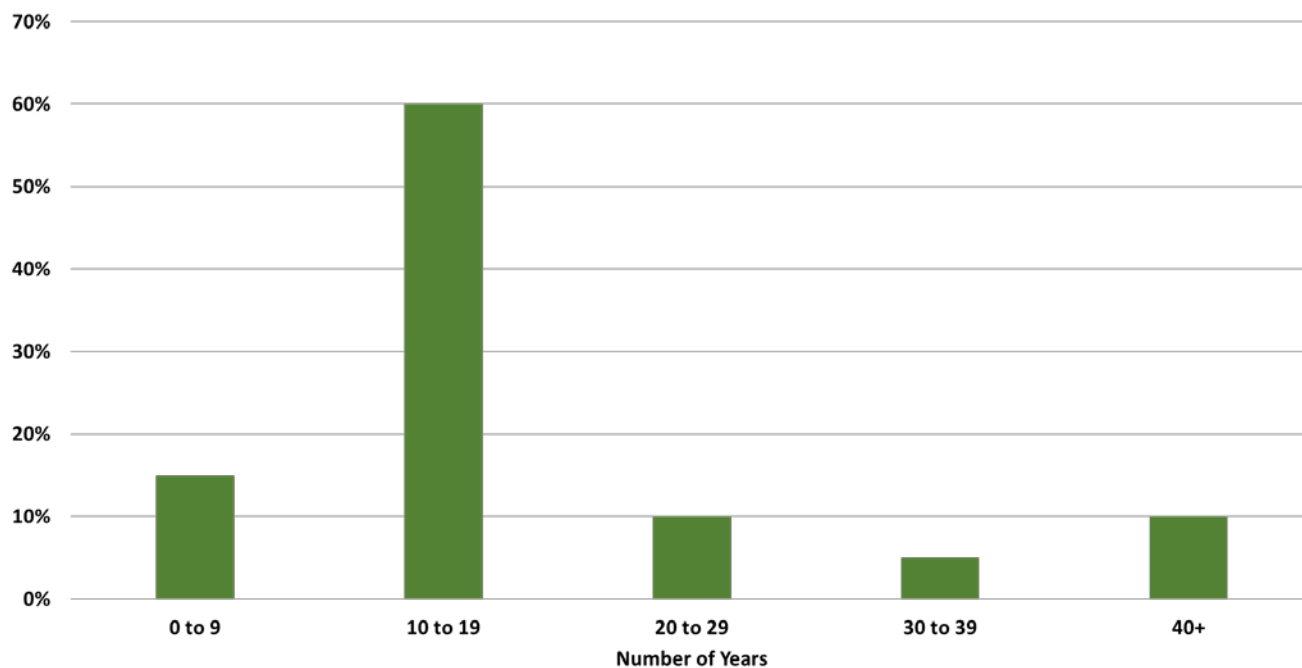
---

Appendix I- Results of Survey Questions

**Responses by Gender**



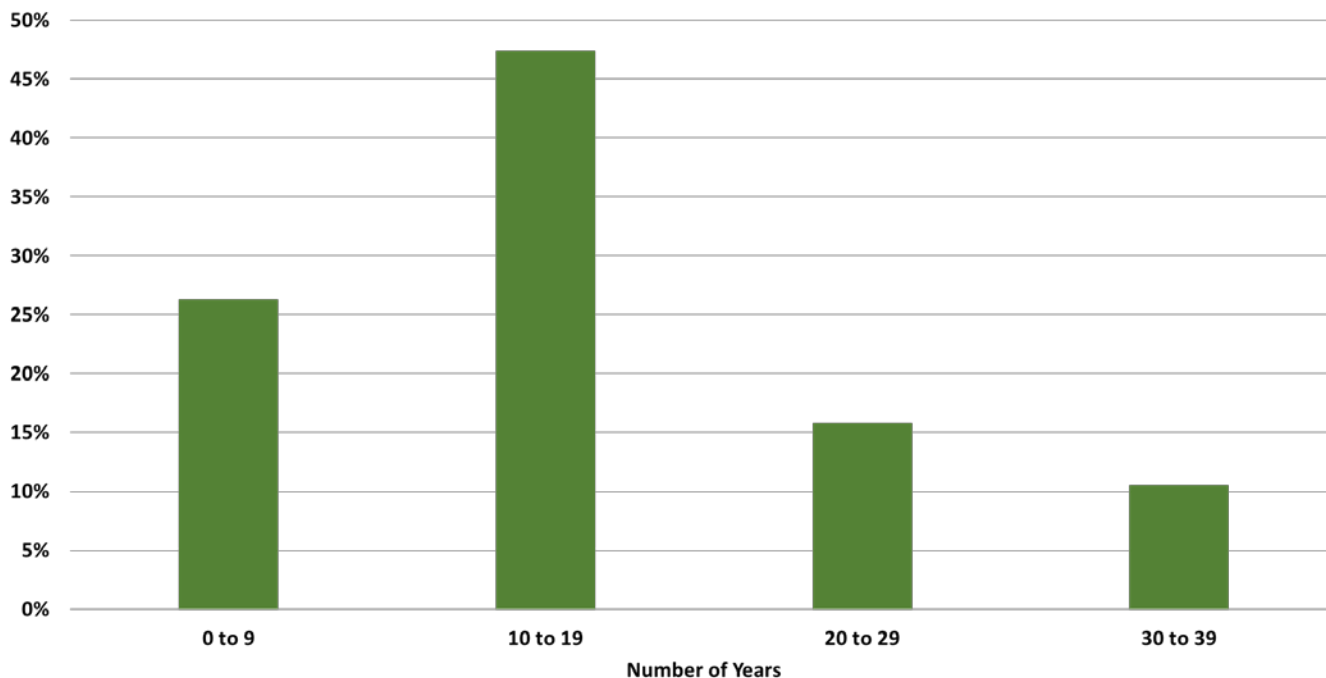
**Years of Teaching Experience**



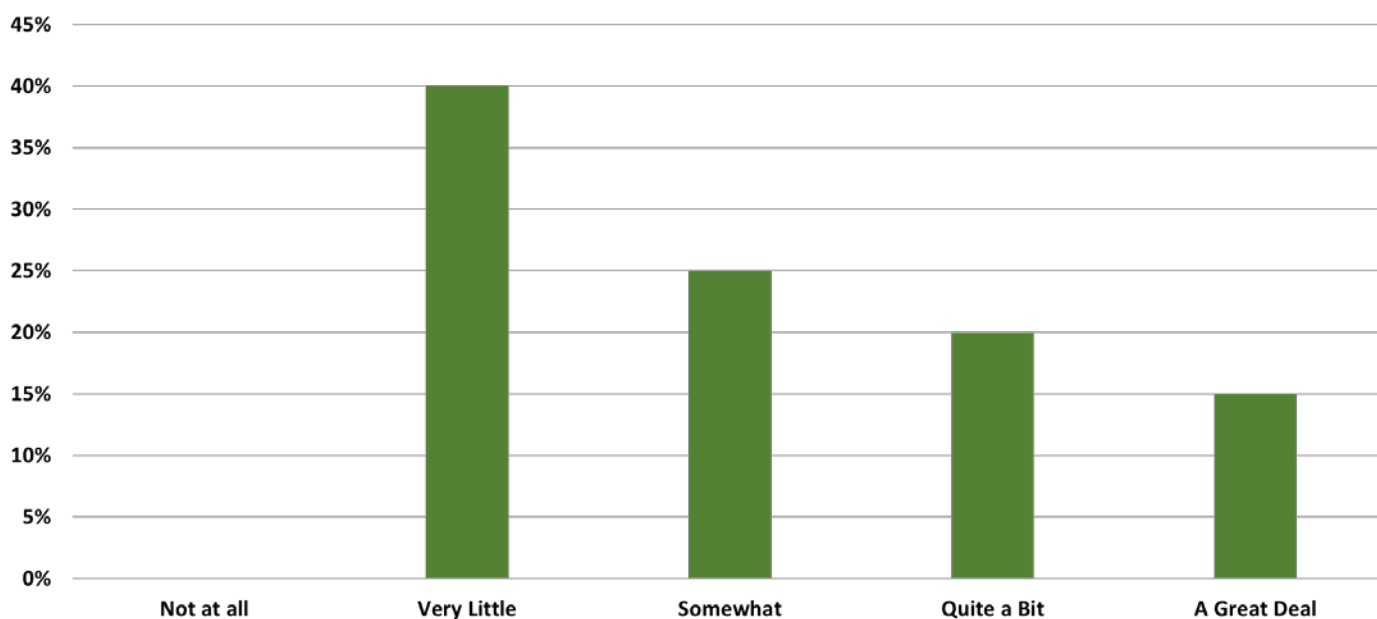
**Standardized Testing: What is it Good For?**  
*Senior Capstone Project for Megan Mapp*

---

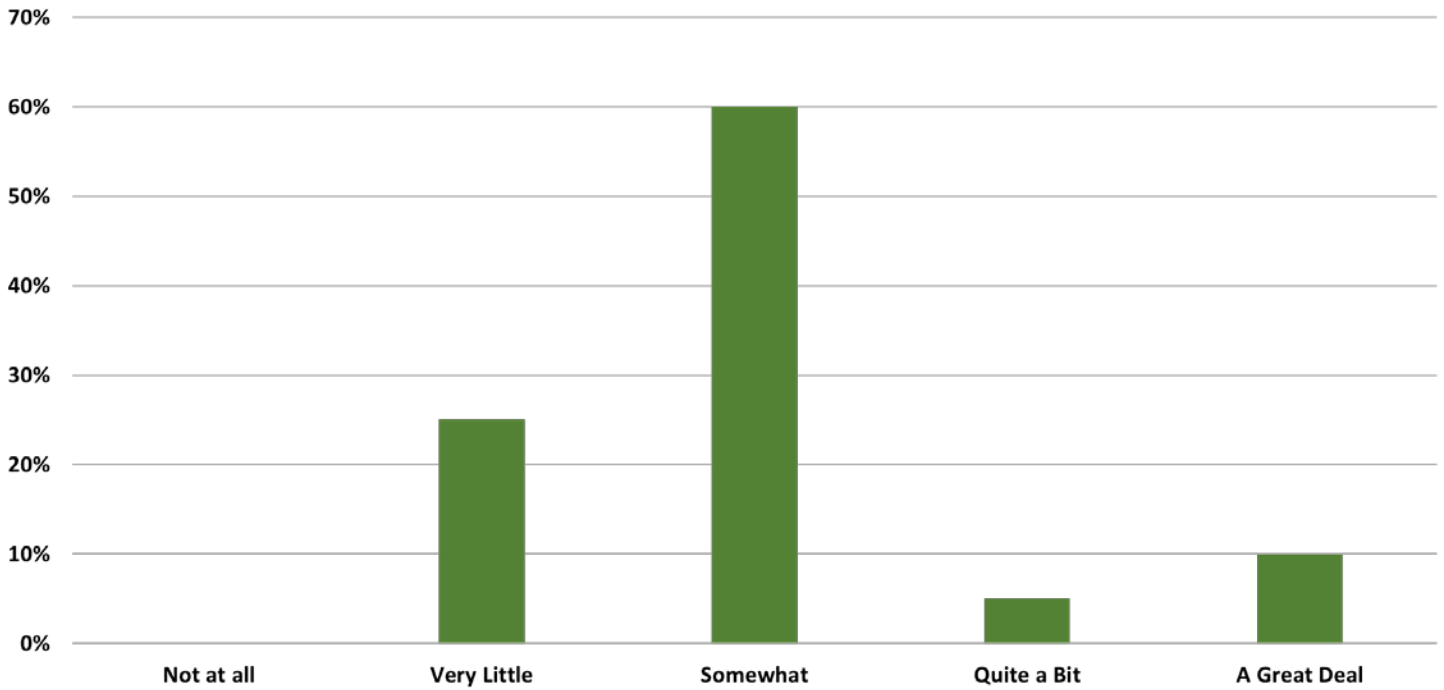
**Number of Years Teaching at Current School**



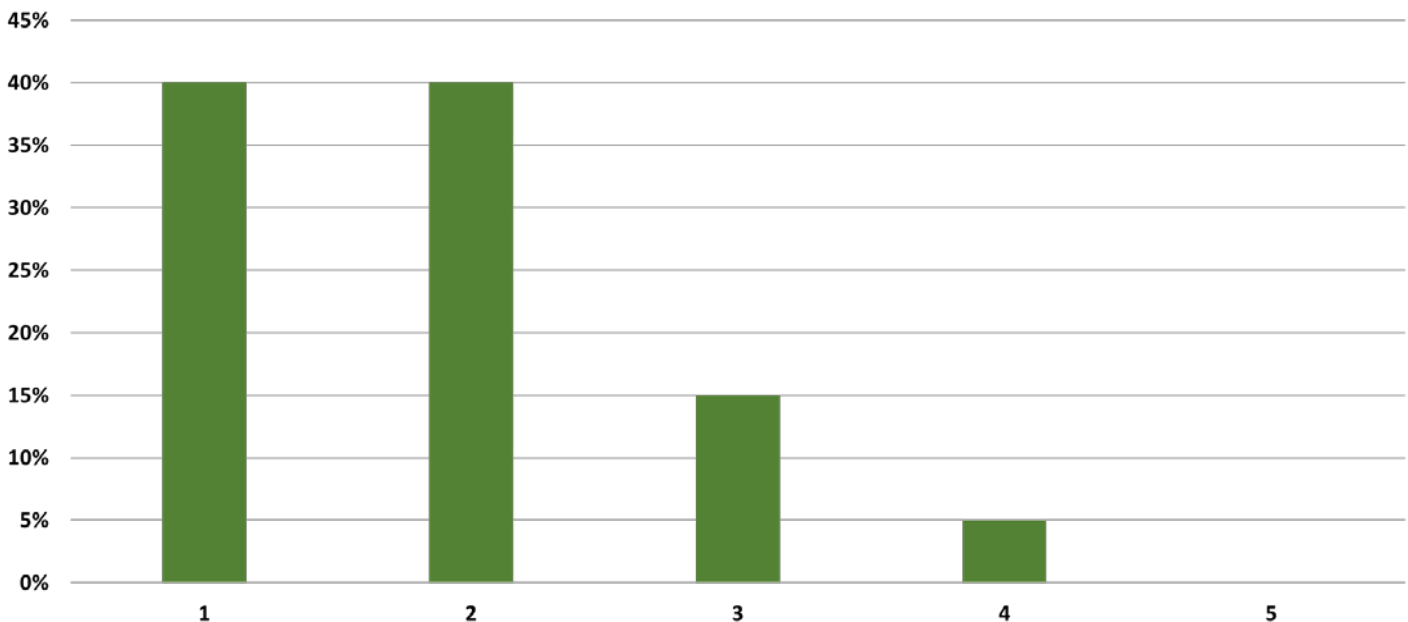
**Approximately how much classroom time of teaching is spent on preparation for state standardized tests (including directions and practice)?**



**Approximately how much personal preparation time is taken from the preparation for standardized testing?**



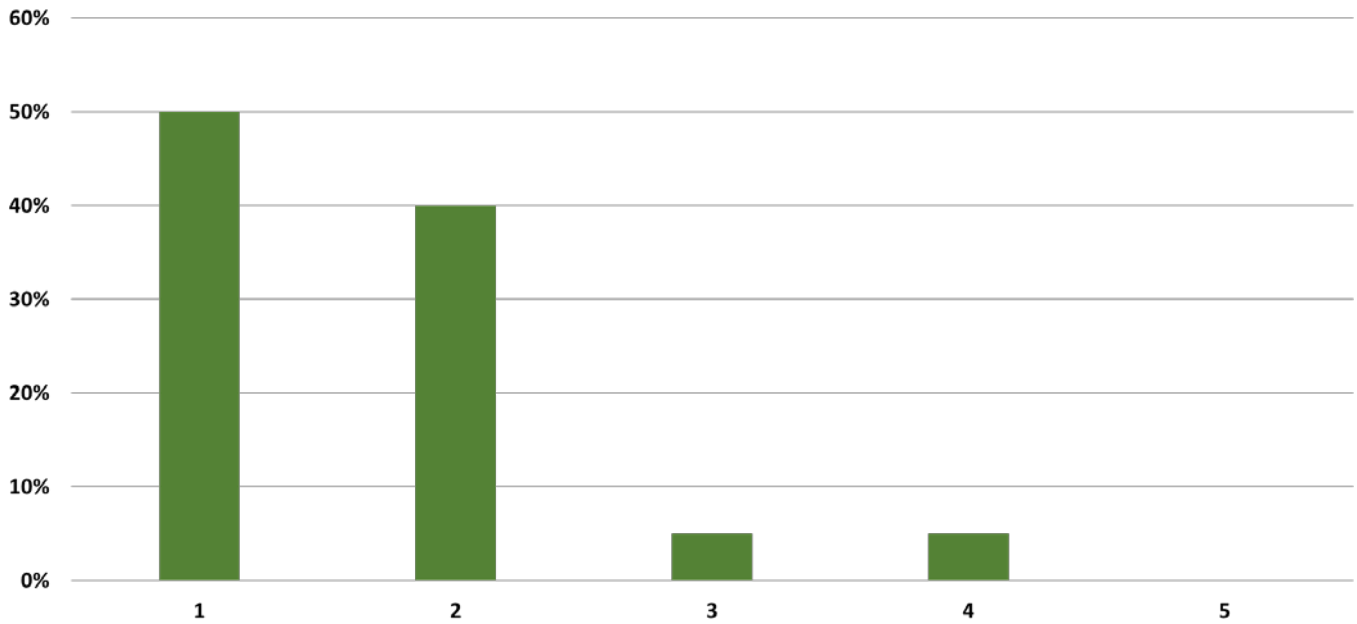
**On a scale from 1 to 5, with 1= unhappy and 5 = happy, how would you rate your happiness and overall job satisfaction as it relates to the use of standardized tests?**



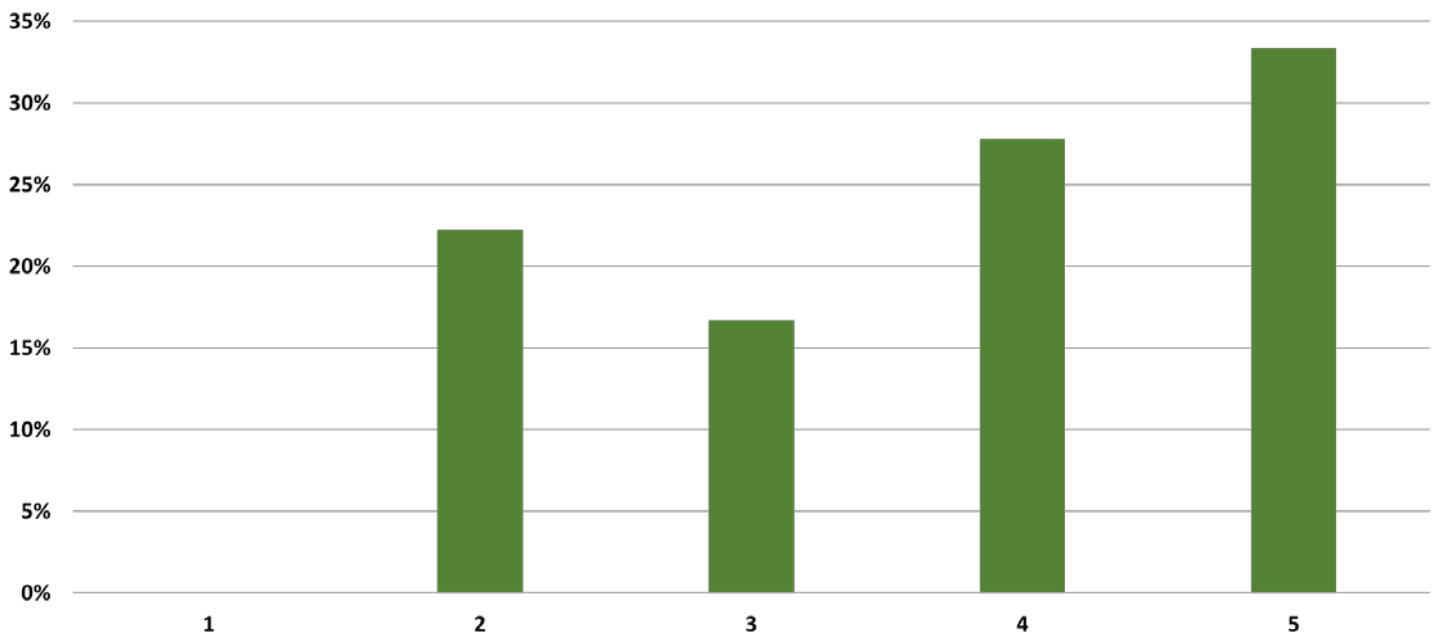
**Standardized Testing: What is it Good For?**  
*Senior Capstone Project for Megan Mapp*

---

**On a scale from 1 to 5, with 1= less and 5=more, how much time do you believe should be spent focusing on standardized testing?**



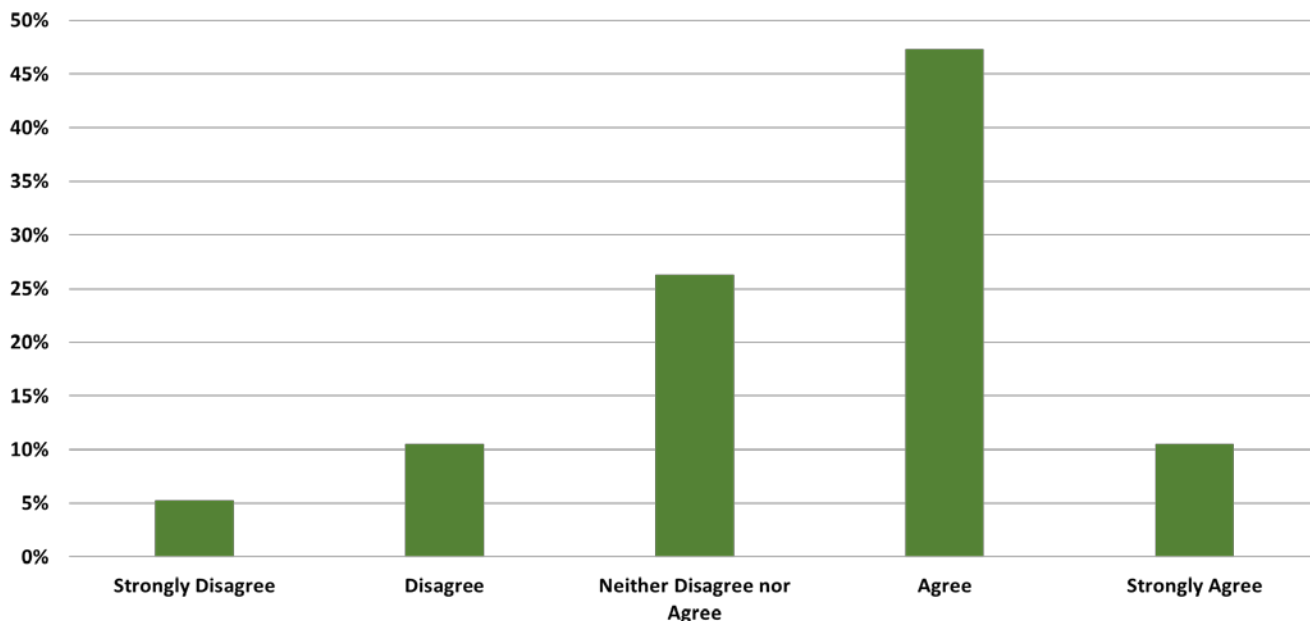
**On a scale from 1 to 5, with 1= no pressure and 5= significant pressure, do you feel pressured by state requirements to teach a certain way?**



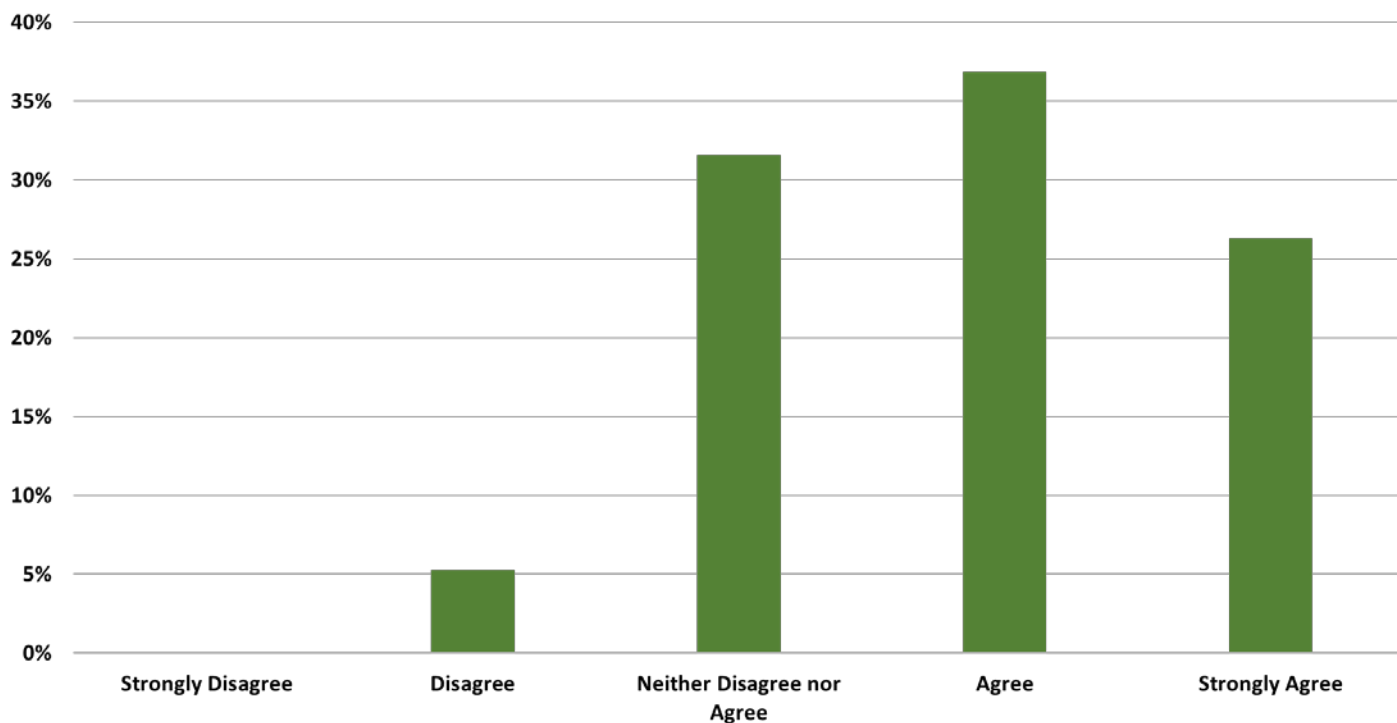
**Standardized Testing: What is it Good For?**  
*Senior Capstone Project for Megan Mapp*

---

**You have changed your teaching curriculum in response to CAPT results.**



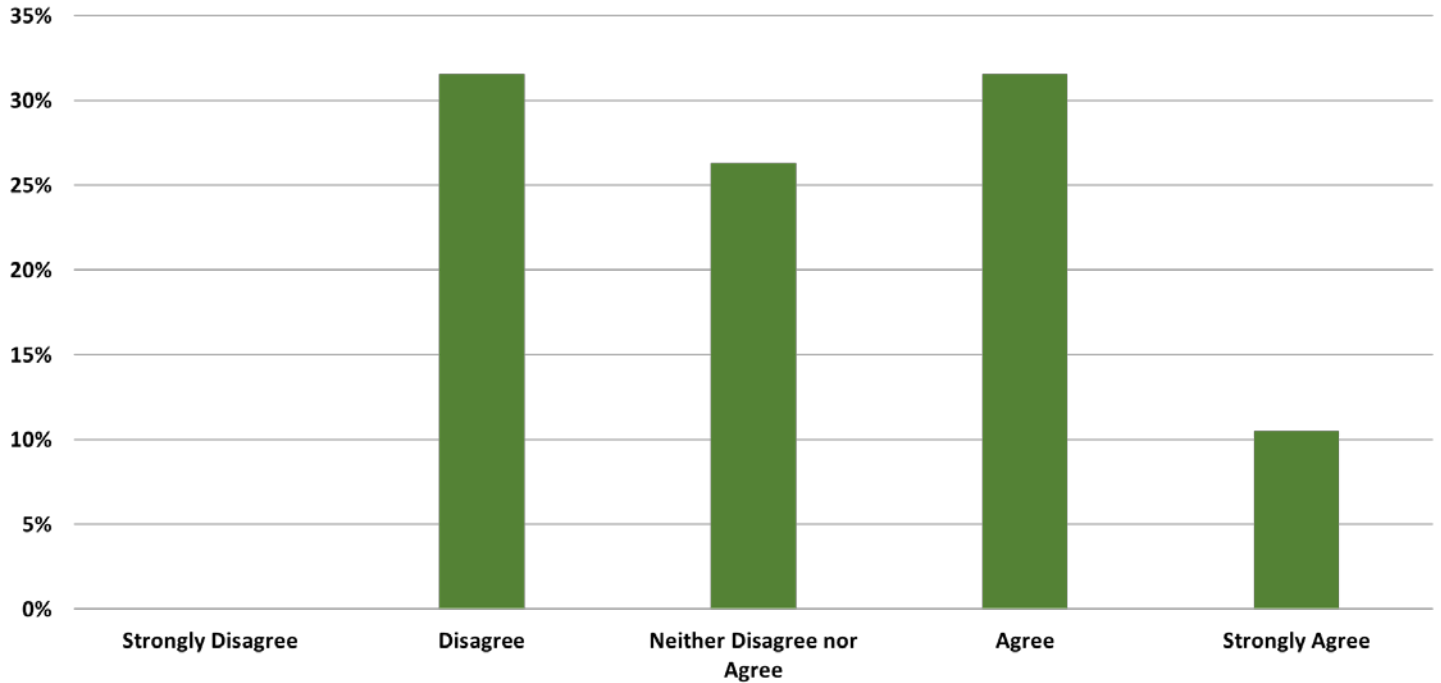
**There is a strong focus on CAPT/standardized testing in school.**



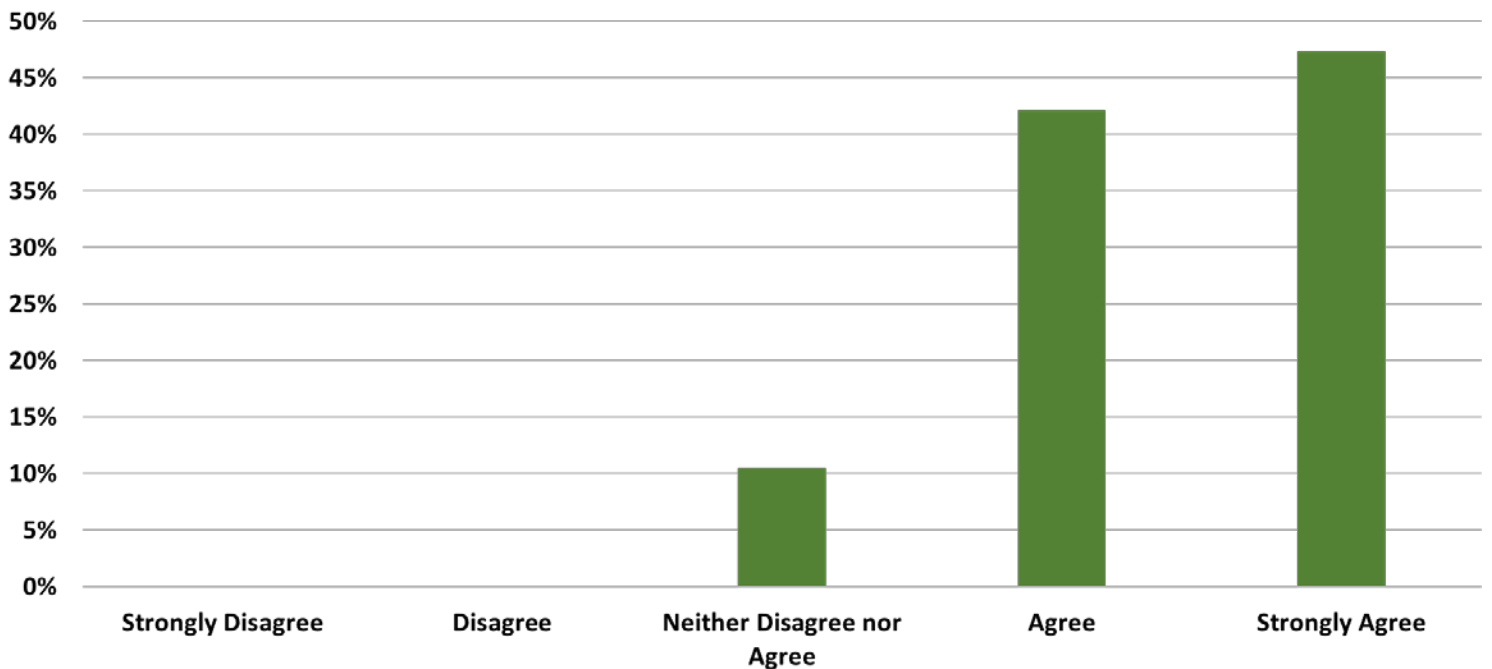
**Standardized Testing: What is it Good For?**  
*Senior Capstone Project for Megan Mapp*

---

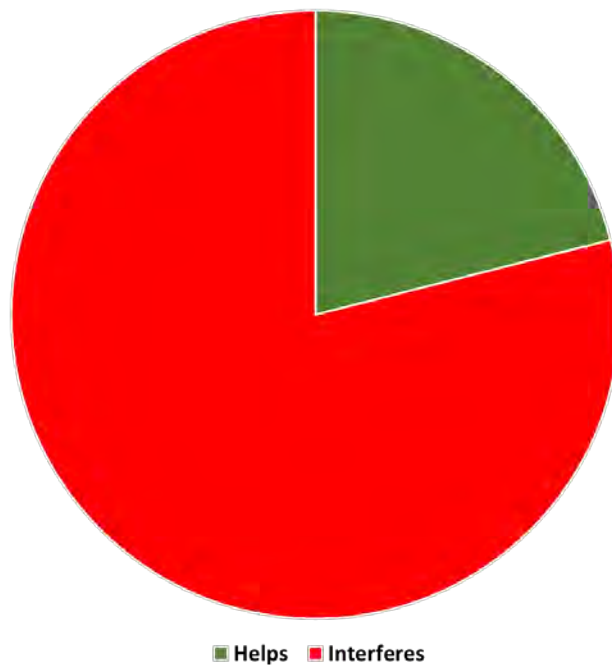
**You have a say in what curriculum changes are made by the school  
in response to CAPT results.**



**Standardized testing has added more responsibilities to your  
job.**



**Do you think the time spent and use of the CAPT standardized testing helps students or interferes with their learning?**

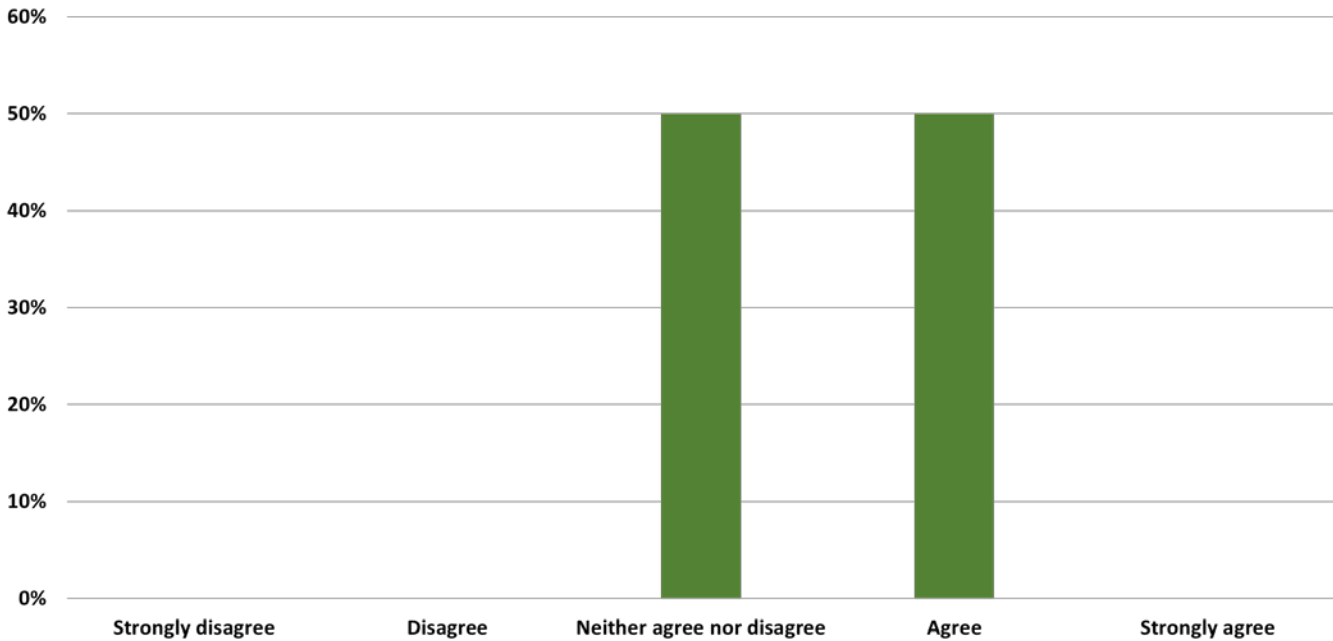




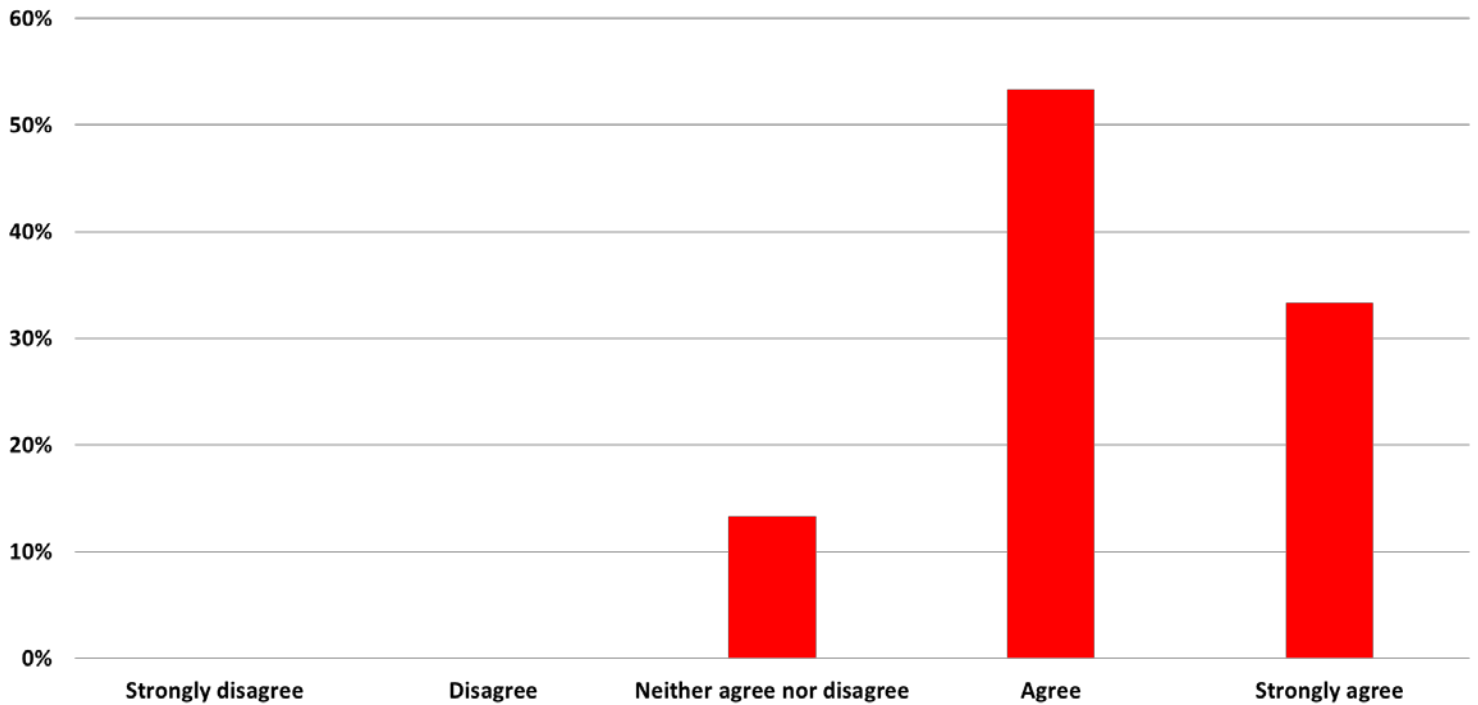
**Standardized Testing: What is it Good For?**  
*Senior Capstone Project for Megan Mapp*

---

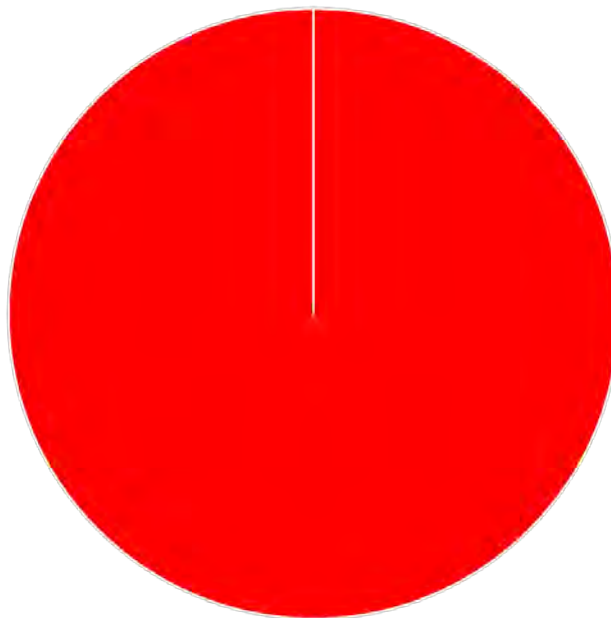
**For those who responded "Helps", to what extent do you support this statement?**



**For those who responded "Interferes", to what extent do you support this statement?**



**With the movement to the Common Core, do you believe this will be a better assessment than the CAPT or do you find similar issues will arise?**



■ Yes (will be better) ■ No (similar issues)

## Standardized Testing: What is it Good For?

### *Senior Capstone Project for Megan Mapp*

---

#### Appendix J- Teacher Comments

<p>We lose 3 block days (equivalent of 6 class periods) to MAP testing (another standardized computer based test). We lose about a week's worth of time due to SBAC testing (the 'new' CAPT). I don't mind testing our kids to see where their gaps may be, but we are spending too much class time doing it and 'worrying' about it.</p>
<p>The Common Core Standards are not being implemented properly. Content that was previously taught in Algebra I has moved to 7th and 8th grade. Students currently taking Geometry and Algebra II have gaps in their knowledge.</p>
<p>Teaching to the test is not needed. A solid curriculum with good teaching instruction is needed.</p>
<p>I am concerned that SBAC testing will be used as part of my evaluation score.</p>
<p>More testing means less time for instruction. We spend a lot of time preparing for a specific TYPE of test (Smarter Balanced), and need to prepare students for the format, including the computerized tools they'll be allowed to use. I believe this is wasted time.</p>
<p>Teachers will always be teaching to the test. As long as teacher ratings will be tied to students success on these assessments, students will lose out on real world learning to prepare for these assessments</p>
<p>The biggest challenge is not to teach to the test or adjust curriculum but to get the students to take it seriously. Last year's practice SBAC was a joke.</p>
<p>The last question... Not better and Not similar issues, but no other choices. Different issues &amp; poorly made test. The CCSS has barely been implemented and our students are going to be tested on it, using a test for which the practice materials are barely developed in time. Then, we will be evaluated using the results of the test. That is very poor planning - no wonder so many teachers feel it is a "Gotcha" situation.</p>
<p>In the middle school environment I felt much more pressure to get students prepared for CMTs. In the high school setting I mainly teach freshman and sophomores. There is not as much pressure to prepare them for CAPT (when we administered CAPT), however there is a greater push as they enter their junior year to prepare them for CAPT. With the CC type assessments we are challenging all students in all grades to think at a higher level but based on the sample assessments I have seen, I don't know that they will be prepared for the CC testing. Our high school students have not been 'raised' on the CC type of testing so it is a new experience for them being introduced just a few years before they need to take the assessment.</p>
<p>We spent so much time on CAPT and almost none on SBA. SBA testing schedule is horrible compared to CAPT and the students hate the interface. There is so much untried on it and very few are familiar. Common Core changes the way we teach and SBA is not a test you can teach to but it will disrupt our teaching for at least a month. A test that does not count yet and no student is ready for</p>
<p>Students are expressing frustration with all the testing. Testing takes time away from student learning.</p>
<p>Computer based assessments are developmentally inappropriate for students in elementary and middle schools. The CAPT and, now, the SBAC tests have no impact on our students' education and, for them, it makes no difference whether they do well or poorly on the test. Overall, all Common Core curriculum documents are impossibly abstract and bureaucratic -- aimed at a single theoretical student.</p>

**Standardized Testing: What is it Good For?**  
***Senior Capstone Project for Megan Mapp***

---

**REFERENCES**

- Amrein, Audrey, and David Berliner. "The Impact of High-Stakes Tests on Student Academic Performance: An Analysis of NAEP Results in States With High-Stakes Tests and ACT,SAT, and AP Test Results in States With High School Graduation Exams." Dec. 2002. Web. 20 Apr. 2014.
- Chubb, John. "Giving Assessment a Better Name." *Independent School* 73.2 (2014): 9–11. Print.
- Chingos, Matthew M. "Strength in Numbers: State Spending on K-12 Assessment Systems." *The Brookings Institution*. N.p., n.d. Web. 21 Apr. 2015.
- Carnoy, Martin, and Susana Loeb. "Does External Accountability Affect Student Outcomes? A Cross State Analysis." *Educational Evaluation and Policy Analysis* 204.4 (2002): 305–331. Print
- Fairtest. "The Dangerous Consequences of High-Stakes Standardized Testing | FairTest." *FairTest: The National Center for Fair and Open Testing*. N. p., 17 Dec. 2007. Web. 19 Apr. 2014.
- Herman, Joan L., and Shairi Golan. *Effects of Standardized Testing on Teachers and Learning-Another Look*. UCLA Graduate School of Education: National center for Research on Evaluation, Standards, and Student Testing, 1991. Web. 20 Apr. 2014.
- Hess, Frederick M., and Michael Q. McShane. "Common Core in the Real World." *Phi Delta Kappan* 95.3 (2013): 61–66. Print.
- "High-Stakes Test Definition." *The Glossary of Education Reform*. N. p., n.d. Web. 18 Apr. 2014.
- Holmes, Sarah. "Standardized Testing and the No Child Left Behind Act: A Failing Attempt at Reform." 2009. Web. 20 Apr. 2014.

**Standardized Testing: What is it Good For?**  
***Senior Capstone Project for Megan Mapp***

---

Linden, Stephanie. "The Impact of Standardized Testing on Student Performance in the United States." 1 May 2007. Web. 23 Apr. 2014.

Neill, Monty. "A Better Way to Assess Students and Evaluate Schools." *Education Week* 18 June 2010. *Education Week*. Web. 17 Apr. 2014.

"New Era Of Student Testing Set To Begin In Connecticut: But Anxiety And Criticism Cloud The Debut Of 'Smarter Balanced' Computerized Test." *The Hartford Courant*. N.p., n.d. Web. 23 Apr. 2015.

Nichols, Sharon, Gene Glass, and David Berliner. "High-Stakes Testing and Student Achievement: Problems for the No Child Left Behind Act-Appendices." Sept. 2005. Web. 20 Apr. 2014.

Rebora, Anthony. "Teachers Place Little Value on Standardized Testing." *Education Week* 31.26 (2012): 14–14. Print.

Schneider, Stephanie, and Matt Christison. "Are Exams Bad for Children?" *New Internationalist* 464 (2013): 30–32. Print.

Sternberg, Betty J. "Real Improvement for Real Students: Test Smarter, Serve Better." *Harvard Educational Review* 76.4 (2006): 557,563,727. *ProQuest*. Web. 21 Apr. 2015.

Strickland, Bonnie, ed. "Standardized Test." *The Gale Encyclopedia of Psychology*. 2nd ed. Detroit: Gale, 2001. 626–627. *Gale Virtual Reference Library*. Web. 19 Apr. 2014.

Thompson, Van. "Do Standardized Test Scores Factor in to How Much Money a School Will Receive?" *Everyday Life - Global Post*. N. p., n.d. Web. 21 Apr. 2014.

US Congress, Office of Technology Assessment. "Lessons From the Past:" *Testing in American Schools: Asking the Right Questions*. Washington, DC: U.S. Government Printing Office, 1992. Web. 10 Apr. 2015. OTA-SET-519.