BEYOND EXIT EXAMS: THE EFFECTS OF STANDARDIZED TESTS ON STUDENT ACADEMIC SUCCESS

By: Taylor Alette Gamez

Submitted in partial fulfillment of the
requirements for Departmental Honors in
the Department of Political Science
Texas Christian University
Fort Worth, Texas

May 4, 2015

BEYOND EXIT EXAMS: THE EFFECTS OF STANDARDIZED TESTS ON STUDENT ACADEMIC SUCCESS

Project Approved:

Supervising Professor: Vanessa Bouche, Ph.D.

Department of Political Science

Emily Farris, Ph.D.

Department of Political Science

Cynthia Williams, Ph.D.

College of Education

ABSTRACT

There is much debate over the best way to measure academic achievement of students in public schools. One of the primary methods being used is administering state-wide exams each year to measure progress. This process became federally mandated following the No Child Left Behind Act of 2001. The purpose of this paper is to discern the effects of standardized testing by asking the question: Is there a relationship between the number of state-wide standardized tests and student academic success at the state level? I rely upon the independent variables of state SAT scores and state graduation rates for the academic years of 2011 and 2012 along with the dependent variable of the total number of tests taken by a student during their years of public schooling. I collect the number of tests variable by back-tracking the grade level in which a graduate of the class 2011 or 2012 *would have been in* during each academic year to reach the total number of tests taken. The results show that an increase in testing has a negative relationship with graduation rates and a negative and significant relationship with SAT scores per state. Based upon these measurements, the conclusion is that there are no positive effects of standardized testing being seen by the students.

Table of Contents

INTRODUCTION

Promoting more effective and better quality public education is largely a bi-partisan issue, but there is significant debate about what programs and policies yield the best outcomes for the students. According to Alfie Kohn, a well-known author and lecturer in the educational realm, "Our children are tested to an extent that is unprecedented in our history and unparalleled anywhere else in the world. While previous generations of American students have had to sit through tests, never have the tests been given so frequently and never have they played such a prominent role in schooling." The shocking truth is that these tests take up large amounts of time and money while simultaneously affecting the education of students. The large issue presently is that no quantitative study has been conducted which looks at how the *number* of tests a student takes effects their future level of academic success.

Is there a relationship between the number of state-wide standardized tests and academic success at the state level? This is an important question to answer for several reasons. First, education policy varies significantly across the United States. For example, there are currently 24 states in which graduation is only permitted once a student passes the state-mandated high school exit exam (HSEE). As other states consider adopting this policy, it is important for them to know the impact of such a policy on student academic outcomes. States also vary significantly in the scope and frequency of standardized testing. For instance, Idaho administers standardized mathematics and English tests every year from grades 3 through 10, as well as science assessments for grades 5, 7 and 10. Texas, in comparison, administers mathematics and English tests from grade 3 through 11, writing tests for grades 4 and 7, science exams for grades 5, 8, 10 and 11 and social studies tests for grades 8, 10 and 11. In other words, over the course of a student's academic career in Idaho,

they will take 19 tests, whereas in Texas they will take 29.[1] As states continue to revise and reform their education policies, it is imperative that they take into consideration the impact of these standardized tests on student academic success.

The second reason this is an important question to answer is that there are serious implications beyond student success of these policies. Tests cost a significant amount of money, paid for by taxpayer dollars. A recent study found that 45 states spent a combined $669 million per year on their primary assessments with an estimate of $1.7 billion spent nationwide (Chingos, 2012). If increasing the number of tests given is either not effective, or even counterproductive, then these tax dollars should be diverted to other educational programs that are more effective. Therefore, discerning the effectiveness of testing is an important issue to parents, students and the citizenry as a whole.

There are two key hypotheses in this study. I expect to find that states with a higher number of tests administered to students during their academic careers will have a lower graduation rate. Secondly, I hypothesize that an increase in tests will have a positive effect on SAT scores per state based on exposure to testing environments and the test-retest theories. In order to test these hypothesize, I collect data on graduation rates per each state for years 2011 and 2012 from the National Center for Educational Statistics. I also collected state math and reading SAT scores for these same years from College Board. I then referred to each state education website individually in order to compile a total number of tests taken by a student who graduated in 2011 and 2012. In short, I find that an increase in the number of tests per state has a negative relationship with both graduation rates and SAT scores.

[1] In 2013, the Texas Senate approved a bill to reduce the number of standardized tests a student needs to take in order to graduate.

DEVELOPMENT AND DISCOVERIES OF STANDARDIZED TESTING

The dominance of state-wide testing expanded when President George W. Bush signed the No Child Left Behind (NCLB) Act of 2001. This act aimed to increase English and math skills across all states and meant that public schools would be held accountable for the students' success and would be subject to consequences if students did not perform well (Schiller and Muller, 2003). Specifically, the provisions of NCLB require that reading and math be tested yearly in grades 3-8 and then a minimum of one time in high school while science tests must be given a minimum of 3 times throughout a child's educational career. NCLB utilizes high-stakes testing as a means of examining if schools are making progress and students are gaining proficiency in these subjects (Lee and Reeves, 2012). Under NCLB, states must achieve "adequate yearly progress" in order to receive federal funding for their schools.

Despite the requirements of NCLB, there is no overarching national achievement standard. As of 2014, 44 states are members of the Common Core State Standards Initiative in which the member states create tests that fall in line with these standards. However, the standards for math and English were first released in 2010 which means that they were not in effect during the majority of the time period of this study.[2] In addition to the variety of standards, NCLB does not require that students pass the exams in order to advance to the next grade or graduate from high school. In other words, whether or not an individual state implements HSEEs is a decision made at the state level.

NCLB's silence on HSEEs and mandatory passage is interesting given the research on these issues. When passing NCLB-mandated exams is required by the state in order to advance to the next grade, Warren, Jenkins and Kulick (2006) find that HSEEs lead to an

[2] All policies mentioned were current during the time of the compiling of this study.

increased number of GED tests and lower high school completion rates. More specifically, they find an additional 2.1% of students do not graduate when HSEEs are in place. These exams act as a final hoop that students must jump through in order to complete their high school careers. However, for students who do not initially pass these exams on their first attempt, it does not appear to have a significant effect on achievement, persistence or graduation rates if the student was near passing score (Reardon *et al,* 2010). They find that failing the math portion of state tests lowers the chance that a student will take a course higher than geometry in the 11^{th} grade, a course typically taken in 9^{th} or 10^{th} grade, but it does not have an effect on their score when the take the exam again in the 11^{th} grade. States in which such high-stakes exit exams are in place see a higher dropout rate if they do not have alternate routes for students to receive a diploma such as relying on overall GPA or previous test scores (Hemelt and Marcotte, 2013). In addition, a report conducted for the National Board of Education prior to NCLB found that the states with the highest dropout rates were using minimum competency testing as a means of high stakes testing (Clarke, Haney and Madaus, 2000).

One of the outcomes following the implementation of NCLB was that both district spending and total per-pupil expenditure increased in order to meet the achievement requirements set forth by the act (Dee, Jacob and Schwartz, 2012). And given that previous literature has found that expenditure per student is positively related to graduation rates, test scores and college continuation (Daun-Barnett and St. John 2012), it can be said that both NCLB and the subsequent increased spending are important to educational achievement. In other words, if increasing testing leads to an increase in overall expenditure, and an increase in spending has also been linked to an increase in in test scores, graduation rates and college

continuation, examining the relationship in another manner may be able to give an answer as to whether or not NLCB is having a positive effect on students.

While these tests have been discussed in the context of teacher and school accountability, questions about the impact of standardized testing on the academic success of the students seems to be an afterthought. Much of the literature that exists analyzes how teachers adapt to standardized testing (Marchant and Paulson, 2005), how "high stakes" standardized testing encourages teachers to "teach to the test" and to spend less time teaching other subjects that are not being tested (Dee, Jacob and Schwartz, 2012), how they are rewarded or punished based on their students' performances, and how public policy makers may benefit (Holme and Heilig, 2012). There has been much less attention on the impact of standardized testing on the students.

Those studies that have examined the impact of standardized testing on the students have focused on the personal/emotional impact, such as increased stress, anxiety and self-esteem issues (Putwain *et al* 2012), while others have focused specifically on the impact of mandatory high school exit exams on graduation rates (Marchant and Paulson, 2005; Warren Jenkins and Kulick, 2006; Hemelt and Marcotte, 2013) and college continuation rates (Daun-Barnett and St. John, 2012). This research that examines the effects of testing on student outcomes motivates the present study, the purpose of which is to assess how the total number of standardized tests administered in public schools, per state, from grades K-12 impacts the academic success of students.

This evidence suggests that, if these state policies are in place to limit graduation of students who are not meeting minimum requirements, it would only make sense that an increase in testing overall, whether exit exams or not, would decrease graduation rates.

However, if these negative findings by scholars are true (i.e. that standardized tests increase stress levels of students and cause teachers to teach to the test and limit curriculum), I would argue that there must be some type of positive outcome in respect to the benefit of the students for implementing more testing. Otherwise, it would raise many questions about the purpose and goals of standardized testing.

Sorting out these results to determine the effects of testing on academic outcomes is therefore important. Based on the balance of the studies reviewed above, I hypothesize that:

(H1) Implementation of more standardized testing throughout a student's years of schooling decreases the graduation rates per state.

STANDARDIZED TESTS AND SAT SCORES

In addition to graduation rates, another measurement of the educational attainment of students is standardized tests such as the SAT. There have been studies that examine how exit exams influence SAT scores (Marchant and Paulson 2005), but none have looked at how the number of standardized tests, whether exit level or not, have influenced children's success or improvements on SAT scores. Since the SAT is one of the most widely administered college entrance examinations (Schneider, 2003), it is important to explore the factors that influence these scores. With evidence that both high school GPA and SAT scores are indicators of future academic achievement of students (Hoffman and Lowitzki, 2005), it is important to discover if there is a positive relationship between standardized testing and average SAT scores in order to measure and grow academic success. The evidence thus far is conflicting.

Some evidence suggests that standardized tests have a positive impact on college readiness. Daun-Barnett and St. John (2012) found that policies mandating exit exams increase college continuation rates 2 percentage points. Qualitative evidence from teachers suggests that exit exams benefit students' academic preparation for college, and that the skills learned in preparation for these exams may be transferred to college admission exams (Perna and Thomas 2008). According to Butler and Roediger (2007), students who are exposed to repeated testing exhibit higher long-term retention rates than those students who only study the same material.

This evidence suggests that the repeat testing methods of standardized tests may allow for students to better recall the information learned in school, which can then be used on other exams in the future. This test and retest method has proven to be effective at both the collegiate level as well as the 3rd grade level (Lipowski, Pyc, Dunlosky, and Rawson 2013). Studies of high school students have also shown that testing students in this manner leads to better application of principals and procedures on following tests (Dirkx, Kester and Kirschner 2014). Based on these types of psychological as well as educational based research, it appears that an increase in testing leads to better performance on future tests. This suggests that state administered tests may increase students' preparedness for standardized tests in general as well as their pathways to college.

On the other hand, there is contradictory evidence that exit exams and standards-based reforms are not related to SAT scores (Musoba, 2011), or are correlated negatively. Specifically, Venezia and Jaeger (2013) found in a study based on SAT data from 2012 that only 43% of all SAT takers met the SAT College & Career Readiness Benchmark, which indicates a 65% chance of earning a "B-" average in the first year of college.

The verdict is still out as to the impact of standardized testing on college entrance exams. This study investigates this question in the context of the total number of standardized tests a student takes over the course of their academic career prior to college. Does the sheer number of standardized tests have a significant impact on SAT scores either positively or negatively?

I hypothesize that: *(H2) an increase in the number of tests taken throughout a student's school years will lead to an increase in state-level SAT scores* due to exposure to standardized testing environments.

METHODS

In order to test my hypotheses, I examine the state testing policies across all 50 states in a two year period. There are at least three reasons why an investigation of all states over a period of time is important. First, all state policies are unique and there is a mixture of states with and without implementation of HSEEs. Secondly, looking at this two year period gives me the ability to track changes in testing policies following the implementation of NCLB. Changes in testing policy of those states that have mandated HSEEs will be able to be compared to other states that may or may not have changed their standardized testing policies over time. Overall, tracking these changes over time with the most recent data available will provide more up-to-date insight on how standardized testing affects education attainment of students. Finally, there was not adequate data available in order to include more recent years in this study.

My dependent variable is academic success, which I operationalize in two ways: graduation rates and SAT scores. I collected the graduation rates of all 50 states for the

academic years of 2011 and 2012 from the National Center for Education Statistics.[3] Specifically, I collected the public high school 4-year adjusted cohort graduation rate which "measures of the percent of students that successfully complete high school in 4 years with a regular high school diploma."[4] Second, I gathered the mean score of public school students for the critical reading and math sections of the SAT for all 50 states for the years 2011 and 2012 which was provided by College Board. I then combined the critical reading and math scores to give the SAT score variable for each state per each year. This adds an important component of control to my data analysis because the SAT will serve as a common basis of measure. The SAT is very similar each time it is administered, regardless of the state in which it is taken, which is a large contrast to the variety of standardized testing policies per state, and, according to the National Center for Education Statistics, nearly half of all graduates take the SAT.[5]

My main independent variable of interest is the total number of standardized tests a student will take over the course of their academic career if they graduated from high school in 2011 and 2012. I am assuming that the students who graduated in these two years attended public schooling from grade 3 through grade 12 in the same state in which they graduated. Given this assumption, I am looking for testing information from the year of 2002 through 2012, as a student graduating in the class of 2011 would have been in the 3rd grade during the spring of 2002 and testing prior to grade 3 is not common. The time period for this research is also of interest because NCLB was not in effect until the spring of 2003.[6] This means that

[3] The states of Oklahoma and Kentucky were granted extensions on publicizing their graduation rate data by the U.S. Department of Education and were not able to be included in this study.
[4] Institute for Education Sciences for the National Center for Education Statistics
[5] Although some may argue that the SAT is subjective because it is only available to those who can afford the test, College Board administers fee waivers for families that qualify.
[6] The majority of standardized testing is conducted in the spring semester.

the two graduating classes selected for this study will over-lap with the changes in NCLB policies in order to track how states responded to the new legislature.

In order to collect this testing information, I had to refer to each state education websites individually. There is no central location of data for standardized testing on the United States Department of Education website, even though the tests results are sent to the federal government by law. Only a small number of states had data available on their websites that covered the time period of 2002-2012. The rest of the information had to be obtained by communicating directly with representatives of each state via phone calls and emails.

There were seven states for which data about their state-wide standardized testing policies could not be obtained.[7] I attempted to contact each state three times but was unable to receive a response in order to collect information about their testing policies during the time constraint of my data collection period. For this reason, these states were not able to be included in my analysis. Although this lowers the number of observations for the study, I do not believe that these states are systematically different than the other 42 states that have been included and therefore do not believe that the exclusion will have a significant effect on the results.

In addition to the main independent variable of interest, I also control for a variety of factors that may also impact academic success. For this analysis I am going to control for spending per pupil, percent of students receiving reduced or free lunch, student to teacher ratio and whether or not the states are administering HSEEs. The demographic of each state

[7] The 8 states not included are: Hawaii, Iowa, Mississippi, Nevada, Oklahoma, Tennessee and West Virginia

is a very important component to consider when analyzing both graduation rates and student performance on tests in order to be sure that comparisons are accurate and not being affected by other variances within each state. For instance, those states which spend more money per student and have a lower student to teacher ratio have a systematically different approach to education than those states that do not spend as much money educating their children and have larger classrooms with less individual attention. Also, states with a higher amount of the student body qualifying for free or reduced lunch indicates that those families are of lower income which it generally linked to lower test scores and a lower chance of graduation. Finally, I included a control for the percent of the population under the age of 18 in each state that fell into the following racial categories: white, Latino, Asian, and black. This control increased the measure of variability in my model, as indicated by the r-squared values of my results in Table 2.

In order to test my hypotheses, I stack the data so that each observation is one state in one year, constituting a total of 84 observations. I ran ordinary least squares regression analysis because both dependent variables--graduation rates and SAT scores--are continuous variables. Running the same statistical tests on both variables also allows for comparison across the models of the impact of the independent variables on the two different measurements of academic success.

RESULTS

The first step to analyzing the results of this study was to run descriptive statistics of the main independent variable (the number of standardized tests per state), the two dependent variables (SAT score per state and graduation rate per state), and the control variables (percent free lunch, student-teacher ratio, spending per pupil and HSEE). The results indicate

that there is a large variety of testing policies among the states included. For example, Nebraska had a variety of testing done at the district and local level, but only had 4 state-wide exams during the years of this study which gives a minimum value recorded while the maximum value of 36 tests was recorded for the state of Georgia. In addition, the table shows that there is a difference of nearly 30 percentage points between the state with the highest rate of graduation and the state with the lowest graduation rate during these years. These results can be seen in **Table 1**.

[INSERT TABLE 1]

According to the information collected, the total number of standardized tests for the 43 states for which data was available was 756 for the year 2011. One year later, a product of changes in testing policies to align with the government mandate of NCLB, the total number of tests for the same set of 43 states was 836. Looking at the averages once more, for the same 43 states with available information, the average SAT score (math and reading combined) for the year 2011 was 1068.7. In the year 2012 the average SAT score for these 43 states was 1066.2. Presented in this way, we can see that in one year there was nearly a 10% increase nationwide in respect to standardized testing and negative, if any, change in the SAT scores between the two years.

The results for my first hypothesis--that increasing standardized testing will lead to a decrease in graduation rates--can be found in Table 2. Of particular note is that the coefficient for number of standardized tests is in the predicted negative direction; however, the relationship is not statistically significant. In other words, there is no impact of number of standardized tests on the graduation rates. One reason for these results may stem from the

fact that nearly half of the states included in the study implement high school exit exams. A second possibility for the lack of a significant correlation between the number of tests and the graduation rate per state could be that the rate provided simply looks at those students who completed *high school* on time. In other words, there is not measurement for students who were held behind in elementary and middle school from either not passing the exams or not obtaining a mastery of the skills being tested which could be having an effect on the graduation rate.

Also noteworthy is that the HSEE variable has a negative relationship with the graduation rate variable but is also not statistically significant. This is contradictory to the findings of Warren, Jenkins and Kulick from 2006. The variation in the results of this study and the results of previous literature could result from the difference of the years included in the study. On the other hand, the control variables student to teacher ratio and spending per pupil were each found to be statistically significant, and in the negative direction. Interpreting these results indicates that an increase of one more student per teacher has an effect of nearly one percentage point decrease in the overall graduation rate. Interestingly, of each of the controls for race, only the percent Asian variable had a statistically significant relationship with the graduation rate variable.

The results of the second hypothesis are found in Table 2. Contrary to my original hypothesis, the analysis shows that an increase in testing leads to a statistically significant *decrease* in the overall SAT average for the state. Specifically, for every increase of 1 standardized test per state, the state average SAT score decreases by just over 5 points (5.1).

Beyond the number of standardized tests, states that have HSEEs have a statistically significant decrease of 45 points on the average SAT scores for the state. A possible explanation for this relationship is that more focus is being paid towards these exit exams than preparation for college entrance exams in these states. When looking at the control variables, states with a higher percentage of Asian students have a 5-point higher average SAT score but this was not found to be significant. States with higher black and Latino percentages see a significant decrease of 2.2 and 3.6 points on the SAT per each increase of 1%, respectively. Finally, the free lunch and student to teacher ratio variables each had a positive relationship with the SAT variable but were not significant.

DISCUSSION

The results based on the data available pose a very interesting question. If an increase in standardized testing in a given state results in a decrease in graduation rates as well as a statistically significant decrease in the average SAT score for the state, why then, have we seen an upward trend in testing? It appears that increasing testing is not for the benefit of the students, and actually adversely impacts the students.

One possible reason for the lack of a positive relationship between testing and SAT scores could be that teachers are "teaching to the tests" for the state and not preparing students for other types of exams. Dee, Jacobs and Schwartz mention that they found that teachers narrow their curriculum to focus on test preparation (255). Shifting the focus of the classroom from purely curriculum to testing strategies for a specific state test may also be the reason that states with high school exit exams in place have a lower SAT score by an average

of 49 points. It is possible that there is more focus on passing these exams in order to graduate than there is to thoroughly learn and understand the class material.

Another explanation is that the purpose behind standardized testing policies is not to benefit or help the students. One might claim that state-wide testing instead is in place solely so that each state can monitor and track the progress of each of its school districts in a uniform way that can then be reported back to the federal government. In this way, standardized testing is used as a measure of achievement for teachers and administrators rather than for the school children. Since federal funding of public education is also based on achieving adequate yearly progress of the proficiency of students at each grade level, it becomes important for teachers to focus on student performance on the state test in order to ensure that the state keeps its funding and that the teachers keep their jobs. A statement from the work of Duffy *et al* (2009) reiterates this claim by saying that "mandated assessments and high-stakes tests affect the progress and future of individual students; the reputations, salaries, and careers of teachers; and the reputations and status of schools" The results of this study seem to indicate that high numbers of high-stakes tests may have the unintended consequence of affecting student progress negatively, even as they have the potential to positively impact school funding and/or teacher/administrator accountability.

However, section 1001 of the statement of purpose for the NCLB act claims that the purpose of this mandate is to "ensure that all children have a fair, equal, and significant opportunity to obtain a high-quality education and reach, at minimum, proficiency on challenging state academic achievement standards and state academic assessments." This statement uses language that insinuates that testing is not only conducted to monitor proficiency and achievement but also to benefit the educational attainment of the students.

Unfortunately, the results of this study show that the educational attainment based on graduation rates from high school and scores on the SAT test are being negatively affected by the implementation of more standardized testing.

Parents, students and tax payers should be concerned. Large sums of money are delegated to the states for public education in hopes of preparing students for the future and producing competent and capable members of the workforce. Yet, the testing that is required of these students appears to be costly monetarily and academically. This raises a number of questions that policymakers must consider. Is there a better way to ensure that students are learning the required material, possibly on a less frequent scale? Should we rethink the way that states and schools receive their funding as to not let the allocation of money be on the shoulders of students? And, finally, what other methods of classroom learning may be effective in increasing both graduation rates and SAT scores?

Undoubtedly, the results of this study indicate that further research should be done on this topic. Policy makers as well as political scientists should look at this same type of data over longer periods of time to see if this is merely a recent trend or if testing has continued to rise while graduation rates and SAT scores have stagnated. Given the lack of more recent data for this study, it would also be of interest to conduct a similar analysis in the future in order to continually monitor the results of NCLB.

TABLE 1

VARIABLE	DESCRIPTION	MEAN	STANDARD DEVIATION	MIN	MAX
GRADUATION RATE	Percent of students graduated from public high school	**79.76**	**6.119**	**62**	**93**
SAT SCORE	Combined Math and Reading SAT score	**1068.05**	**90.718**	**883**	**1241**
NUMBER OF TESTS	Number of state-wide standardized tests	**18.512**	**5.81**	**4**	**36**
% FREE LUNCH	Percent of students eligible for free or reduced lunch	**42.718**	**10.493**	**17.2**	**70.6**
STUDENT:TEACHER	Average number of students per one teacher	**15.215**	**2.871**	**9.2**	**25.6**
SPENDING PER PUPIL	Average spending per pupil in dollars	**10802.32**	**3046.506**	**1145**	**19552**
HIGH SCHOOL EXIT EXAM	Binary- 1= HSEE in effect. 0= no HSEE	**.469**	**.5016**	**0**	**1**
PERCENT ASIAN	Percent of the state population under the age of 18 that is classified as Asian	**3.966**	**6.232**	**.707**	**51.77**
PERCENT BLACK	Percent of the state population under the age of 18 that is classified as Black	**13.73**	**12.391**	**.973**	**49.87**
PERCENT LATINO	Percent of the state population under the age of 18 that is classified as Latino	**14.87**	**13.677**	**.098**	**59.49**
PERCENT WHITE	Percent of the state population under the age of 18 that is classified as white	**61.99**	**18.098**	**14.46**	**98.82**

TABLE 2

INDEPENDENT VARIABLES	GRADUATION RATE	SAT SCORE
TESTS	-0.071	-5.104***
	(0.134)	(1.49)
FREE LUNCH	-0.07	1.401
	(0.08)	(1.13)
STUDENT TEACHER	-0.88***	3.318
RATIO	(0.32)	(4.58)
SPENDING PER PUPIL	-0.000**	-0.01***
	(0.000)	(0.004)
HSEE	-1.08	-45.68*
	(1.62)	(23.08)
ASIAN	1.132**	5.376
	(.449)	(5.68)
BLACK	0.283***	-2.29**
	(0.052)	(1.05)
LATINO	0.375***	-3.4***
	(0.075)	(1.22)
WHITE	0.488***	-1.091
	(0.081)	(1.36)
N	82	40
R^2	0.5097	0.30

*** P≤ .01
** P≤ .05
* P≤ .10

REFERENCES

Butler, A. C., & Roediger, H. L. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, 19, 514–527.

Clarke, Marguerite, Haney, Walter and George Madaus. "High Stakes Testing and High School Completion" 2000. *The National Board of Education Testing and Public Policy* 1(3): 2-13.

Chingos, Mathew M. "STRENGTH IN NUMBERS State Spending on K-12 Assessment Systems" 2012. *Brown Center on Education Policy.* 1-41.

Daun-Barnett, Nathan and Edward P. St. John. "Constrained Curriculum in High Schools: The Changing Math Standards and Standard Achievement, High School Graduation and College Continuation" 2012. *Education Policy Analysis Archives* 20(5): 1-25.

Dee, Thomas S., Jacob, Brian and Nathaniel L. Schwartz. "The Effects of NCLB on School Resources and Practices" 2013. *Education Evaluation and Policy Analysis* 35(2): 252-279.

Dirkx, Kim J. H., Liesbeth Kester and Paul A. Kirschner. "The Testing Effect for Learning Principles and Procedures from Texts" 2014. *The Journal of Educational Research* 107: 357-364

Duffy, Maureen, Victori A. Giordano, Jill B. Farell, Oneyda M. Paneque and Genae B. Crump. "No child left behind: values and research issues in high-stakes assessments" 2009. *Counseling and Values* 53(1): 53+.

Hemelt, Steven W. and Dave E. Marcotte. "High School Exit Exams and Dropout in an Era of Increased Accountability" 2013. *Journal of Policy Analysis and Management* 32(2): 323-349.

Hoffman, John L. and Katie E. Lowitzki. "Predicting College Success with High School Grades and Test Scores: Limitations for Minority Students" 2005. *The Review of Higher Education* 28(4): 455-474.

Holme, Jennifer J. and Julian V. Heilig "High-Stakes Decisions: The Legal Landscape of High School Exit Exams and the Implications for Schools and Leaders" 2012. *Journal of School Leadership* 22: 1177-1197.

Lee, Jaekyung and Todd Reeves. "Revisiting the Impact of NCLB High-Stakes School Accountability, Capacity and Resources: State NAEP 1990-2009 Reading and Math Acievement Gaps and Trends" 2012. *Education Evaluation and Policy Analysis* 34(2): 209-231.

Lipowski, Stacy L., Mary A. Pyc, John Dunlowsky and Katherine A. Rawson. "Establishing and Explaining the Testing Effect in Free Recall for Young Children" 2013. *Development Psychology* 50(4): 994-1000.

Marchant, Gregory J. and Sharon E. Paulson. "The Relationship of High School Graduation Exams to Graduation Rates and SAT Scores" 2005. *Education Policy Analysis Archives* 13(6):1-17.

Musoba, Glenda D. "Accountability Policies and Readiness for College for Diverse Students" 2011. *Educational Policy* 25(3): 451-487.

Perna, Laura W. and Scott L. Thomas. “Barriers to College Opportunity: The Unintended Consequences of State-Mandated Testing” 2008. *Educational Policy* 23(3): 451-479.

Putwain, Dave W., Liz Connors, Kevin Woods, and Laura J. Nicholson. “Stress and Anxiety Surrounding Forthcoming Standard Assessment Tests in English schoolchildren” 2012. *Pastoral Care in Education* 30(4): 289-302.

Reardon, Sean F., Arshan, Nicole, Atteberry, Allison and Michael Kurlaender. “Effects of Failing a High School Exit Exam on Course Taking, Achievement, Persistence, and Graduation” *Educational Evaluation and Policy Analysis* 32(4): 498-520.

Schiller, Kathryn S. and Chandra Muller. “Raising the Bar and Equity? Effects of State High School Graduation Requirements and Accountability Policies on Students’ Mathematics Course Taking” 2003. *Educational Evaluation and Policy Analysis* 25(3): 299-318.

Schneider, Barbra L. “Strategies for Success: High School and Beyond” 2003. *Brookings Papers on Education Policy* 55-79.

Venezia, Andrea and Laura Jaeger. “Transitions from High School to College” 2013. *The Future of Children* 23(1): 117-136.

Warren, John R., Jenkins, Krista N. and Rachael B. Kulick. “High School Exit Examinations and State-Level Completion and GED Rates, 1975 through 2002” 2006. *Educational Evaluation and Policy Analysis* 28(2): 131-152.