1-1-2009

High Stakes Testing Effects on Graduation Rates

Karen Gail Cummings
mckinney17@marshall.edu

Running Head: HIGH STAKES TESTING

High Stakes Testing Effects on Graduation Rates

Thesis submitted to
The Graduate College of
Marshall University

In partial fulfillment of
the requirements for the degree of
Educational Specialist
in the School Psychology Program

By

Karen Gail Cummings, M.A.

Fred Jay Krieg, Ph.D., Committee Chairperson
Sandra S. Stroebel, Ph.D
Stephen O'Keefe, Ph.D

Marshall University

Spring 2009

ABSTRACT

High Stakes Testing Effects on Graduation Rates
by Karen Gail Cummings

With the inception of No Child Left Behind many states are now requiring students to pass a statewide exam in order to be promoted to the next grade or to graduate from high school. Opponents of these tests argue that high stakes testing decreases graduation rates among these students. They also argue that the tests are biased and unfair for minority students, students of low socioeconomic status, and students with disabilities who often fail the tests. The graduation rates for twenty-nine Appalachian counties in Ohio, a state that requires students to pass the Ohio Graduation Test in order to graduate, and Appalachian West Virginia, which recently implemented the statewide WESTEST but does not use the test as a requirement for graduation, were examined to determine if high stakes testing had an effect on graduation rates. Results indicated that implementation of high stakes testing did not have a significant effect on graduation rates.

ACKNOWLEDGEMENTS

First, I would like to thank Dr. Fred Jay Krieg, Dr. Sandra Stroebel, and Dr. Stephen O'Keefe for your wisdom, guidance, and support throughout this project and also during my graduate years in the School Psychology Program. To my family and friends whose belief in me gave me the motivation to succeed. Finally, I would like to express my deepest gratitude to my husband, Alan. His self-sacrifice, encouragement, and unconditional love helped me achieve what was once, only a dream.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

REVIEW OF LITERATURE

Before the 1970's very few states required students to pass an exam in order to be promoted to the next grade or graduate from high school. However, as the demand for public school accountability increased, so did the prevalence of high-stakes testing. The underlying idea behind high-stakes testing is that withholding promotion or graduation will create an incentive for low-performing students and schools to improve their performance (Reardon & Galindo, 2002). By 1993, twenty-four percent of school districts used high-stakes tests for school accountability (United States General Accounting Office, 1993). Opponents of high-stakes testing believe that creating additional hurdles for students to cross may increase dropout rates especially for minority and low-income students already at risk for dropping out (Reardon & Galindo, 2002; Swanson, 2003) .

A New Era for Public Schools

On January 8, 2002 President George W. Bush signed the No Child Left Behind Act which drastically changed public education in the United States. The NCLB Act reauthorizes the Elementary and Secondary Education Act of 1965 (ESEA) and requires states to implement statewide accountability systems covering all public schools and all their students. Accountability is based on State standards in reading and mathematics, annual testing for all students in grades 3-8 and once in high school, and annual statewide progress objectives ensuring that all groups of students reach proficiency within twelve years (US Department of Education, 2004). Accountability is measured by student scores on statewide assessments. In some states these assessments have high-consequences for both schools and students that fail to make sufficient academic progress as indicated by their test scores.

Adequate Yearly Progress

Accountability is an essential component of the NCLB Act. Under NCLB, each state is required to develop a plan to address adequate yearly progress (AYP) that will be used to assess each school district and school within the state (US Department of Education, 2003). The goal for every school system across the United States is to have all children, including minority students and students with disabilities, proficient in reading and math by 2014 (US Department of Education, 2003). Each state is able to create its own AYP plans for reaching this goal. Accountability is kept by utilizing statewide assessments. Assessment results and state progress objectives must be broken down by student subgroups based on poverty, race and ethnicity, disability, and limited English proficiency. The AYP results must be made available to parents. AYP results allow parents to know how well their child's school is performing and make informed decisions regarding the school.

AYP Rewards and Sanctions

Schools receive many rewards or sanctions based on the results of their adequate yearly progress (AYP). Schools that fail to make AYP face demanding corrective actions (Abrams & Madaus, 2003). If a school fails to make AYP for two consecutive years it will be identified as needing improvement and the school must develop improvement plans. The school district must offer parents the choice of school and provide transportation to that school for students in the failing school. A school failing to make AYP for three consecutive years must continue to offer school choice and use Title I funds to provide low achieving students with supplemental learning services (The White House, 2002). If a school continues failing AYP, it will be subjected to

state take over and restructuring. Schools that achieve AYP are eligible for State Academic Achievement Awards and "No Child Left Behind" bonuses (The White House, 2002).

Students and teachers may also be rewarded for achieving high test scores (Jones, Jones, & Hargrove, 2003). Teachers often receive bonuses and students receive tickets to sporting events and other benefits. However, if test scores are low teachers will not receive bonuses and low achieving students may not be promoted to the next grade and in some school districts may not be eligible for high school graduation.

The Number Game

Schools find many ways to falsify their test scores and AYP. Many low performing students are being suspended and expelled from school or reclassified as exempt from testing before test days (Amrein & Berliner, 2003b). Children may be encouraged to transfer to another school if the school is in danger of failing to make AYP. There have also been reports of low performing students being retained a grade before pivotal testing years to "ensure they are properly prepared to take the high stakes test" (Amrein & Berliner, 2003b, p. 32). Some low achieving students have even been encouraged to dropout so that schools can achieve high test scores (Bushweller, 2004).

One of the most significant problems revolves around ways in which dropout statistics are calculated. Dropout rates have been disguised in many different ways. Recipients of the General Education Development (GED) are often counted as graduates. Some schools have gone so far as to count students who promise to take the GED test later as a graduating student (Chaplin, 2002; Thornburgh, 2006). Many schools and school districts calculate graduations rates based on the number of 12^{th} grade students. However, research suggests that in order to get

accurate data, graduation rates should be based on 9th grade enrollment and the number students who graduate four years later (Haney, Madaus, Abrams, Wheelock, Mio & Gruia., 2004; Thornburgh, 2006). Ninth grade retention is the biggest risk factor for dropping out of high school. Research has shown there is *bulge* of students in the ninth grade compared to the number of students in 8th and 10th grade (Haney et al., 2004).

A recent study conducted by Swanson (2008) used the Cumulative Promotion Index (CPI) method to calculate graduation rates for America's fifty largest cities. This method multiplies grade-specific promotion ratios together to estimate the likelihood that a 9th grade student will complete high school on time, with a regular diploma. According to the study, "The CPI method represents graduating from high school as a process rather than a single event." (Swanson, 2008, p.7). Using the CPI method, researcher found that only 52% of students in the principal school districts of the fifty largest cities graduate from high school (Swanson, 2008). Further investigation into the metropolitan areas of the fifty largest cities indicated that 58% of students served by the urban school districts, within the largest metropolitan areas, graduate from high school. (Swanson, 2008). Although the CPI method is a very accurate method of calculating graduation rates, it is a very detailed and time consuming method. Calculating graduation rates based on number of students that enter the 9th grade and subsequently graduate four years later may be a feasible compromise that presents an accurate, although grim, picture of the graduation rates in the United States.

High Stakes Testing

Since President George W. Bush signed the NCLB Act in 2002, high stakes testing has become increasingly popular. Although some schools have been using high stakes tests for

many years, it was not mandated until states were required to implement statewide testing as a form of accountability to No Child Left Behind. Prior to NCLB, only sixteen states required mandatory high school exit exams. As of the year 2007, twenty-six states are now requiring high school exit exams and those numbers are likely to continue to rise (Center on Education Policy, 2007).

United States graduation rates have been in decline since 1984 and have plunged even further since many states began implementing high stakes testing (Shriberg & Shriberg, 2006). A number of studies have found that graduation exams increase dropout rates. Jacobs (2001) used data from the National Educational Longitudinal Survey (NELS) to examine the effects of graduation tests on dropout rates. He found that graduation tests increased dropout rates among the lowest achieving students. These students are 25% more likely to drop out of high school than their peers. Similarly, Reardon & Galindo (2002) found that students subjected to a high stakes eighth-grade promotion exam increased the odds of dropping out prior to tenth grade. Even after controlling for a number of school-level covariates, high stakes testing remained a strong predictor of dropping out.

Other studies suggests that high stakes testing leads to increased dropout rates, decreased graduation rates, and higher GED rates (Amrein & Berliner, 2003a). Dropout rates are 4-6% higher in schools with graduation exams. Amrein & Berliner (2002b, 2003a) compared the dropout rates in states with high stakes exams to the national average. They found that 88% of states with graduation exams had higher dropout rates than states without the exams and of those states, 62 % had an increase in the dropout rate in comparison with the rest of the nation after the implementation of the high stakes test.

NCLB was designed to narrow the achievement gap between minority and non-minority students (The White House, 2002). However, it appears that the minority students are being left behind as a result of high stakes testing. Research shows that those who fail high stakes exams are more likely to be minorities and students of low socioeconomic status (Marchant & Paulson, 2005). Performance on high-stakes tests have been found to be directly related to socioeconomic status (SES); lower SES equates to lower test scores (Swanson, 2003). Most states that have high school exit exams have areas with larger concentrations of low-income, minority students (Viadero, 2005). All ethnic minorities, for the exception of Asian American, score significantly lower than white Americans on standardized tests. Altshuler & Schmautz (2006) suggest that white, middle class values, norms, testing language and norms of self perception negatively affect the test success of Hispanic students. They believe that test design and procedures discriminate against Hispanic students (Altshuler & Schmautz, 2006). These sentiments may be generalized across all minority students. In 1998, 1 in 4 African American and Latino 9th grade students in Texas were retained. After these retentions, thousands of them dropped out of school (Amrein & Berliner, 2003a).

Students who fail high stakes exams are often retained until they are able to pass the exam. Moreover, students are being retained prior to pivotal testing years to ensure they pass the high stakes exam (Amrein & Berliner, 2003a). The retention is thought to better prepare the student and provide motivation for achievement. However, numerous studies have concluded that retention significantly increases the likelihood that students will drop out of school (Clarke, Haney, & Madaus, 2000; Goldschmidt & Wang, 1999; Roderick, 1995). Retention does not motivate students to learn more or perform better; it motivates them to drop out of school (Amrein & Berliner, 2003a).

Student Motivation and Achievement

High-stakes testing was designed to increase academic achievement and motivation for academic success, however, it has been shown to decrease motivation and increase the number of students who drop out of school (Amrein & Berliner, 2003a; Reardon & Galindo, 2002). Students become less intrinsically motivated when rewards and sanctions are attached to tests. Retaining students and not allowing them to graduate is not a solution to improve academic success. Extensive research has demonstrated that retention does not have a positive impact on self-esteem or overall school adjustment. On the contrary, retention is associated with significant increases in behavior problems. Retention, particularly ninth grade retention, is also one of the most powerful predictors of high school dropouts (National Association of School Psychologist, 2003).

Student learning does not appear to improve due to the implementation of high-stakes testing. Although initial achievement gains may occur during the year in which a student is retained, achievement gains begin to decline and students perform no better than or worse than same grade students or comparable students that were promoted (National Association of School Psychologist, 2003). Amrein & Berliner (2003a) found that when rewards and sanctions are attached to performance on a test, students become less intrinsically motivated. High stakes exams decrease motivation and increase the number of students who drop out of school. Amrein & Berliner (2002a) examined the results of student Scholastic Aptitude Test (SAT) scores after the implementation of high stakes testing. They found that the SAT score improved in 15 cases and declined in 16 cases. Marchant & Paulson (2005) found similar results when examining SAT scores. Amrein & Berliner (2002a) also examined American College Test (ACT) scores, Advanced Placement (AP) scores, and the National Assessment of Educational Progress

(NAEP). In all but one analysis, student learning did not increase; it either remained the same or decreased after the implementation of high stakes testing. These results indicate that high stakes testing policies have resulted in no measurable improvements in student achievement.

At Risk Students

Recent research indicates that at-risk students for high school dropout can be identified as early as the third grade. It is imperative for schools to identify third grade students who meet any of the following four criteria: a student that is one grade-level below his/her peers, been retained one year, resides in a lower socioeconomic household, or attends schools with greater than 50 percent of students receiving free/reduced lunch. Students who meet two or more of the four factors results in the student having no chance of completing high school (Canady, 2008). Sixth grade at-risk students may be identified by the following criteria: failing math, failing language arts, less than 80 percent attendance rate, or two or more behavioral referrals per month. Finally, ninth grade students may be identified as at-risk if: they are over age, they finish the ninth grade without enough credits to graduate with his/her cohort, less than 80 percent attendance rate, fails Algebra I, fails English 9, or is disconnected from school. Any sixth or ninth grade students that meet two or more of these criteria have no chance of completing high school (Canady, 2008).

Other criticisms of high stakes testing include narrowing curriculums. Non-testing areas such as art, music, social studies, and physical education are being pushed out of the curriculum in order to make more time for testable subjects (Amrein & Berliner, 2003a). Students are not receiving a well rounded education. They are only learning the content and items from state

administered tests, leaving little room for higher level academic achievement and learning (Amrein & Berliner, 2002; Marchant & Paulson, 2005).

In addition, reports of increasing numbers of children suffering from sleep disorders and other stress related maladies have been connected to high-stakes testing. Some parents believe that high stakes testing puts "undo pressure on children" (Dounay, 2000, p. 10). Students experience a range of different emotions when faced with testing that can cause headaches, stomach aches, irritability, and aggression (Amrein & Berliner, 2003a; Jones et al., 2003).

In summary, high-stakes testing appears to be related to higher student dropouts and lower graduation rates. Minority students, students of low socioeconomic status and students with disabilities are the most in danger of failing a high-stakes exam and subsequently dropping out of school. These lower achieving students are often asked to be absent on the day of testing and in some cases encourage to dropout of school. Finally, high-stakes tests have not been found to improve the overall academic achievement and learning of students. Teachers can easily improve test scores by teaching the test, however, this limits the amount of supplemental information a student learns in testable subjects and severely reduces learning time in non-testable subjects such as social studies, art and music.

West Virginia

The state of West Virginia created the West Virginia Educational Standards Test (WESTEST) to comply with the statewide assessment requirements of No Child Left Behind. The WESTEST is a criterion referenced test designed for West Virginia students in grades 3-8 and 10 (West Virginia Department of Education, 2007). The test was implemented and became a requirement for all students in grades 3-8 and 10 during the 2003- 2004 school year. Although

this may be considered a high stakes test for educators, due to West Virginia AYP requirements, it does not carry serious consequences for students. Students in West Virginia are not required to pass the exam in order to be promoted to the next grade or graduate from high school.

During the 2006 school year the West Virginia Department of Education (2007), reported that 69% of tenth grade students received proficient scores on the mathematics section of the WESTEST, 72% were proficient in reading, and 87% were proficient in science. The lowest achieving students were African American, Hispanic, and American Indian (West Virginia Department of Education, 2007).

Ohio

Ohio students are required to take the Ohio Achievement Test (OAT) in grades 3-8 and the Ohio Graduation Test (OGT) in grade 10. There are five sections on the Ohio Graduation Test and students must pass all five sections in order to graduate from high school (Ohio Department of Education, 2007). The Ohio Graduation Test replaced the Ninth Grade Proficiency Test as the high school graduation requirement beginning in 2007. During its first year of implementation as a graduation requirement 65.3% of students taking all five tests achieved a proficient level or above (Ohio Department of Education, 2007). The results indicate that 34.7% of Ohio students did not achieve the proficient level needed to graduate. American Indian, African American, and Hispanic students achieved the lowest percentages of proficiency while White and Asian students achieved the highest percentages of proficiency (Ohio Department of Education, 2007). Results dating back to 2004 show similar trends.

Statement of the Problem

A review of the literature revealed that there may be higher dropout rates and lower graduation rate among students who are required to take high stakes tests in order to be promoted or graduate from high school. The perceived causes for these higher dropout rates may be linked to research on risk factors for high school dropouts that indicate students who are retained, especially in the ninth grade, have a very little chance of completing high school. Students that are at the highest risk for failure are minority students, students with disabilities, and students in low socioeconomic areas. Therefore, it may be ascertained that low socioeconomic areas, such as Appalachian regions, may have an increased risk for higher dropouts and lower graduation rates. The question remains does high-stakes testing increase the number of students who dropout, thereby decreasing graduation rates?

Purpose of the Study

Previous research indicates that implementation of high stakes testing leads to lower graduation rates (Amrein & Berliner, 2002b, Jacobs 2001, Reardon and Galindo, 2002). Jacobs (2001) and Reardon & Galindo (2002) used NELS data to analyze the effects of high stakes testing on a single cohort group. As Jacobs (2001) discussed, a more preferable approach would be to compare multiple cohort groups before and after implementation of high stakes exams. In addition, Amrein & Berliner (2002b) compared the graduation rates of states before and after implementation of high stakes exams to the national average. A stronger approach would be to compare changes in graduation rates of high stakes testing states to those of non-high stakes testing states.

A recent study attempted to analyze the effects of high stakes testing by comparing the dropout rates in West Virginia for three years before and three years after implementation of the WESTEST (Dodd, 2007). Results indicated no significant differences in the dropout rates after implementation of the exam. Limitations of Dodd's (2007) study included the way in which dropout rates were calculated and the fact that there are no consequences for students who do not pass the WESTEST. Although the WESTEST has high stakes implications for teachers and administrators, students are not required to pass the exam to be promoted to the next grade or to graduate from high school.

The purpose of this study is to examine the relationship between high stakes testing and graduation rates. It is a follow-up and expansion of Dodd's (2007) study conducted in West Virginia. The current study will use ninth grade enrollment data to calculate cohort group graduation rates. It will look at two Appalachian region states, one with a high stakes exam the other without, and attempt to determine whether or not high stakes exams leads to decreased graduation rates.

The research hypotheses for this study are: 1. Graduation rates in West Virginia decrease with implementation of the WESTEST. 2. Graduation rates in the 29 Appalachian counties of Ohio decrease with implementation of the Ohio Graduation Test. 3. The 29 Appalachian counties of Ohio will have lower graduation rates than West Virginia because Ohio utilizes the Ohio Graduation Test as a high stakes exam. The null hypothesis is: There is no difference in graduation rates in West Virginia or the 29 Appalachian counties of Ohio with implementation of high stakes testing.

Significance of the Problem

Statewide testing, as mandated by the No Child Left Behind Act, is used to determine students' academic achievement status. In many states, assessment scores are used to decide whether or not a student is promoted to the next grade or graduates from high school.

Students with disabilities and minority students are subject to failing high-stakes exams at a higher rate than other students (Jacobs, 2001; Jones, et al., 2003). Yet, No Child Left Behind is designed to improve the academic achievement of all students, including disability and minority students. Is No Child Left Behind really leaving no child behind with only 50-60% graduation rates in some areas (Swanson, 2003; Swanson, 2008)? It is very alarming that academic achievement is based on tests that doom certain students to failure and increase the risk of them dropping out of school.

CHAPTER TWO

METHOD

Two Appalachian regions, the twenty-nine Appalachian counties of Ohio and the fifty-five counties of Appalachian West Virginia, were chosen for this study because of their similar demographic information and lower migration rates. According to the Ohio Department of Development (2008) and the United States Census Bureau (2008), the total population for the twenty-nine Appalachian Ohio counties in the year 2000 was 1,455,313. The total population for the state of West Virginia in the year 2000 census was 1,808,344. Clermont, the largest county within the twenty-nine Appalachian counties of Ohio, census population was 177,977, as compared to West Virginia's Kanawha county population of 200,073. Ohio's largest Appalachian area of Union Township had a population of 42,322 in 2000, compared to Charleston, West Virginia's population of 53,422 (United States Census Bureau, 2008).

West Virginia and the 29 Appalachian counties of Ohio fall within the Central sub-region of Appalachia. Central Appalachia boasts the lowest migration rates of all Appalachian regions. Obermiller & Howe (2004) reported that Central Appalachia's gross migration rate was 12.9 percent in 2000 as compared to the Non-Appalachian United States rate of 18.3 percent. Migration in and out of the Central Appalachian sub-region appears to remains fairly stable. Only a 0.6 percent gain in population was reported from 1995-2000 (Obermiller & Howe, 2004). These trends allow for a more reliable and valid representation of graduation rates calculated in this study, as students in the West Virginia and the 29 Appalachian counties of Ohio are less likely to transfer out of the Central Appalachian sub-region.

Research Design

A time-series research design was used to examine the effects of implementation of high stakes testing on graduation rates. According to Campbell and Stanley (1963, p. 37) "the essence of a time series design is the presences of a periodic measurement process on some group or individual and an introduction of an experimental change into this time series of measurements, the results of which are indicated by a discontinuity in the measurements recorded in the time series." To determine if there is a relationship between high stakes testing (independent variable) and high school graduation rates (dependant variable), the graduation rates in Ohio's twenty-nine Appalachian counties were calculated for the years before and after implementation of the Ohio Graduation Test (OGT) and compared to a non-high stakes test Appalachian state, West Virginia, using multiple Chi Squares.

Participants

The participants for the study consisted of students who attended high school in Ohio's twenty-nine Appalachian counties between the school years 1996- 2007 and students who attended high school in West Virginia between the school years 1996-2007. The focus was on those students who graduated from 2000-2007. Graduation rates were calculated by dividing the number of students that graduated by the ninth grade enrollment numbers four years prior to their graduation.

Procedures

Ninth grade enrollment and the number of students who graduate in Ohio and West Virginia are reported by the Ohio Department of Education and the West Virginia Department of

Education. The data collected included all students who were enrolled in the ninth grade from 1996-2003 in Ohio's twenty-nine Appalachian counties and all fifty-five Appalachian counties of West Virginia, and the number of students who graduated from Ohio's Appalachian counties and West Virginia from 2000-2007. Students' names, grades, and test scores were not used for this study. The graduation rates for West Virginia were calculated by dividing the total number of graduates by the total number of students enrolled in the 9^{th} grade four years prior to their graduation year (see Table 1). In addition, graduation rates were calculated for the 29 Ohio Appalachian counties (see Table 2). Graduation rates were then compared for the four years prior to implementation of high stakes testing and then for the four years during and after implementation.

CHAPTER THREE

RESULTS

The graduation data for both Ohio and West Virginia was analyzed using the Chi Square statistic. The Chi Square allows data to be analyzed in the form of proportions, such as the 9^{th} grade enrollment to graduation proportion used to calculate the graduation percentage rates used in this study. First, a Chi Square was run on the data from each school year to determine if there was a significant difference in the graduation rate before and after implementation of the WESTEST. Results of the Chi Square test showed no significant differences in the West Virginia graduation rates on a year by year basis at $p < 0.05$ (see Table 3). In addition, no significant differences were seen when making a multiple year analysis of the data, indicating that the statewide assessment in West Virginia had no effect on graduation rates (see Table 3).

As a comparison group, the graduation rates in the 29 Appalachian counties of Ohio were examined after implementation of the Ohio Graduation Test. The Chi Square test was used to analyze the Ohio Appalachian counties graduation rate. Result showed no significant differences between the Ohio Appalachian counties graduation rates on a year by year basis or multiple year basis at $p < 0.05$ (see Table 4), thus, implementation of high stakes testing did not appear to have an effect on graduation rates in Ohio.

Finally, for a state by state comparison of graduation rates the 29 Appalachian counties of Ohio was compared to West Virginia. A Chi Square test was used to analyze the graduation rates of the 29 Ohio Appalachian counties and West Virginia. Results of the year by year and multiple year analysis indicated no significant differences in the graduation rates of the twenty-nine Appalachian counties of Ohio and the state of West Virginia at $p < 0.05$ (see Table 5).

Given that no significant differences were found during any portion of the data analysis, the null hypothesis is upheld. High stakes testing does not appear to have a significant effect on graduation rates in Appalachian West Virginia or the 29 Appalachian counties of Ohio.

CHAPTER 4

DISCUSSION

The purpose of this study was to examine the effects of high stakes testing on high school graduation rates. A longitudinal study was conducted in two states, West Virginia and the Appalachian counties of Ohio, to determine if implementation of high stakes testing decreases graduation rates. Research has suggested that high stakes testing leads to decreased graduation rates and higher dropout rates (Amrein & Berliner, 2003a; Jacobs, 2001, Shriberg & Shriberg, 2006). Ohio implemented the Ohio Graduation Test in 2004. Although West Virginia implemented the WESTEST in 2004, students are not required to pass the exam in order to graduate. Given the research that suggests high stakes testing decreases graduation rates, decreases in graduation rates should be seen after implementation of high stakes testing. Although slight decreases were noted in the year after implementation in both Ohio and West Virginia, the decreases were not significantly different from the other years. When comparing Ohio's graduation rates to West Virginia's graduation rates, the non high stakes testing state of West Virginia had lower graduation rates in all years except for the years 2000 and 2001. Again, none of the years analyzed approached levels of significance. These findings suggest that, contrary to previous research, high stakes testing does not significantly impact graduation rates.

One possible reason for the findings in this study may be related to Ohio's retest policy regarding student who do not pass the high stakes exam. Ohio does not retain students for failing the Ohio Graduation Test. Students have the opportunity to take the exam three times, once in the tenth, eleventh, and twelve grades before being denied graduation. School districts that only give one opportunity to pass statewide exams before retention may have more significant decreases in graduation rates, than those found in this study. Future research may wish to

examine the graduation rates in states that retain high school students for failing the graduation exam. Given that ninth grade retention is a very strong predictor of high school dropouts, it would be of particular interest to examine the graduation rates in a state that retains ninth grade students for failing a high stakes exam.

A second factor that may have influenced the results of this study encompasses the fact that Ohio has had a form of a graduation test since 1994. Prior to the 2004 implementation of the Ohio Graduation Test, students were required to pass the Ninth Grade Proficiencies Tests. Therefore, this study does not reflect a true introduction of experimental change in the Ohio group. In order to obtain more accurate information about the effects of high stakes testing implementation on graduation rates, research would need to look at years before there was any form of high stakes testing and comparing it to the years after implementation of a high stakes exam.

A third factor that may have influenced the results of this study involves the weaknesses of the statistic used to analyze the data. The Chi Square is not as powerful as parametric statistics. The Chi Square lacks power and sometimes underestimates small differences. However, in analyzing graduation rates, a nonparametric statistic such as the Chi Square is the only valid statistic that can be used. This non-parametric nature of statistical analysis is a measure of central tendency and therefore does not account for variance within each state. To account for possible outlying counties within each state, graduation rates were plotted for all 55 counties in West Virginia and the 29 Appalachian counties in Ohio. Although, there were a couple of years in which a county in West Virginia and Ohio had a higher or lower graduation rate than the average, none were significantly different from the other counties within the state.

An advantage of utilizing a time-series design, allows researchers to plot data and analyze the directional implication of the line (Campbell & Stanley, 1963). Beginning in 2002, West Virginia's graduation rates were on the rise. Then in 2005, the year after WESTTEST implementation, the graduation rate dropped and has yet to fully recover (See Figure 1). A similar pattern was seen in Ohio. However, Ohio's Appalachian counties graduation rates have recovered, as the year 2006 rate of 81.5 percent was the highest seen from 2000-2007 (See Figure 2). Ohio's first class required to pass the Ohio Graduation Test was the class of 2007. It will be of interest to follow the graduation rates over the next several years to see if the new version of the high stakes test effects graduation rates in Ohio.

Another important finding from this study, involves the formula used for calculating graduation rates. School districts across the United States utilize different criteria to calculate their graduation rates. The current study looked at cohort groups of students entering ninth grade and those graduating four years later. Rates calculated in this study were often ten or more percentage points lower than those reported on the Ohio and West Virginia State Department of Education websites (Ohio Department of Education, 2009; West Virginia Department of Education, 2009). Although the formula used in this study does not account for variables (i.e. transfers, deaths, etc.) other than students who do not graduate, it may provide a more realistic view of graduation rates.

West Virginia's ninth grade enrollment significantly increased from 24,041 students in 1997 to 32,507 in 1998. Although the ninth grade enrollment increase in 1998, the number of students graduating four years later remained consistent with the number of students graduating across all other comparison years. The increased enrollment caused the 2002 cohort graduation rate to drop to 52 percent. Though this may appear to be a dramatic decrease in the graduation

rate, the 2002 graduation rate was not found to be significantly different from any other year. There are numerous possible hypotheses for the increased ninth grade enrollment, including those related to the restructuring of a junior high 7-9th grade model to a middle school 6-8th grade model and changes in private school law. Although there are several possible explanations, the nature and causes of the increased enrollment is beyond the scope of the present study. Future research may wish to examine ninth grade enrollment trends in West Virginia and the subsequent effects on graduation rates.

The results obtained in this study are significantly different from much of the previous research. One reason for this difference may be related to teacher anxiety regarding the WESTEST in West Virginia. Although the WESTEST does not have high stakes implications for students, teachers' perception of the test may influence how students view the test. Teachers begin planning for the WESTEST very early in the year. Much of the academic instruction throughout the year is focused on preparing students for the test. Even before third grade, teachers talk to their students about the importance of WESTEST. After the testing week, schools have numerous activities and parties to celebrate the end of testing. Students who score high are often recognized and rewarded by their school. Students may perceive the WESTEST to be a high stakes test because teachers present the test to student as if it were high stakes. It may be of interest to examine teacher anxiety produced by high stakes testing and the effects that anxiety has on students.

In conclusion, previous research suggests that high stakes testing decreases graduation rates. Results of this study do not support those findings. Although, Ohio recently implemented the Ohio Graduation Test, it was not a new implementation of high stakes testing. This may explain why no significant differences were found in the graduation rates before and after

implementation of the Ohio Graduation Test. Further research examining the graduation rates in a state that did not have a high stakes test before the statewide testing mandates of No Child Left Behind, but now requires students to pass an exam to graduate needs to be conducted in order to obtain a more accurate view of the effects of high stakes testing implementation on graduation rates. It may also be beneficial to determine whether or not differences exists between states that allow students to take the graduation exam several times before denying graduation and those states which retain students the year the exam is failed. High stakes testing was designed as an accountability measure for United States public schools. If new research finding prove to be consistent with the results found in the present study, high stakes testing may prove to be an ineffective measure of accountability, as it appears to produce no positive or negative effects on student education.

References

Abrams, L.M. & Madaus, G. F. (2003). The lessons of high-stakes testing. [Electronic Version]. *Educational Leadership, 6, 5*, 31-35.

Altshuler, S. J., & Schmautz, T. (2006) No hispanic student left behind: The consequences of high stakes testing. [Electronic Version]. *Children and School, 28, 5*, 1-14.

Amrein, A. L., & Berliner, D. C. (2002a). High stakes testing, uncertainty, and student learning. *Educational Policy Analysis Achieves, 10, 18*. Retrieved October 8, 2007 from http://epaa.asu.edu/epaa/v10n18/.

Amrein, A.L., & Berliner, D.C. (2002b). *An analysis of some unintended and negative consequences of high-stakes testing.* Arizona State University: Educational Policy Research Unit.

Amrein, A. L. & Berliner, D. C. (2003a). The effects of high-stakes testing on student motivation and learning. [Electronic Version]. *Educational Leadership, 60, 5*, 32-38.

Amrein, A. L. & Berliner, D. C. (2003b). The testing divided: new research on the intended and unintended impact of high-stakes testing. *Peer Review, 5, 2*, 31-32.

Bushweller, K. (2004). Study of exit exams notes failure issues. [Electronic Version]. *Education Week, 24, 1*, 18.

Campbell, D.T. & Stanley, J.C. (1963). *Experimental and Quasi-Experimental Designs for Reaserach.* Reprinted from Handbook of Research on Teaching. Boston, MA: Houghton Mifflin Company.

Canady, R.L. (2008, January). Early identification of at-risk students, 3rd, 6th, and 9th grade. Presented at the Ohio Appalachian Center for Higher Education 2008 Middle School Initiatives workshop, Huntington, West Virginia.

Center On Education Policy. (2007). *State high school exit exams: working to raise test scores*. September 2007. Retrieved October 5, 2007 from http://www.cep-dc.org/index.cfm?fuseaction=Page.viewPage&pageId=493&parentID=481.

Clarke, M., Haney, W., & Madaus, G. (2000). High stakes testing and high school completion. [Electronic Version]. *The National Board on Educational Testing and Public Policy*. Lynch School of Education. Boston College.

Chaplin, D. (2002). Tassels on the cheap. *Education Next, 2, 3*, 24-29.

Dodd, A.S. (2007). *High stakes testing effects graduation rates.* (Masters thesis, Marshall University 2007). Available at http://www.marshall.edu/etd/ETD-List.html.<http://www.marshall.edu/etd/ETD-List.html>

Dounay, J. (2000). High-stakes testing is high-stress, too. [Electronic Version]. *Education Digest, 65, 9*, 9-13.

Goldsmith, P. & Wang, J. (1999). When can school affect dropout behavior: a longitudinal multilevel analysis. [Electronic Version]. *American Education Research Journal, 26, 4*, 715-738.

Haney, W., Madaus, G., Abrams, L., Wheelock, A., Miao, J. & Gruia I. (2004). The education pipeline in the united states, 1970-2000. *The National Board On Education Testing and Public Policy*. Lynch School of Education. Boston College. Retrieved on February 11, 2008 from http://www.bc.edu/research/nbetpp/statements/nbr3.pdf.

Jacobs, B. A. (2001). Getting tough? The impact of high school graduation exams. [Electronic Version]. *Educational Evaluation and Policy analysis, 23, 2,* 99-121.

Jones, M. G., Jones, B. D., & Hargrove, T. (2003). *The Unintended Consequences of High Stakes Testing*. Lanham, Maryland: Roman & Littlefield Publishers, Inc.

Marchant, G. J. & Paulson, S. E. (2005). The relationship of high school graduation exams to graduation rates and SAT scores. Educational Policy Analysis Archives, 13, 6. Retrieved September 29, 2007 from http://eppa.asu.edu/eppa/v13n16/.

National Association of School Psychologist. (2003). *Position Statement on Student Grade Retention and Social Promotions.* Retrieved February 11, 2008 from http://www.nasponline.org/about_nasp/pospaper_graderetent.aspx.

Obermiller, P.J, & Howe, S.R. (2004). Moving mountains: Appalachian migration patterns, 1995-2000. *Journal of Appalachian Studies, 10, 3,* 359-371.

Ohio Department of Development. (2008). *Ohio County Profiles: Appalachia*. Retrieved February 5, 2009 from http://www.odod.state.oh.us/research/files/s0/appalachia.pdf.

Ohio Department of Education. (2007). *State of Ohio proficiency tests*. Retrieved October 8, 2007 from www.ode.state.oh.us.

Ohio Department of Education. (2009). *Power User Reports*. Retrieved March 23, 2009 from http://ilrc.ode.state.oh.us/Power_users.asp.

Reardon, S. F. & Galindo, C. (2002). Do high stakes testing affect students' decision to drop out of school? Evidence from NELS. *Department of Education Policy Studies*.

Roderick, M. (1995). Grade retention and school dropout: Policy debate and research. [Electronic Version]. *Phi Delta Kappa Research Bulletin, 15*, 1-6.

Shriberg, D. & Shriberg, A. B. (2006). High-stakes testing and dropout rates. [Electronic Version]. *Dissent, 53, 4*, 76-80.

Swanson, C.B. (2003*). Keeping count and loosing count: Calculating graduation rates for all students under NCLB accountability*. [Electronic Version]. Education Week, 23, 20, 1.

Swanson, C. B. (2008). Cities in crisis: a special analytic report on high school graduation. *Editorial Projects in Education Research.* Retrieved April 1, 2008 from http://www.americaspromise.org/uploadedFiles/AmericasPromiseAlliance/Dropout_Crisis/SWANSONCitiesInCrisis040108.pdf

Thornburgh. N. (2006) Dropout nation. *Time, 167, 16,* 30-40.

The White House (2002). *Fact Sheet: No Child Left Behind*. Retrieved September 27, 2007 from http://www.whitehouse.gov/news/releases/2002/01/20020108.html.

United States Census Bureau. (2008). *United States Census 2000*. Retrieved February 5, 2009 from http://quickfacts.census.gov/qfd/states/54000lk.html.

United States Department of Education (2003). *Key Policy Letters Signed by the Education Secretary or Deputy Secretary*. Retrieved September 27, 2007 from http://www.ed.gov/policy/elsec/guid/secletter/020724.html.

United States Department of Education (2004). *Executive Summary*. Retrieved September 27, 2007 from http://www.ed.gov/nclb/overview/intro/execsumm.html.

United States General Accounting Office. (1993). Student testing: current extend and expenditures, with cost estimates for national examination. Report to Congressional Requesters. Retrieved October 1, 2007 from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/13/91/30.pdf

Viadero, D. (2005). Researchers connect lower achievement, high school exit tests. [Electronic Version]. *Education Week, 24, 21*, 10.

West Virginia Department of Education. (2007). WESTEST. Retrieved October 10, 2007 from http://westest.k12.wv.us/.

West Virginia Department of Education. (2009). *West Virginia Education Information System*. Retrieved March 23, 2009 from http://wveis.k12.wv.us/nclb/pub/enroll/repstatgrar.cfm?xrep=1&sy=07.

Table 1

West Virginia Graduation Rates from 2000-2007

Year	9th Grade Enrollment	Year	#Graduates	Graduation Rate
1996	25,124	2000	19,440	0.774
1997	24,041	2001	18,447	0.767
1998	32,507	2002	17,147	0.527
1999	27,266	2003	17,272	0.633
2000	23,041	2004	17,018	0.739
2001	23,211	2005	16,336	0.704
2002	22,896	2006	16,408	0.717
2003	23,578	2007	17,024	0.722

Table 2

Ohio-Appalachian Counties Graduation Rate from 2000-2007

Year	9th Grade Enrollment	Year	#Graduates	Graduation Rate
1996	22,049	2000	16,076	0.729
1997	21,731	2001	15,452	0.711
1998	20,929	2002	15,709	0.751
1999	20,179	2003	16,310	0.808
2000	20,574	2004	15,953	0.775
2001	19,881	2005	15,605	0.785
2002	19,413	2006	15,823	0.815
2003	19,384	2007	15,404	0.795

Table 3

Chi Square Value and Significance Level for West Virginia 2000-2007

Year	df=1 Significance Level
2000-2001	No
2000-2002	No
2000-2003	No
2000-2004	No
2000-2005	No
2000-2006	No
2000-2007	No
2001-2000	No
2001-2002	No
2001-2003	No
2001-2004	No
2001-2005	No
2001-2006	No
2001-2007	No
2002-2000	No
2002-2001	No
2002-2003	No
2002-2004	No
2002-2005	No
2002-2006	No
2002-2007	No
2003-2000	No
2003-2001	No
2003-2002	No
2003-2004	No
2003-2005	No
2003-2006	No
2003-2007	No
2004-2000	No
2004-2001	No
2004-2002	No
2004-2003	No
2004-2005	No
2004-2006	No
2004-2007	No
2005-2000	No
2005-2001	No
2005-2002	No
2005-2003	No
2005-2004	No
2005-2006	No
2005-2007	No
2006-2000	No
2006-2001	No
2006-2002	No
2006-2003	No
2006-2004	No
2006-2005	No
2006-2007	No

Table 4

Chi Square Value and Significance Level for 29 Appalachian Ohio Counties 2000-2007

Year	df=1 Significance Level
2000-2001	No
2000-2002	No
2000-2003	No
2000-2004	No
2000-2005	No
2000-2006	No
2000-2007	No
2001-2000	No
2001-2002	No
2001-2003	No
2001-2004	No
2001-2005	No
2001-2006	No
2001-2007	No
2002-2000	No
2002-2001	No
2002-2003	No
2002-2004	No
2002-2005	No
2002-2006	No
2002-2007	No
2003-2000	No
2003-2001	No
2003-2002	No
2003-2004	No
2003-2005	No
2003-2006	No
2003-2007	No
2004-2000	No
2004-2001	No
2004-2002	No
2004-2003	No
2004-2005	No
2004-2006	No
2004-2007	No
2005-2000	No
2005-2001	No
2005-2002	No
2005-2003	No
2005-2004	No
2005-2006	No
2005-2007	No
2006-2000	No
2006-2001	No
2006-2002	No
2006-2003	No
2006-2004	No
2006-2005	No
2006-2007	No

Table 5

Chi Square Value and Significance Level for 29 Appalachian Ohio Counties and West Virginia 2000-2007

Year	df=1 Significance Level
2000-2000	No
2000-2001	No
2000-2002	No
2000-2003	No
2000-2004	No
2000-2005	No
2000-2006	No
2000-2007	No
2001-2000	No
2001-2001	No
2001-2002	No
2001-2003	No
2001-2004	No
2001-2005	No
2001-2006	No
2001-2007	No
2002-2000	No
2002-2001	No
2002-2002	No
2002-2003	No
2002-2004	No
2002-2005	No
2002-2006	No
2002-2007	No
2003-2000	No
2003-2001	No
2003-2002	No
2003-2003	No
2003-2004	No
2003-2005	No
2003-2006	No
2003-2007	No
2004-2000	No
2004-2001	No
2004-2002	No
2004-2003	No
2004-2004	No
2004-2005	No
2004-2006	No
2004-2007	No
2005-2000	No
2005-2001	No
2005-2002	No
2005-2003	No
2005-2004	No
2005-2005	No
2005-2006	No
2005-2007	No

Table 5 continued

Chi Square Value and Significance Level for 29 Appalachian Ohio Counties and West Virginia 2000-2007

Year	df=1 Significance Level
2006-2000	No
2006-2001	No
2006-2002	No
2006-2003	No
2006-2004	No
2006-2005	No
2006-2006	No
2006-2007	No
2007-2000	No
2007-2001	No
2007-2002	No
2007-2003	No
2007-2004	No
2007-2005	No
2007-2006	No
2007-2007	No

Figure 1

Percentage of West Virginia Graduates 2000-2007

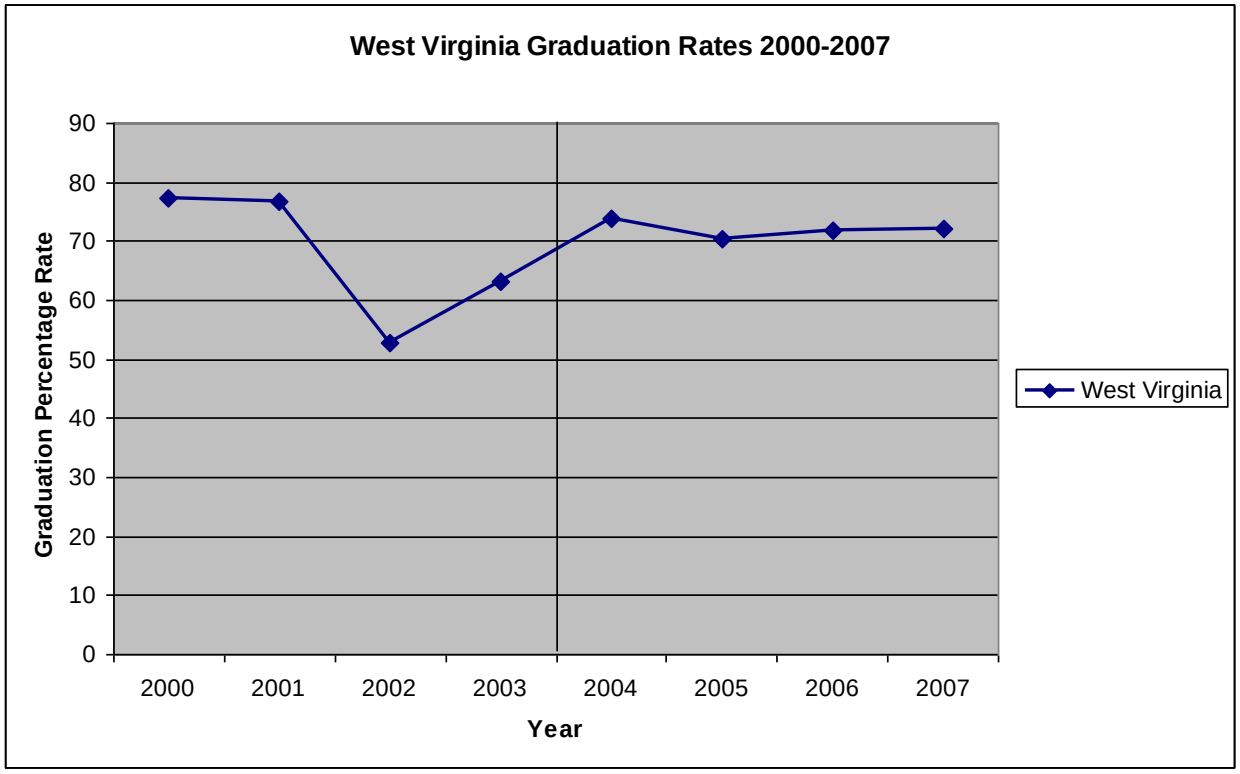

Figure 2

Percentage of Ohio Appalachian Counties Graduates 2000-2007

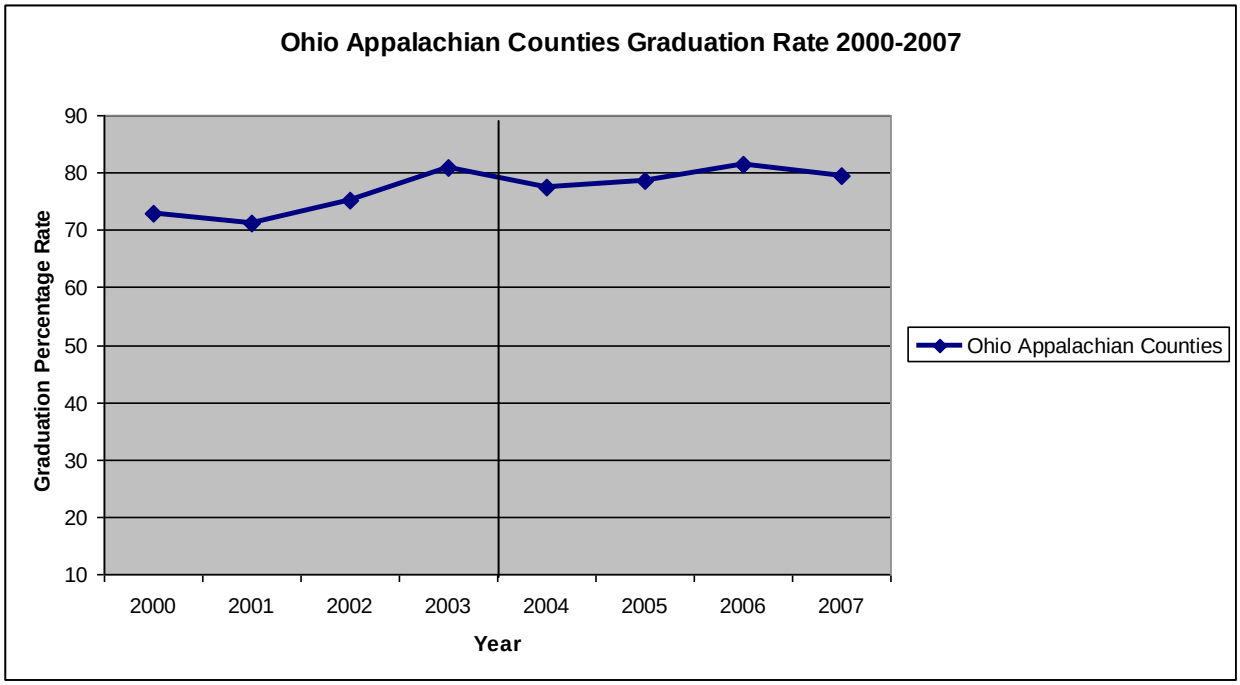

Figure 3

Percentage of Ohio Appalachian Counties and West Virginia Graduates 2000-2007 Comparison