

No School Left Uncorrupted:
How Cheating, High-Stakes Testing, and Declining Budgets Affected Student
Achievement in Philadelphia

by
James Sadler

Dr. Marc Meredith, Advisor

A thesis submitted in partial fulfillment
Of the requirement for the
Degree of Bachelor of Arts
In Political Science with Distinction

University of Pennsylvania
Philadelphia, PA

April 1, 2013

Abstract

The No Child Left Behind education act, signed by President Bush in 2002, imposed high-stakes testing standards on all schools in the nation. A decade later, amidst a cheating scandal and a budget crisis, the School District of Philadelphia experienced dramatic standardized test score declines after nine years of increases. This study aims to place these declines in the context of national, state, and local education policy and provide statistical evidence for the cause of the declines. School climate among schools flagged for cheating and budget decreases experienced by all Philadelphia schools significantly contributed to the declines. Nevertheless, the major finding faults the elimination of cheating, enforced by the use of increased testing security in 2012, for the lower test scores. The analysis supports existing theory that high stakes testing encourages administrative cheating and hinders the educational achievement of students.

Table of Contents

I. Introduction	Pg. 4
II. National Education Policy and Cheating Theory	Pg. 10
III. The State of Education in Pennsylvania and Philadelphia	Pg. 20
IV. The Philadelphia Schools Cheating Scandal	Pg. 30
V. Data	Pg. 38
VI. Hypotheses	Pg. 42
VII. Research Design	Pg. 45
VIII. Results	Pg. 54
IX. Analysis	Pg. 66
X. Discussion	Pg. 72
XI. Conclusion	Pg. 82
Acknowledgements	Pg. 85
Appendix: Detailed Table of Flagged Schools in Philadelphia	Pg. 86
Bibliography	Pg. 90

I. Introduction

In 2010, Roosevelt Middle School, a school with an 85% poverty concentration located in the Germantown neighborhood of Philadelphia, Pennsylvania, gathered the attention of several education advocates for impressive improvements in academic achievement. In 2008, only 25.6% of Roosevelt's students were proficient in math while 29.7% were proficient in reading. In just two years, these figures jumped to 75% proficient in math and 78.2% proficient in reading. The school's teachers and staff were applauded for beating the odds and overcoming obstacles such as poverty, violence, and teenage pregnancy. Roosevelt's principal noted, "if you think miracles can't happen, look at us."¹ Then-Pennsylvania Governor Ed Rendell commented how Roosevelt proved politician's doubts about Philadelphia wrong. Even the US Secretary of Education, Arne Duncan, joined in the celebrations by personally congratulating Philadelphia's superintendant and Roosevelt's principal.

Roosevelt's success was part of a citywide trend of progress across several key indicators measured in 2011. Standardized test scores in Philadelphia increased for the ninth straight year in both reading and math. The graduation rate for public school students was at an all time high. More schools met federal No Child Left Behind standards than ever before. Although then-Superintendent Arlene Ackerman admitted that the District "[had] a long way to go," the public viewed

¹ Graham, Kristen; Purcell, Dylan. "City school's fast rising test scores questioned". *The Philadelphia Inquirer*. May 1, 2011.

Philadelphia as a District where—slowly, but surely—progress had occurred.² In an editorial, *The Philadelphia Inquirer* claimed that progress made in Philadelphia schools was “encouraging and worth applauding...they show that efforts to improve academic achievement in the city’s public schools are slowly making a difference.”³ The Pew Charitable Trusts noted “Philadelphia’s education numbers tell a story of much that has been accomplished...”⁴ The Council of Great City Schools named Ackerman the best urban superintendent in the country.⁵ Much like the city itself, the School District successfully combated its reputation of being broken and unfixable.

Unfortunately, the cautious optimism surrounding the District—one that took years to build—evaporated quickly over the course of the next year. First, the District spent the entire year trying to close a \$715 million budget shortfall that arose due to both state budget cuts and District mismanagement.⁶ The District also dealt with a transition of leadership as four different people led the District between August 2011 and October 2012. Under extreme pressure from her employers, Ackerman left the District in August. Her Deputy, Leroy Nunery, led as Acting Superintendent until January 2012, when Thomas Knudsen was then temporarily brought in as Chief Recovery Officer (with the power of both a superintendent and chief financial officer) in an attempt to curb the District’s rapidly deteriorating

² Graham, Kristen. “A test-score milestone for Phila. schools”. *The Philadelphia Inquirer*. June 17, 2010.

³ “Editorial: Applaud, but lightly”. *The Philadelphia Inquirer*. June 19, 2010.

⁴ “Philadelphia 2011: The State of the City”. *The Pew Charitable Trusts*. Apr. 9, 2011.

⁵ “Philadelphia Superintendent Wins Nation’s Top Prize”. *Council of the Great City Schools*. Oct. 21, 2010.

⁶ Mezzacappa, Dale. “Budget ‘ominous,’ further cuts required; District to hire recovery officer”. *The Philadelphia Public School Notebook*. Jan 19, 2012.

budget situation. Ackerman's permanent successor, Dr. William Hite, was appointed in the summer of 2012 but did not officially start until October. Thus, for over a year, Philadelphia schools were left with a limbo of leadership without clear direction.

Yet even with these already negative occurrences, the most damaging effect on Philadelphia's education outlook arrived with a cheating scandal that first broke out in the summer of 2011. Across the entire state, several schools and Districts contained suspicious amounts of corrective erasure markings on standardized tests. The vast majority of these suspicions came from Philadelphia, where more than 25% of all District run schools were publicly flagged for multiple erasure irregularities between 2009-2011. Roosevelt Middle School—whose academic rise had seemed too good to be true—was flagged for every grade level in every subject across all three years that were investigated. Doubts on the credibility of the previous nine years of test score increases immediately surfaced as both the state and the District continued to investigate cheating accusations—an effort that is still ongoing at the time of this study.

In an effort to combat this alleged cheating, the School District increased security measures for all schools under their jurisdiction in 2012, with particular scrutiny on the flagged schools. Teachers could not administer tests to their own students. Boxes containing test booklets remained locked and unopened until hours before test time. Central Office staff aided the process by monitoring school hallways and classrooms during tests.

When all the 2012 tests were taken and scores calculated, Philadelphia's proficiency rate plummeted across every grade and student subgroup. Overall math

proficiency rates dropped 8.7 percentage points (from 57.8% in 2011 to 49.1% in 2012). Similarly, reading proficiency rates dropped 7.1 percentage points (from 51.6% to 44.5%). The 2012 test score declines put Philadelphia back on 2008 test-proficiency levels. Meanwhile, Roosevelt's scores plummeted by around 50% in both subjects, eliminating the impressive gains made the previous two years. The Philadelphia miracle appeared completely forced and fabricated.

While the elimination of cheating seemed to contribute heavily on the academic achievement declines in 2012, there are other plausible alternative explanations. During the nine years of test score increases, Philadelphia schools also experience nine years of funding increases. However, school budgets were drastically cut at an unprecedented rate during the 2011-2012 school year—teachers were fired, extra curricular activities suspended, and resources eliminated. Perhaps the test score decreases were simply the result of Harrisburg failing to provide schools with enough resources to succeed.

On the other hand, the cheating scandal had a significant effect on Philadelphia schools. Morale was at an all time low, and all schools—flagged or not—were forced to defend their credibility to the public. Teachers and administrators complained that the intense security measures placed by the State and District created a climate that affected children on test day.

Of course, the most common analysis blamed schools themselves for the declines. Because of the increased security measures, schools that got away with cheating in previous years were forced to comply with proper testing protocol in

2012. In this way, the 2012 test scores showed Philadelphia's true achievement for the first time.

This study aims to understand the context of education policy leading up to the 2012 declines and statistically determine what caused them. Did cheating in Philadelphia occur due to something particular about national, state, or local education policy, or did a minority of administrators misrepresent test data simply because they had the opportunity to do so? How much did cheating actually affect student achievement? Did flagged schools and grade levels experience sharper declines than their non-flagged counterparts? In many of the flagged schools, certain grade levels were not marked for cheating but still received the same level of public scrutiny and security because other grade levels in the schools *were* flagged. Did these non-flagged grade levels in flagged schools experience similar declines to their non-flagged counterparts? Were patterns different between the math and reading tests?

The qualitative analysis portion in the study finds that the No Child Left Behind act—signed into law by President Bush in 2002—is guilty of creating a hostile environment in schools as it placed too much importance and influence on high-stakes standardized tests. The proficiency rates of students potentially determined the job status of teachers and administrators and inevitably led to cheating.

On the statistical side, this study finds that the 2012 budget decreases experienced in Philadelphia schools significantly contributed to the declines in test scores, even when controlling for cheating. It also suggests that environmental

factors within cheating schools may have contributed to the overall score declines, since test scores in non-flagged grades dropped significantly more if they were part of a flagged school. Nevertheless, the overall conclusion finds that even when taking the above findings in consideration, the elimination of massive cheating contributed significantly to the rapid decline in Philadelphia test scores.

The paper is organized as follows: The first half of the study aims to present a story of the context of the District on the day students took their test in March 2012. Part II will describe the existing theory on why public school administrators cheat and how cheating is connected to national education policy initiatives. Part III will provide an overview of the state of education in both Pennsylvania and the School District of Philadelphia. Part IV will specifically focus on the details of the cheating scandal in Philadelphia. The later half will shift the paper from qualitative description to quantitative analysis. Parts V – VII will introduce the data used, hypotheses tested, and equations manipulated. Part VIII will outline the results from statistical tests while Part IX will analyze the implications of what the data reveals. Finally, Part X will discuss who or what should be blamed for the test score declines and what the data reveals about the actions taken by local, state, and national education policy makers.

II. National Education Policy and Cheating Theory

The Birth of No Child Left Behind

Education was firmly on the map during the 2000 presidential election, and the issues that divided the candidates continue to split the education debate today. Like many presidential candidates before and after him, President Bush fought hard to paint himself as an education pioneer during the 2000 presidential campaign. In one particular commercial, his campaign claimed that “as president, George W. Bush will challenge the status quo with a crusade to improve education. He’ll fight for reforms hailed as the most fundamental in a generation.”⁷ Indeed, as Governor of Texas, Bush oversaw reforms that *The Economist* claimed “improved [educational] standards...and narrowed the achievement gap between white and non-white students.”⁸

Bush was an advocate for placing more accountability on schools and school districts with high stakes standardized testing. The argument stated that chronically failing schools did not deserve public funding and parents had the right to send their children to higher performing schools. Parents would make an educated decision on where to enroll their children based on standardized test scores made available to the public. High stakes testing thus gave school districts the incentive to perform well while giving parents reliable school choices. If schools did not make the grade, this quasi market-based system would lead them to close while higher performing schools would grow.

⁷ Rosenbaum, David. “The 2000 Campaign: The Ad Campaign; Bush on Education”. *The New York Times*. March, 18, 2000.

⁸ “How bad is Texas, really?” *The Economist*. Oct. 19th, 2000.

On the other side of the political spectrum, Democratic candidate Al Gore's proposal for reforming schools focused much more on development rather than punishment. Under his plan, low performing schools would receive a large influx of funding accompanied by school turnaround experts who would work with the school for at least two years. If the school continued to demonstrate low performance, only then would staff turnover occur.⁹ Bush's accountability-based reform initiatives won out over Gore's investment-based approach and Bush was quick to implement his ideas in his new office.

One of Bush's first presidential acts was to propose what would become the No Child Left Behind Act (NCLB), "arguably the most far-reaching education policy initiative in the United States over the last four decades".¹⁰ Passed with bipartisan support, NCLB created national standards for growth in proficiency in math and reading. The goal of the act was for every school in the country to meet 100% proficiency in both subjects by 2014. Each individual state was allowed to create their own tests, standards of proficiency, and targets for yearly growth. If a school or district met the target proficiency rate for a given year, the institution would meet Adequate Yearly Progress (AYP).¹¹ Schools that make AYP are granted greater autonomy with their budgets and policies. However, schools that miss AYP over multiple consecutive years are forced to overhaul their school, which could possibly mean conversion to a charter school or complete staff turnover.

⁹ Goldstein, Andrew. "Bush and Gore: Who's The Education President?". *TIME Magazine*. Nov. 6, 2000.

¹⁰ Dee, Thomas S., and Brian A. Jacob. "The Impact of No Child Left Behind on Students, Teachers, and Schools." *Brookings Papers on Economic Activity* 2010.2 (2010): 149.

¹¹ Institutions must also meet a 95% test participation target as well as varying attendance and graduation targets, depending on if the school has a graduating class or not.

True to its name, the No Child Left Behind Act was the first of its kind to not only stress accountability, but also emphasize improvement for multiple student subgroups: white students, minority students, English language learners, special education students, and economically disadvantaged students. An institution only meets AYP if *all* student groups of at least 40 students reach performance targets.¹² Undoubtedly, the No Child Left Behind Act created a high-stake testing environment in all fifty states.

Education advocates and researchers have criticized NCLB for a multitude of reasons, including shifting away resources from other subjects not tested (social studies, science, art, etc.) and encouraging instruction that focuses on test-taking training instead of valuable instruction.¹³ Furthermore, some critics are skeptical of state-reported claims of large proficiency rate increases.

One way of analyzing the effectiveness of NCLB measures is to look at results from low-stake testing assessments like the National Assessment of Educational Progress (NAEP). The NAEP tests students every two years on a variety of subjects by using the same test nationwide, making a national proficiency rate and state-by-state comparisons possible. Grades 4th and 8th are tested on the math and reading NAEPs with three possible achievement levels: Basic, Proficient, and Advanced. Basic students “denotes partial mastery” of the subject while Proficient students

¹² There are multiple ways for a subgroup to make a specific target, including meeting the target outright, meeting the target over a two-year average, decreasing the amount of non-proficient students by 10%, or coming close enough to meeting the target through any of these measures that a margin for error surpasses the gap of success.

¹³ Dee & Jacob, “The Impact of No Child Left Behind...”. Pg. 150.

“have demonstrated competency over challenging subject matter”.¹⁴ While there is no definition, students may also perform below the Basic standard.

While proficiency rate on the NAEP have increased since NCLB’s implementation in 2002, they are nowhere near the 100% proficiency goal for 2014 (see Table 1). Furthermore, proficiency rates increased before NCLB just as much as after.

Table 1: Nation-wide Proficiency Rate on 4th/8th Grade Math/Reading NAEP¹⁵

	1992	2003	2007	2011
4th G. Reading	29%	31%	33%	34%
8th G. Reading	29%	32%	31%	34%
4th G. Math	18%	32%	39%	40%
8th G. Math	21%	29%	32%	35%

These trends are not inline with indications from several state-level standardized tests that show high proficiency rates and increased gains. Pennsylvania tests their students using the Pennsylvania System of School Assessment (PSSA) tests.

Proficiency standards for the PSSA are far behind standards set by the NAEP. For example, over the past three NAEP testing cycles, proficiency rates for 8th grade math tests are significantly higher on the PSSA than on the NAEP (see Table 2). The difference between the two tests has recently increased as the 2011 PSSA proficiency rate was over twice as high as the NAEP equivalent. These lower

¹⁴ Definitions retrieved from <http://nces.ed.gov/nationsreportcard/about/nathowreport.asp>

¹⁵ U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, 2002, 2003, 2005, 2007, 2009 and 2011 Reading and Mathematics Assessments.

proficiency standards allow most Pennsylvania districts and schools to label themselves as successful.

Table 2: Pennsylvania 8th Grade Math Proficiency Rate Based on PSSA and NAEP¹⁶

	2007	2009	2011
NAEP	38%	40%	38%
PSSA	68%	71%	77%
Difference	30%	31%	39%

Pennsylvania is certainly not alone in demonstrating a higher proficiency rate than the nationally administered NAEP. In fact, all 50 states had proficiency standards that were lower than NAEP's standards in both reading and math, with the exception of Massachusetts in math.¹⁷ Those critical of NCLB state that this gap demonstrates that proficiency standards for state tests mask true learning and that any real gains made on these high stakes test is a product of teachers focusing on how to take the test itself rather than on teaching the broad subject matter tested. Even with these higher proficiency rates, some districts and schools resorted to other means to boost their test proficiency rates.

Cheating Scandals: Part of the New American School System

¹⁶ Ibid.

PSSA scores taken from the Pennsylvania Department of Education website.

¹⁷ U.S. Department of Education: Mapping State Proficiency Standards Onto the NAEP Scales: Variation and Change in State Standards for Reading and Mathematics, 2005-2009. National Center for Education Statistics. 2011. pg. 14

A key question in this study will analyze the claim that “high-stakes testing may lead school personnel to intentionally manipulate student test scores.”¹⁸ While Gore’s model of education reform would have supported struggling schools and teachers, Bush’s reform plan placed extremely high pressure on institutions to raise test scores. As previously stated, schools that don’t meet AYP for multiple years in a row under No Child Left Behind must undergo swift transformation that often includes hiring new teachers and administrative staff. High pressure combined with lax security provides ample opportunity for testing improprieties.

Cheating scandals that have occurred across the country since NCLB’s inception have supported this concern about manipulation. For example, in Atlanta, a state investigation into cheating in 2011 found that:

“cheating was caused by a number of factors but primarily by the pressure to meet targets in the data-driven environment...a culture of fear, intimidation and retaliation existed in APS [Atlanta Public Schools], which created a conspiracy of silence and deniability with respect to standardized test misconduct.”¹⁹

Over forty schools were found guilty of cheating. Thirty-eight principals were implicated while 82 teachers and staff members confessed to misrepresenting student test score data. In the Western half of the country, 10% of 3,000 Arizona teachers surveyed claimed that they knew of other teachers participating in cheating, including replacing incorrect answers with correct ones and refusing to test low-performing students.²⁰ The DC Public School system, previously run by

¹⁸ Dee & Jacob.

¹⁹ Resmovits, Joy. “Atlanta Public Schools Shaken By Cheating Report”. *The Huffington Post*. Sep. 4, 2011.

²⁰ Amrein-Beardsley, A.; Berliner, D.C.; Rideau, S.. “Cheating in the first, second, and third degree: Educators’ responses to high-stakes testing”. *Education Policy Analysis Archives*. Vol 18, No. 14. 2010.

Michelle Rhee, a strong high-stakes accountability advocate, is also currently under investigation for testing irregularities while Rhee was at the helm.²¹ Other scandals have occurred in Texas, Ohio, California, Chicago, and more—all after No Child Left Behind came into effect.²² Cheating has occurred at such a high frequency, that Diane Ravitch—former US Assistant Secretary of Education and leading education advocate—admitted that she has “never seen so many cheating scandals as there have been in the last few years.”²³ Ravitch believes the rise in cheating is directly related to the fact that NCLB’s 100% proficiency goal looms closer and closer.

Undoubtedly these Districts and schools in question had an incentive to perform well on tests. Consistent negative results on standardized tests would often lead to the elimination of teachers, administrators, and at times entire schools. However, the question of whether high-stakes testing is completely to blame is debatable.

Bad Apples or Bad Policy?

The question of why cheating occurs has been heavily discussed before and after NCLB: is cheating the result of a wayward minority trying to game the system, or do certain policies encourage and inevitably lead to corruption? In the context of education policy, did No Child Left Behind’s high-stakes testing atmosphere do more

²¹ Gillum, Jack and Bello, Marisol. “When standardized test scores soared in D.C., were the gains?” *USA Today*. Mar. 30, 2011.

²² Samuels, Christina. “Cheating Scandals Intensify Focus on Test Pressures”. *Education Week*. Aug. 4, 2011.

²³ Resmovits, Joy. “Atlanta Public Schools...”

harm than help to students by causing unintended consequences like cheating by educators?

Some place the blame of cheating on certain individuals who were given the opportunity to do so. Professor Cizek of University of North Carolina at Chapel Hill assisted in the cheating investigation in Atlanta public schools and believes high stakes testing has zero blame in all cheating scandals. Cizek believes that “it’s a huge mistake...to think that the incentives to cheat in school settings are fundamentally any different from those in any other context.”²⁴ In this argument, pointing the blame on testing policies would be equivalent to sympathizing with baseball players who use performance-enhancing drugs to gain an advantage in a pressure-filled occupation. For Cizek, cheating will always occur “as long as important educational decisions are informed in part by test results,” but only because the same people who “cut corners” as sportsmen, politicians, and journalists exist in all facets of society—including education.²⁵

But how much is the act of cutting-corners a factor of one’s own behavior and how much is due to the demands of a situation? Cambell’s Law (named after social scientist Donald Campbell) is often used to describe the relationship between external pressure and corruption:

The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.²⁶

²⁴ Cizek, Gregory J. “Cheating on Tests and Other Dumb Ideas”. *Education Week*. July 25, 2011.

²⁵ Ibid.

²⁶ Anderman, Eric, & Murdock, Tamera. *Psychology of Academic Cheating*. Elsevier Academic Press. 2007. Pg. 292.

According to Eric Anderman and Tamera Murdock, Campbell's Law is applicable to both the student and teacher side of the classroom when referring to the magnitude of the weight attached to a single measure of aptitude and competency. Students who cheat "do so as a function of the situations in which they find themselves rather than because of some general flaw in character."²⁷ Similarly, teachers are pushed to perform unprofessional actions when policies exert unfair expectations. The authors find that NCLB "has sown seeds of corruption in a once morally trustworthy, highly respected profession."²⁸ Furthermore, the declining trust of teachers and administrators as a result of cheating scandals "is too great a price to pay for an accountability system that appears not to work."²⁹ In their opinion, standardized tests should take a Gore-like approach and be used to foster development among teachers and staff instead of hostility and degradation.

Nichols and Berliner take this argument a step further by forgiving teachers for certain types of cheating:

In some cases, cheating is difficult to condemn. For example, when a teacher "helps" a struggling student with one or two challenging test items, we may view it as a small and forgivable infraction when compared to the potential motivational and psychological costs of that student failing another test.³⁰

The authors believe that cheating by teachers and administrators is at worst morally vague when not directly connected to a teacher's personal motivations and gains.

According to Nichols and Berliner, teachers may be looking out for the best interests of their students when they decide to participate in cheating. No Child Left Behind

²⁷ Ibid. pg. 289

²⁸ Ibid. pg. 307

²⁹ Ibid. pg. 308

³⁰ Nichols, Sharon L. and Berliner, David C.. *Collateral Damage: How High-Stakes Testing Corrupts America's Schools*. Harvard Education Press. 2008. Pg. 34.

puts teachers in a moral dilemma where they “may come to the conclusion that cheating on a test, like stealing unaffordable medicine [in order to save a dying loved one], can be justified.”³¹ In this case, it is the national education policy brought down from Washington politicians that must answer for the disturbing trends of cheating in school districts across the country.

As this study delves deeper into the details of the vast amounts of cheating that occurred in Philadelphia, one must consider the implications of education policy in the classroom and how unintended consequences can nullify good intentions.

³¹ Ibid. pg. 35

III. The State of Education in Pennsylvania and Philadelphia

NCLB In Pennsylvania: The PSSA

Before No Child Left Behind was signed into law in 2002, Pennsylvania already had a statewide assessment measure known as the Pennsylvania System of School Assessment, or the PSSA. Instituted in 1992, the PSSA test changed to its current form in 2000. Pennsylvania continues to use the PSSAs in accordance with NCLB. Students in grades 3-8 and 11 take the test every year in math and reading. Some of those grades also take tests in science and writing; however, these tests are not counted for AYP purposes. Students' scaled scores are translated into four categories: Below Basic ("inadequate academic performance"), Basic ("marginal"), Proficient ("satisfactory") and Advanced ("superior"). The Department of Education states that proficiency "indicates a solid understanding and adequate display of the skills included in the Pennsylvania Academic Content Standards".³² Students scoring either Proficient or Advanced are counted towards AYP targets.

Across the entire state of Pennsylvania in 2012, 74% of students tested scored Proficient or above in math while 71% did so in reading—both small declines from the previous year (these scores are roughly 25 percentage points higher than Philadelphia's rates). Translated into NCLB and Adequate Yearly progress terms, the majority of the 500 school districts and 153 charter schools in

³² "Technical Report for the 2012 Pennsylvania System of School Assessment". *Data Recognition Corporation*. 2012. pg. 274. Available on the Pennsylvania Department of Education website.

the state made AYP in 2012 (58%) while 91% have made it within the past two years.³³

Philadelphia's experience with No Child Left Behind has lagged behind compared to the rest of the state. Only three school districts have failed to meet AYP standards every year since No Child Left Behind's inception: Harrisburg City, Chester Upland, and the School District of Philadelphia. Only seven schools (.2% of all schools in the state) have failed to make AYP every single year—everyone of them in Philadelphia. The label of failure given to these institutions by No Child Left Behind has contributed to Philadelphia's reputation of a crumbling school system beyond repair. Additionally, as performance targets have risen in anticipation of the 2014 goal, state standards are becoming increasingly unattainable for Philadelphia schools.

As the 100% proficiency target in 2014 approaches, Philadelphia—and indeed the rest of the state—is suffering due to the way the state defined its yearly targets. Because NCLB allowed individual states to create their own proficiency targets, Pennsylvania—along with several other states—took a “backloaded” approach to reaching the 100% proficiency target in 2014 (See Table 3). For example, targets in math rose 21 percentage points between 2002 and 2007, which means proficiency rates need to rise 55 percentage points between 2007 and 2014. Some have speculated that states took this approach with the belief that NCLB would be replaced by another policy as the seemingly impractical 100% goal of

³³ Data taken from the Pennsylvania Department of Education's AYP website: <http://paayp.emetric.net/StateReport#pie>

2014 neared.³⁴ While this notion has come to fruition for the majority of states in the form of NCLB waivers given by the federal Department of Education, Pennsylvania is one of six states that has yet to apply for a waiver (though has plans to do so at the time of this writing).³⁵

Table 3: Pennsylvania AYP Proficiency Targets

Year	Math Proficiency Target	Reading Prof. Target
2002	35%	45%
2003	35%	45%
2004	35%	45%
2005	45%	54%
2006	45%	54%
2007	45%	54%
2008	56%	63%
2009	56%	63%
2010	56%	63%
2011	67%	72%
2012	78%	81%
2013	89%	91%
2014	100%	100%

³⁴ Chudowsky, Naomi; Chudowsky, Victor. "Many States Have Take a 'Backloaded' Approach to No Child Left Behind Goal of All Students Scoring 'Proficient'". *Center on Education Policy*. May 19, 2008. Pg. 5.

³⁵ Niederberger, Mary. "No Child Left Behind waiver to be sought by Pennsylvania". *Pittsburgh Post-Gazette*. Nov. 28, 2012.

The increased targets have caused more pressure for schools deemed failing and needing restructuring. Although test scores across the District rose in 2011, the number of schools making AYP dropped from 158 schools to 110 schools due to the increase of the target proficiency rate.³⁶ Schools that fail to make AYP for multiple years in a row are forced to restructure—a move that often leads to teacher and administration turnover.

In Philadelphia, the two main routes of reform are turning failing schools into Promise Academies or Renaissance Charter Schools; each of these systems lead to teacher and administration turn-over. Promise Academies are District-run schools with extra resources that allow for a longer school day, among other things. The Principal turnover rate at Promise Academies is high. Whoever is named principal must turnover more than 50% of teaching staff. The teachers that are not chosen to stay are either forced-transferred to another school or let go by the District, depending on District availability and the teacher's seniority. Under the current agreement between the School District and the Philadelphia Federation of Teachers (Philadelphia's main teaching union), teachers are laid off based on a system of seniority: younger teachers are laid off first.

Schools that are restructured as Renaissance Charter Schools effectively turn into a neighborhood charter school: independently managed yet publicly funded schools that must take students from the neighborhood boundaries. In this situation, the hiring of teachers and administrators is completely out of District control and is decided by whichever Charter Management Organization (CMO) runs

³⁶ Only 33 schools in the District made AYP in 2012. Herold, Benjamin. "110 schools meet AYP targets" *The Philadelphia Public School Notebook*. Aug. 18, 2011.

the school. Teachers at the school are employed by the CMO (not the District), and are thus not unionized. Because of these factors, teacher and principal turnover rates are high. While these turnover models both aim to bring resources and new perspectives to under-performing schools, teachers and principals have an incentive to increase their students test scores before their jobs can possibly be placed in jeopardy.

The School District of Philadelphia: A History of Fiscal Mismanagement

The School District's poor reputation across the state of Pennsylvania derives largely from recent tensions over fiscal mismanagement by District administrators. The School District has long had a rocky relationship with the state government, and this was exacerbated with the creation of the School Reform Commission (SRC) in 2001. The SRC was created in a bargain between the District and the state government, where the District ceded the state greater control over its governance in exchange for more state funding.³⁷ What resulted was the SRC, a five-member school board—three appointed by the governor and two appointed by the mayor. While the SRC would function as a normal school board, it allowed the State to have greater power over the financial dealings of the School District as well as the appointment of the Superintendent. However, the power of appointing superintendents did not eliminate conflict between District leaders and the SRC.

In its first summer of its existence, the SRC unanimously appointed Paul Vallas—a superstar with a reputation for fixing the public school system in

³⁷ "Who runs the School District of Philadelphia?". *Education Voters of Pennsylvania*. 2013.

Chicago—as Superintendent.³⁸ Vallas would go on to spearhead important reform movements including expanding the charter school system, creating smaller high schools, and a \$1.7 billion capital expansion plan.³⁹ However, he suddenly lost support from the SRC during the 2006-2007 school year when a \$73.3 million budget deficit surprisingly arose from seemingly nowhere.⁴⁰ This SRC, apparently as surprised as everyone else, claimed that they were “betrayed” and “disappointed” in Vallas in allowing such a deficit to occur.⁴¹ Vallas claimed that the SRC were just as much to blame as he was as they both had the same amount of access to budgetary information.⁴² The in-fighting eventually led to Vallas leaving Philadelphia to successfully lead public schools in New Orleans in the summer of 2007.

Arlene Ackerman replaced Vallas as Superintendent a year later and created her own reform plan called Imagine 2014. The initiative called for turning around lower performing schools by either converting them to charters or providing extra funding with direct District supervision.⁴³ This part of the plan, called the “Renaissance Schools Initiative”, along with many other parts of the plan, were funded by federal stimulus money that was intended to be used on one-time expenditures instead of recurring programs.⁴⁴ When stimulus funds ran dry and the State drastically decreased Philadelphia’s funding in 2011, Ackerman and her

³⁸ Fifield, Adam. “Power: Reformer R.I.P.”. *Philadelphia Magazine*. September 2007.

³⁹ Graham, Kristen. “Vallas on Ackerman”. *The Philadelphia Inquirer: Philly School Files*. Sep. 6, 2011.

⁴⁰ Basham, Nick. “The Five-Year Legacy of Paul Vallas in Philadelphia”. *The Philadelphia Inquirer*. June 2, 2007.

⁴¹ Snyder, Susan. “Vallas facing sharp criticism”. *The Philadelphia Inquirer*. Nov. 11, 2006.

⁴² Fifield, “Power: Reformer R.I.P”

⁴³ Socolar, Paul. “Imagine 2014 gets a green light from the SRC”. *The Philadelphia Public School Notebook*. Summer 2009. Vol. 16. No. 4.

⁴⁴ Owens, Erika. “Ackerman’s tenure in Philadelphia: 2008-11 key events”. *The Philadelphia Public School Notebook*. October 2011. Vol. 19. No. 2.

administration were facing a \$715 million budget shortfall—almost ten times the amount that had forced Vallas to leave.⁴⁵ Ackerman was an extremely controversial and polarizing figure whose tenure ended bitterly in August 2011 after feuds between herself, the SRC, and Philadelphia mayor Michael Nutter proved too heated to reconcile.⁴⁶ The SRC bought out the rest of her contract by paying her \$905,000 to leave her post.⁴⁷

Philadelphia: Tragically Underfunded

As previously mentioned, the School District faced an unprecedented budget crisis during the 2011-2012 school year. While the School District has certainly suffered from internal financial mismanagement, the financial neglect from the State of Pennsylvania has also left the District with fewer resources than were needed to adequately educate their students.

Like in many other states, Pennsylvania funds its schools through a combination of state funding and local property tax revenue. Because property tax revenues are based on the value of the property owned by residents within a district's boundaries, wealthier districts can tax low to spend high while poorer districts (such as Philadelphia) must tax high merely to spend low. Ideally, basic education funding from the State reconciles this gap.

In 2006, the General Assembly of Pennsylvania commissioned a study to determine how much money the state needed to fund their school districts to

⁴⁵ Mezzacappa, Dale. "Budget 'ominous,' ..."

⁴⁶ Kerkstra, Patrick. "Nutter, Ackerman Split". *The Philly Post*. June 7, 2011.

⁴⁷ Owens, Erika "Ackerman's tenure..."

“provide an education that will permit [students] to meet the State’s academic standards and assessments”.⁴⁸ The completed study, which was released a year later, found that the State was underfunding schools by \$4.57 billion if they wanted to meet their achievement goals. Specifically, Philadelphia needed a 42% funding increase from the state to have adequate resources to succeed by the states’ standards.⁴⁹ In 2008, Governor Ed Rendell made a six-year plan to close the vast majority of this gap by creating funding formulas that targeted poorer school districts while dramatically increasing basic education funding.⁵⁰ The 2009 economic recession put a dent in those plans as Rendell was forced to use one-time federal stimulus money to increase education funding.

When Governor Corbett came into office in 2011, the General Assembly (under the advice of the Governor) decreased K-12 basic education funding by almost \$900 million for the 2011-2012 Fiscal Year.⁵¹ The cuts included the loss of federal stimulus money and the elimination of grants and funding formulas that benefited poorer schools.⁵² Poorer districts that depended on the state more for funding lost much more money per-pupil than wealthier districts. For example, while the School District of Philadelphia lost \$533 per-pupil, Lower Merion School

⁴⁸ “Costing Out the Resources Needed to Meet Pennsylvania’s Public Education Goals”. *Augenblick, Palaich and Associates, Inc.* December, 2007. Pg. ii.

⁴⁹ *Ibid.* Pg. 71.

⁵⁰ Venkataramanan, Rajiv; Mezzacappa, Dale. “How state funding has changed since the costing-out study”. *The Philadelphia Public School Notebook*. October 2010. Vol. 18. No. 2.

⁵¹ Worden, Amy. “PA budget highlights: On time, slimmed down”. *The Philadelphia Inquirer: Commonwealth Confidential*. June 30, 2012.

⁵² White, Bill. “Pennsylvania Gov. Tom Corbett bombing public education into Stone Age”. *The Morning Call*. Sept. 26, 2012.

District—just a ten-minute drive away—lost less than \$20 per-pupil.⁵³ Although Philadelphia educated 10% of the states’ students, they suffered 30% of the basic education funding cuts.

The drastic decreases in funding were unprecedented for Philadelphia and contributed substantially to the record \$715 million shortfall. As Figure 1 shows, Philadelphia had not experienced a funding decrease in the previous two decades, never mind one so severe. The District was forced to make drastic cuts to programs and staffs. Around 4000 teachers, nurses, secretaries, reading specialists, and other personnel were let go.⁵⁴ Half of the central office staff was laid off.⁵⁵ Twenty-six schools lost their full-service kitchens.⁵⁶ Seven schools closed in the summer of 2012 while 23 more will close this summer.⁵⁷ While AYP targets continue to rise rapidly towards 100% proficiency, Philadelphia schools have not been in worse financial shape in its long history.

⁵³ Hardy, Dan. “Pa. budget cuts hit poorer districts hardest, funding equity advocate says”. *The Philadelphia Inquirer*. July 11, 2011.

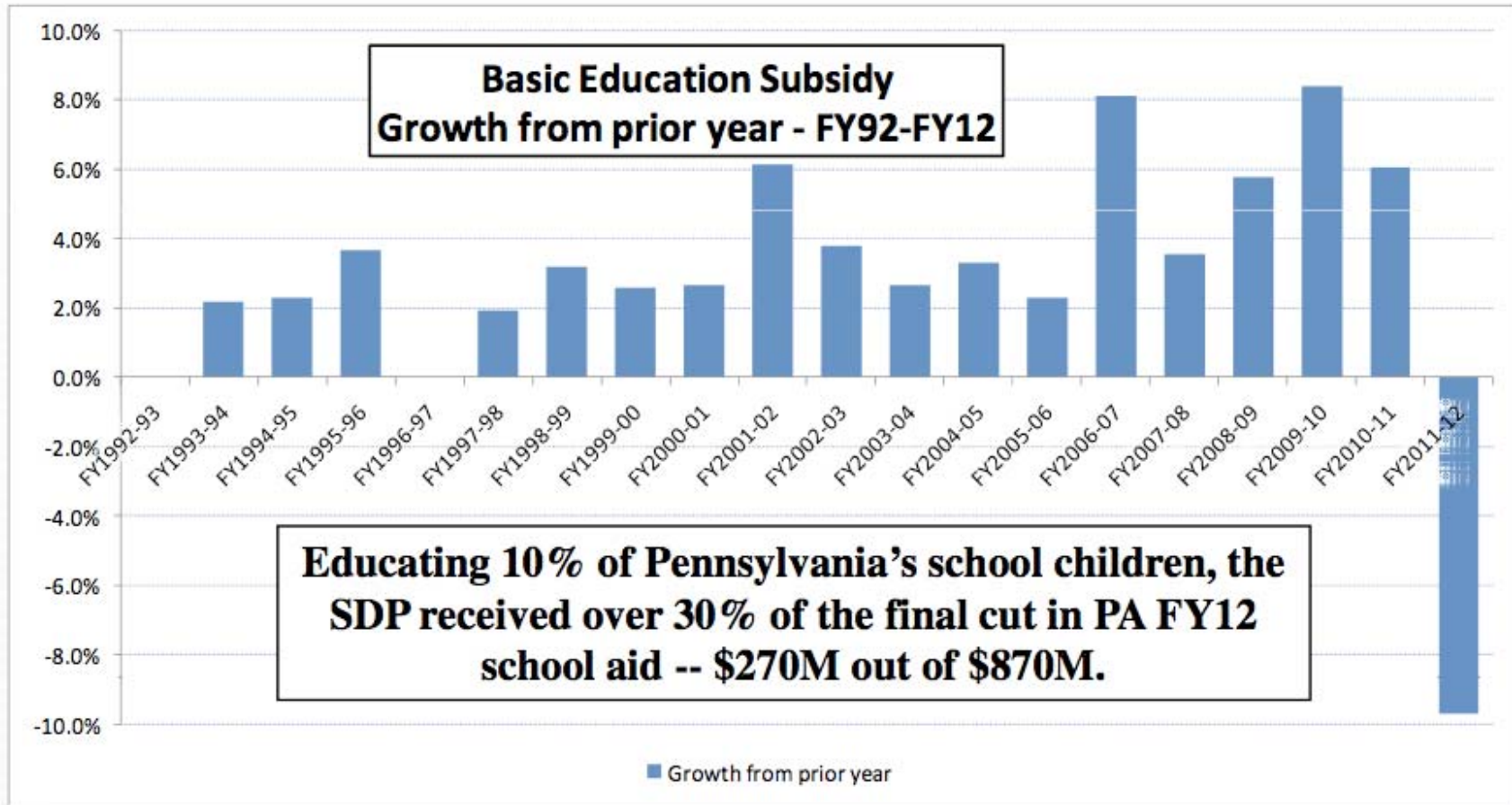
⁵⁴ Mezzacappa, Dale. “Just in time for the holidays: more layoffs”. *The Philadelphia Public School Notebook*. Dec. 22, 2011.

⁵⁵ Scott, Katherine. “3,800 Philadelphia school personnel face layoffs”. *6ABC.com*. June 6, 2011.

⁵⁶ Graham, Kristen. “26 Philadelphia schools losing full-service kitchens”. *The Philadelphia Inquirer*. June 30, 2011.

⁵⁷ Mezzacappa, Dale. “Hite: New plan will result in fewer students moving to lower-performing schools”. *The Philadelphia Public School Notebook*. Feb. 19, 2013.

Figure 1: Pennsylvania Basic Education Subsidy growth for the School District of Philadelphia, 1992-2012⁵⁸



⁵⁸ This chart was presented at the August 3rd, 2011 School Reform Commission meeting by then School District of Philadelphia Chief Financial Officer Michael Masch. Masch, Michael. "SDP Budget Update: Changes Since Initial Adoption on May 31, 2011". *School District of Philadelphia*. August 3, 2011. Slide 13.

IV. The Philadelphia Schools Cheating Scandal

Erasure Results Collect Dust

Though the Philadelphia schools cheating scandal erupted in 2011, evidence for the scandal was available years before. In the summer of 2009, the Pennsylvania Department of Education paid the Data Recognition Corporation (the company that collects and aggregates data on the PSSA) \$108,000 to conduct a forensic analysis on the 2009 PSSA tests.⁵⁹ The final report included four different statistical tests to determine the “integrity of test results”.⁶⁰ The clearest of these four measures is the erasure analysis.

Test scanners are capable of determining if someone (a student or administrator) changed a response to a question by erasing one answer choice and filling in another one. The Data Recognition Corporation aggregated the number of erasures in every reading and math PSSA taken by Pennsylvania public school students in 2009. By using the mean number of wrong-to-right erasures (erasing an incorrect answer and replacing it with a correct one) the analysis flagged schools that had suspiciously high numbers of wrong-to-right erasures in specific grade levels. In other words, if the probability of a school’s wrong-to-right erasures occurring due to chance exceeded 1 in 10,000, the school was flagged.⁶¹ In the end, 225 schools in the state were flagged in 2009, 88 (39%) of which were in the School

⁵⁹ Graham, Kristen. “Pa. education secretary orders reviews of all state exams since 2009, especially Phila.’s”. *The Philadelphia Inquirer*. July 15, 2011.

⁶⁰ “Technical Report for the Pennsylvania System of School Assessment: 2009 Reading and Mathematics Grades 3, 4, 5, 6, 7, 8, and 11”. *Data Recognition Corporation*. July 2009. Page 3.

⁶¹ *Ibid.* Pg. 7.

District of Philadelphia.⁶² Some Philadelphia schools greatly exceeded the minimum requirement for a flagged status. For example, 55% of Strawberry Mansion High School 11th graders had suspiciously high numbers of wrong-to-right erasures on their 2009 Math PSSAs. The likely hood of this pattern occurring by chance was 1 in 10,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000—roughly the equivalent of winning the Mega Millions lottery jackpot five times in a row.⁶³

For reasons that are still largely unclear, the report did not cause even the simplest of follow-ups. The State Secretary of Education at the time, Gerald Zahorchak, later claimed that “the report never reached my desk...I don’t have any idea why it languished.”⁶⁴ When the report eventually became public in 2011 (under a different administration), the School District of Philadelphia asserted that they were never informed of the report’s data.⁶⁵ The spokesperson for the Secretary of State in 2011 summed up the situation by stating, “[the report] was basically put on the shelf and ignored”.⁶⁶ No forensic analyses were completed after the 2010 PSSA tests as the state discontinued them due to budgetary concerns.

Roosevelt and FitzSimons: The Calm Before the Storm

⁶² Herold, Benjamin. “One-third of District schools had pattern of suspicious erasures”. *The Philadelphia Public School Notebook*. July 11, 2011.

⁶³ Lecher, Colin. “What Are Your Odds Of Winning The Lottery? [Infographic]”. *Popular Science*. Nov. 29, 2012.

⁶⁴ Herold, Benjamin. “Pa. wants 90 schools investigated for 2009 cheating”. *The Philadelphia Public School Notebook*. July 13, 2011.

⁶⁵ Socolar, Paul. “District statement: We never got 2009 report”. *The Philadelphia Public School Notebook*. July 9, 2011.

⁶⁶ King, Larry. “Fairly or not, spikes in test scores haunt some schools”. *The Philadelphia Inquirer*. July 16, 2011.

Any hint of cheating in Philadelphia schools remained at a minimum until May 2011, when *The Philadelphia Inquirer* published a story about teachers anonymously reporting cheating at Roosevelt Middle School and FitzSimons High School.⁶⁷ Before the article was written, Roosevelt was hailed as a District success story where “disadvantaged students could overcome enormous challenges to realize their full potential.”⁶⁸ Under then-Principal Stefanie Ressler, Roosevelt increased proficiency rates by over 50% in both Math and Reading between 2008 and 2010. At a Principal’s meeting in 2010, Ressler was singled out by Superintendent Arlene Ackerman and Governor Ed Rendell for her accomplishments.⁶⁹ As described in this paper’s introduction, Ressler herself commented, “if you think miracles can’t happen, look at us.”⁷⁰ In the summer of 2011, Ressler was reassigned to Wilson Middle School, “considered a better assignment” as Wilson is a high achieving school.⁷¹

Several teachers at Roosevelt contacted *The Inquirer* and detailed multiple test-security violations. According to their accounts, “answers [were] written on a blackboard...senior staff encourag[ed] teachers to drill certain concepts they knew appeared on the exam,” and students “with test booklets and answer sheets out [were] engaged in conversation with the principal.”⁷² The School District began an internal investigation, which included interviews with several teachers at the

⁶⁷ Graham, Kristen; Purcell, Dylan. “City school’s fast rising test scores questioned”. *The Philadelphia Inquirer*. May 1, 2011.

⁶⁸ Ibid.

⁶⁹ Graham, Kristen. “Philadelphia schools celebrate successful scores”. *The Philadelphia Inquirer*. Aug. 21, 2010.

⁷⁰ Ibid.

⁷¹ Graham, Kristen. “Phila. District rejects claims of cheating at 2 schools”. *The Philadelphia Inquirer*. July 7, 2011.

⁷² Graham & Purcell, “City school’s fast-rising test scores questioned”.

school. One teacher commented, “what we all said [of Ressler and Roosevelt] was pretty damning”.⁷³

FitzSimons High School also had a teacher come forward to *The Inquirer* to report wrong-doings at their school. The teacher walked into the newsroom with a copy of two PSSA exams that would be administered to students the following day—a clear violation of a policy that states test booklets must be “kept under lock and key” until the day of the exam.⁷⁴ Furthermore, administrators allegedly asked teachers to look through tests and coach students on specific topics or problems that they knew would be tested.

On July 7, 2011, the School District concluded its investigation and stated that “allegations of state test-score cheating at [Roosevelt and FitzSimons] were unfounded.”⁷⁵

Cheating Scandal Unravels

Finally, two years after the original publication of the erasure analysis report, the details of the cheating scandal were released to the public. Presumably as a result of *The Inquirer’s* reports, another publication—*The Philadelphia Public School Notebook*—asked the Pennsylvania Department of Education (PDE) in May 2011 if they had a history of commissioning forensic analyses of PSSA tests. “To our surprise,” *The Notebook’s* Paul Socolar would later comment, “we received a wealth

⁷³ Graham, “Phila. District rejects claims”

⁷⁴ Graham & Purcell, “City school’s fast-rising test scores questioned”.

⁷⁵ Graham, “Phila. District rejects claims”

of statistical information regarding the 2009 test administration.”⁷⁶ The PDE staff member who responded to *The Notebook* “apparently [was not] fully aware” that he handed over the 2009 erasure analysis report.⁷⁷ On July 8th, just a day after the District found no wrongdoing at Roosevelt, *The Notebook* released the erasure analysis report and its implications to the public. Roosevelt was flagged for suspicious erasures in 2009 for both grade levels that they tested in reading and math. The probability that the school’s 7th grade reading erasure marks occurred by chance was 1 in 100 trillion.⁷⁸

PDE Secretary Ronald Tomalis quickly ordered forensic analyses from PSSA tests conducted in 2010 and 2011.⁷⁹ After analyzing the forensic data from all three years (2009-2011), the state derived new criteria for flagging schools and focused their investigations to 53 schools in Philadelphia. The Data Recognition Corporation’s analyses had led to 88 schools marked for suspicious erasures in Philadelphia for 2009 alone. For unknown reasons, the State later flagged a school-grade-subject combination if “at least 10 percent of student response sheets had five or more wrong-to-right erasures”, according to *The Notebook* (the average test sheet has less than one).⁸⁰ It is likely that the State dropped investigations for schools flagged originally in 2009 that did not show suspicious trends in either 2010 or

⁷⁶ Socolar, Paul. “Cheating coverage honored”. *The Philadelphia Public School Notebook*. April 2012. Vol 19. No. 5.

⁷⁷ Herold, Benjamin. “Pa. wants 90 schools investigated for 2009 cheating”. *The Philadelphia Public School Notebook*.

⁷⁸ Herold, Benjamin; Mezzacappa, Dale. “2009 report identified dozens of PA schools for possible cheating”. *The Philadelphia Public School Notebook*. July 8, 2011.

⁷⁹ Graham, Kristen. “Education chief orders reviews of state exams since ‘09”. *The Philadelphia Inquirer*. July 14, 2011.

⁸⁰ Herold, Benjamin; Mezzacappa, Dale. “New evidence: Philadelphia test cheating is likely far worse than previously revealed”. *The Philadelphia Public School Notebook*. Aug. 16, 2012.

2011. Although PDE initially stated that the 2010 and 2011 forensic analyses reports would be released to the public, the data are still confidential in what has turned into a secretive investigation.⁸¹

The School District worked to increase test monitoring at all of its 250 schools (non-charters). While training and monitoring increased at all schools, special focus was given to the 53 state-flagged schools. During 2012 PSSA testing, staff from the Office of Accountability monitored flagged schools directly; at some schools, “testing material [was] limited” to these central office employees.⁸² Furthermore, PDE prohibited teachers from proctoring the PSSAs to their own students—a measure only given to Philadelphia.⁸³ The State and the District intended the identity of the 53 flagged schools to remain unknown; however, just days before students took the 2012 PSSA tests for reading and math, the *Inquirer* released and widely publicized the names of all 53 schools.⁸⁴

The Bottom Line

When the dust cleared and the 2012 PSSA test results were finally released in September of 2012, Philadelphia schools saw overall declines for the first time in nine years (see Figure 2). Scores declined for every grade level and student demographic group for both reading and math. The next parts of this study aim to

⁸¹ Herold, Benjamin. “State cheating probe was widened to involve more than 50 District schools”. *The Philadelphia Public School Notebook*. Feb. 16, 2012.

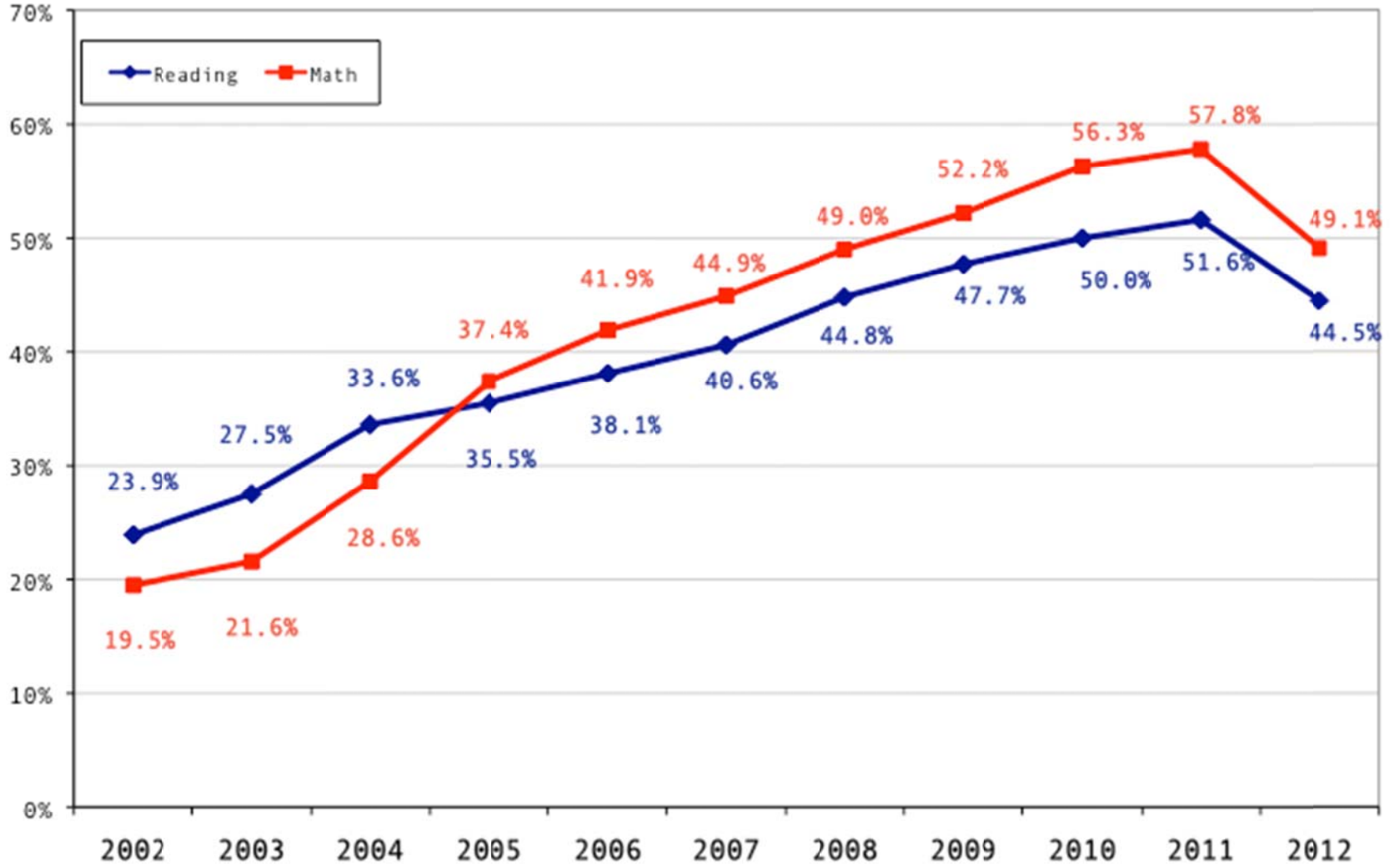
⁸² Ibid.

⁸³ Mezzacappa, Dale. “State prohibits Philly teachers from administering PSSA to their own students”. *The Philadelphia Public School Notebook*. Feb. 28, 2012.

⁸⁴ Graham, Kristen; Purcell, Dylan. “‘Compelling’ evidence of cheating in many Phila. Schools”. *The Philadelphia Inquirer*. Mar. 11, 2012.

shed light on the statistical relationships between a variety of variables and the declines.

Figure 2: PSSA Scores for the School District of Philadelphia (2002-2012)



V. Data

The data used in this study can be placed in three separate categories: test score data, school information data, and cheating data.

Test Score Data

PSSA test proficiency rates on the district, school, and school-grade level are available to the public. Test data from 2011-2012 school year is available in the form of detailed charts and tables on Pennsylvania's AYP website (<http://paayp.emetric.net>). Test data going back to the 1995-1996 school year is available on the Pennsylvania Department of Education's website.

This study analyzes test scores from 2007-2012 for all 246 District-run schools in Philadelphia for both reading and math in order to compare scores across schools before, during, and after proven testing irregularities took place. Test scores were collected on the school level and on the school-grade level. For example, Samuel Powel Elementary School (K-4), has only two tested grades (grades 3 and 4). On the school level, Powel's combined proficiency rates for all tested grades are taken into account. On the school-grade level, changes in scores within specific grades from year to year are analyzed (i.e. Grade 3 reading scores in 2011 compared to Grade 3 reading scores in 2012). For high schools, proficiency rates on the school and school-grade level are equivalent as Grade 11 is the only tested grade. Pennsylvania tests grades 3-8 and 11 every year for reading and math.

School-Specific Data

School information included in this study includes budget data, enrollment data, and AYP status.

Yearly school budgets from 2009 to 2012 were collected from the School District of Philadelphia’s website. The School District lists each school’s funding allotment into three categories: full-time staff subsidy, non-full time subsidy, and per-teacher allotment subsidy. This study uses the combined total of these three subsidies. Also included is the change in budgets from year to year. Principal data from 2009-2011 was similarly attained from the School District’s website.

Table 4: Summary Statistics of School Budgets

	2009	2010	2011	2012
Budget (Mean, N = 246)	\$4,589,385	\$4,887,915	\$5,081,875	\$4,397,290
\$ Change in Budget (Mean, N= 246)		+\$298,530	+\$193,969	-\$684,584
% Change in Budget (Mean, N = 246)		+6.42%	+5.01%	-14.09%

Note that while test score data was available in the school and school-grade level, budget data in this study is limited to the school level. Furthermore, the focus on this study is limited to District-run public schools in Philadelphia. Thus, charter schools—public schools that are independently run—are not considered.

Cheating Scandal: Flagged School and Grade Levels

The data on which schools and grade levels within schools were flagged for cheating came from *The Philadelphia Public School Notebook*, which published the details of the investigation last August.⁸⁵ As previously stated, 53 schools are currently under investigation for having suspiciously high erasure markings on PSSA tests for multiple grades/subjects between 2009-2011.

As discussed in the following sections, these schools have a wide variety of percentage of grade levels that are flagged (see the Appendix (pg. 86) for a comprehensive list of schools and grade levels flagged). Some schools, such as Roosevelt Middle School, were flagged in every grade-subject-year category possible. Since Roosevelt tested two grades (7th and 8th) over three years (2009-2011) in two subjects (reading and math), they had twelve opportunities to receive flags. Table 10 in the Appendix lists every school under investigation for cheating as well as which grade-subject-year combinations were flagged. Blacked out cells indicate that those grades are not tested at the specific schools. Schools are sorted first by school type (elementary, middle, and high school) and then by alphabetical order.

On school level analyses, even though schools had differing number of flags within their grade levels, they are each treated the same under the subset of flagged-schools. When the *Inquirer* released the names of the 53 schools before the start of 2012 PSSA testing, they did not release the details listed in the Table 10—only the names. Treating flagged schools as one category allows the study room to separate

⁸⁵ "Suspicious erasure patterns at 53 schools". *The Philadelphia Public School Notebook*. Aug. 16, 2012.

the effects of decreases due to the elimination of cheating and decreases due to a negative climate at the publicly shamed and privately monitored 53 schools.

VI. Hypotheses

This study aims to test five main hypotheses about the nature of the 2012 PSSA test score declines in Philadelphia.

#1: Test scores in flagged schools will drop by a greater amount than test scores in non-flagged schools

Strong evidence of testing irregularities was found at 53 District-run schools. Increased security and monitoring was placed on these schools in an attempt to curb this cheating. Since the cheating likely led to an artificial inflation of test scores, the elimination of such cheating should have led to flagged schools having larger proficiency rate decreases than non-flagged schools.

#2: Test scores will drop by a greater amount in those schools that received a greater reduction in funding between 2011 and 2012

The massive District budget cuts in 2012 affected all District-run schools. Because school budgets directly affect the amount of resources a school can attain and many of those resources are connected to student achievement (class sizes, a healthy and safe school environment, etc.), the change in funding in 2012 for a specific school should be directly connected to the change in test scores. Schools that had significant decreases in their budgets should have lower test scores.

#3: Within a flagged school, test scores in flagged grades will drop more than in non-flagged grades

While some flagged schools had 100% of their grades flagged for all three years, most schools had some grades that did not show suspicious erasure markings. Theoretically, because cheating occurred in these schools in the flagged grades, test scores in these flagged grades should drop by more than in the non-flagged grades.

#4: Across non-flagged grades, test scores will drop by more in flagged schools than in non-flagged schools

While non-flagged grades in flagged schools did not experience an elimination of cheating that occurred in flagged grades in the same school, they still experienced the same increased security, monitoring, and public scrutiny because they were part of a flagged school. On another note, while non-flagged grades had no suspicious erasure marks in previous tests, they might have taken advantage of other cheating methods (coaching, answer-pointing, pre-test reveals) not accounted for in the erasure analysis. For these two reasons, these “clean” grade levels at “dirty” schools should have experienced a sharper drop in test scores than grade levels at non-flagged schools.

#5: A change in principal will cause change in the percentage of flagged grades in a school

This hypothesis aims to test the notion that administrators (not just individual teachers) were at fault for the cheating occurring at the 53 flagged schools. If the amount of cheating at a school is connected to the person in charge,

then a change in leadership should lead to a higher or lower percentage of grades flagged (depending on the incoming or outgoing principal).

VII. Research Design

These hypotheses will be tested in four separate steps: a pictorial analysis, regression analyses on the school level, regression analyses on the school-grade level, and a detailed look at how a change in principal affected the percentage of flagged grades in a flagged school.

The pictorial analysis will show quantitatively how scores dropped in flagged schools in 2012 compared to non-flagged schools. Then on the school-grade level, the study will focus on three flagged schools and show how flagged grades scores dropped compared to non-flagged grades. Whilst these graphs and figures will show absolute differences between flagged/non-flagged schools/grades, the conclusions drawn from them are limited because of the small amount of variables present. For example, simple graphs and figures cannot distinguish between the effects of cheating and budget cuts on the declining proficiency rate. Multivariate regressions, however, can isolate, for example, the effects of cheating by controlling for several other variables.

Regressions are used in statistical analyses as a means to understand the effect that one variable has on a dependent variable while holding constant the other variables included in the regression. In this study, the dependent variables are a school's (or school-grade-level's) proficiency rates in math and reading. The independent variables used in this study include the school year, the percent change of budget, flagged status, and specific school or school-grade level. A regression will produce a numerical coefficient for each independent variable that describes how the independent variable affects the dependent variable (proficiency rate),

essentially creating a “best-fit” line equation for a multi-dimensional graph. The more uncorrelated independent variables that are added to the equation, the more each coefficient accurately describes the true relationship between the independent and dependent variables.

Fixed effects are included as explanatory variables to control for the differences in test scores across schools that are constant across time and the differences in test scores across years that are constant across schools. For example, when year-fixed effects are included, every single school year has its own coefficient attached to it. Similarly, when school-fixed effects are used, each school is used as a separate independent variable with its own coefficient. Each of these independent variables is only relevant to the school and school year tested that corresponds to them. When these fixed effects are added, the significance of other coefficients, such as one that relates to a school’s flagged status, is much higher. This allows for stronger conclusions, for example, on how cheating and the elimination of cheating affected schools’ proficiency rates.

The regression tests used in this study are described below, whilst the numerical values of the coefficients are revealed in the next section. Each test equation is listed before discussing its relation to the hypotheses of the previous section. Fixed effects and coefficients are written in italics whilst their corresponding independent variables are written in normal text. Sometimes these independent variables also contain specific identifiers to demonstrate which schools they apply to. Consider the following part of *Test #1*’s regression equation for an explicit explanation:

$(\text{Flagged School Drop in 2012 Coefficient}) * (\text{Flagged School Indicator}) * (\text{Year} == 2012)$

The independent variable in this case is Flagged Schools in 2012, as the coefficient only applies to flagged schools (Flagged School Indicator) if the test was taken in 2012 (Year == 2012). The coefficient will demonstrate how much more scores dropped in flagged schools than in non-flagged schools because it only applies to flagged schools.

Each regression will contain a baseline constant that applies to all schools (or school-grade levels). Unexplained factors account for the gaps in predicted proficiency rates from the regressions and actual proficiency rates experienced. The first seven tests pertain to variables that affect a schools' proficiency rate, while the last test aims to shed light on the possible perpetrators of cheating.

School Level Regressions

Using the following equation, the first test will use schools' scores from 2007-2012 and find a relationship between a school's proficiency rates and whether or not the school was flagged for cheating.

Test 1:

School Proficiency Rate =

Constant + (Year Fixed Effects) + (Flagged School Coefficient)*(Flagged School Indicator) + (Flagged School Trend)*(Flagged School Indicator)*(Year - 2012) +

$(\textit{Flagged School Drop in 2012 Coefficient}) * (\textit{Flagged School Indicator}) * (\textit{Year} == 2012)$
+ Unexplained Factors

The year fixed effect will account for year-to-year fluctuations in proficiency rates (e.g. the general rise in test scores between 2007-2011 and the general decline in 2012 will be controlled). The next three coefficients will only apply to the 53 flagged schools. The *Flagged School* coefficient applies to all 53 flagged schools regardless of the year (If a school is flagged, the Flagged School Indicator equals 1, otherwise, 0). The *Flagged School Trend* coefficient will demonstrate how much more flagged school's scores increased over time before 2012. A positive coefficient signifies that the gap between proficiency rates in flagged schools and non-flagged schools increased every year before 2012. The *Flagged School Drop in 2012* coefficient only applies to the 53 flagged schools if the test was taken in 2012, the year of the declines. To isolate the effects of being a flagged school in 2012, the *Flagged School* coefficient will be added to the *Flagged School in 2012* coefficient. Thus, this equation should show how much advantage flagged schools had during the time cheating was found (2009-2011), but should also show how much more they declined in 2012 compared to non-flagged schools (testing Hypothesis #1).

Test 2:

School Proficiency Rate =

Constant + (*Year Fixed Effects*) + (*Flagged School Coefficient*)*(Flagged School Indicator) + (*Flagged School Drop in 2012 Coefficient*)*(Flagged School Indicator)*(Year == 2012) + Unexplained Factors

Test #2 is very similar to Test #1. In anticipation of Test #3, this test will only include scores from 2011 and 2012. Because only two years are accounted for, the *Flagged School Trend* coefficient is excluded. Like the previous test, Hypothesis #1 is relevant as the combination of the two *Flagged School* coefficients will reveal how much flagged school's scores dropped compared to non-flagged schools in 2012.

Test 3:

School Proficiency Rate =

Constant + (*Year Fixed Effects*) + (*Flagged School Coefficient*)*(Flagged School Indicator) + (*Flagged School Drop in 2012 Coefficient*)*(Flagged School Indicator)*(Year == 2012) + (*Change in Budget Coefficient*)*(Percent Change in Budget) + Unexplained Factors

Test #3 adds one crucial factor to the school level regression equations: the percent change in budget from the previous year. The test will only use scores and budgets changes in 2011 and 2012 to focus on the effects in the last year. Since the test includes separate coefficients for budget changes and cheating status, Hypothesis #1 and #2 will be analyzed using the results.

School-Grade Level Regressions

To test hypotheses #4 and #5, proficiency rates must be analyzed on the school grade level instead of just the aggregate school results. Using test scores aggregated for each grade level at all District run-schools—along with the flagged status for each grade in the 53 flagged schools (see Appendix)—the following regressions should reveal more details about how cheating in 2011 affected schools in 2012.

Test #4:

School-Grade Proficiency Rate =

Constant + (*Year Fixed Effects*) + (*School Fixed Effects*) + (*2011 Flagged Grade Coefficient*)*(Flagged Grade Indicator) + (*2011 Flagged Grade in 2012 Coefficient*)*(Flagged Grade Indicator)*(Year == 2012) + Unexplained Factors

As in the previous tests, *Year Fixed Effects* are included. The *School Fixed Effects* are included in this analysis because proficiency rates in schools between grades and tested years are likely similar and strongly correlated. Having a coefficient for every school examined in the analysis controls for this effect. Because this study aims to focus on the drops in scores between 2011 and 2012 with an emphasis on allegedly cheating schools, the regression tests will contain predictive coefficients for grades flagged in 2011 only. As in the previous tests, the score change in 2012 for grades flagged in 2011 will be the sum of the *2011 Flagged Grade*

and *2011 Flagged Grade in 2012* coefficients. Test #4 will only analyze data from 2011 and 2012.

Test #5:

School-Grade Proficiency Rate =

Constant + (*Year Fixed Effect*) + (*School-Grade Fixed Effect*) + (*2011 Flagged Grade in 2011 Coefficient*)*(Flagged Grade Indicator)*(Year == 2012) + Unexplained Factors

Instead of controlling for similar scores within schools, Test #5 controls for similar schools within school-grade combinations. For example, Roosevelt Middle School has a separate fixed coefficient for Grade 7 and Grade 8. This more precise test uses scores from 2007-2012.

Test #6:

School-Grade Proficiency Rate =

Constant + (*Year Fixed Effect*) + (*School-Grade Fixed Effect*) + (*Flagged School Coefficient*)*(Flagged School Indicator)*(Year == 2012) + (*2011 Flagged Grade in 2012 Coefficient*)*(Flagged Grade Indicator)*(Year == 2012) + Unexplained Factors

Test #6 adds one more variable from Test #5: the school's flagged status. This allows differentiation between non-flagged grades in flagged schools and non-flagged grades in non-flagged schools (Hypothesis #4). This test will use scores from 2007-2012.

Test #7:

School-Grade Proficiency Rate =

Constant + (*Year Fixed Effect*) + (*School-Grade Fixed Effect*) + (*School Fixed Effect in 2012*) + (*2011 Flagged Grade in 2012 Coefficient*)*(Flagged Grade Indicator)*(Year == 2012) + Unexplained Factors

Test #7 uses another measure to control for non-flagged grades in flagged schools (Hypotheses #4). Instead of adding a flagged school coefficient, this test adds a fixed effect for schools specifically for 2012. Thus, for tests taken in 2012, all school-grade proficiency rates have a baseline coefficient along with the *School-Grade Fixed Effect*.

Change in Principal Regression

The last regression test will attempt to discover if principals may have directly participated in cheating. The effect of a change in principal on the percentage of flags in a school will be tested with a regression with only one dependent variable.

Test #8:

Change in Percentage of Flags = (*Change in Principal Coefficient*) + Constant

Note that the change in percentage of flags will be a positive value and can represent both a decrease and an increase in flags. What this test will show is whether a change in principal from the previous year led to a significant change in flags in either direction. Significant effects would imply that some administrations may carry their cheating habits to other schools. For example, if a cheating principal leaves a flagged school in 2011, then the percentage of flagged grades should decrease. The school that receives the cheating principal would then experience a rise in percentage of flagged grades.

VIII. Results

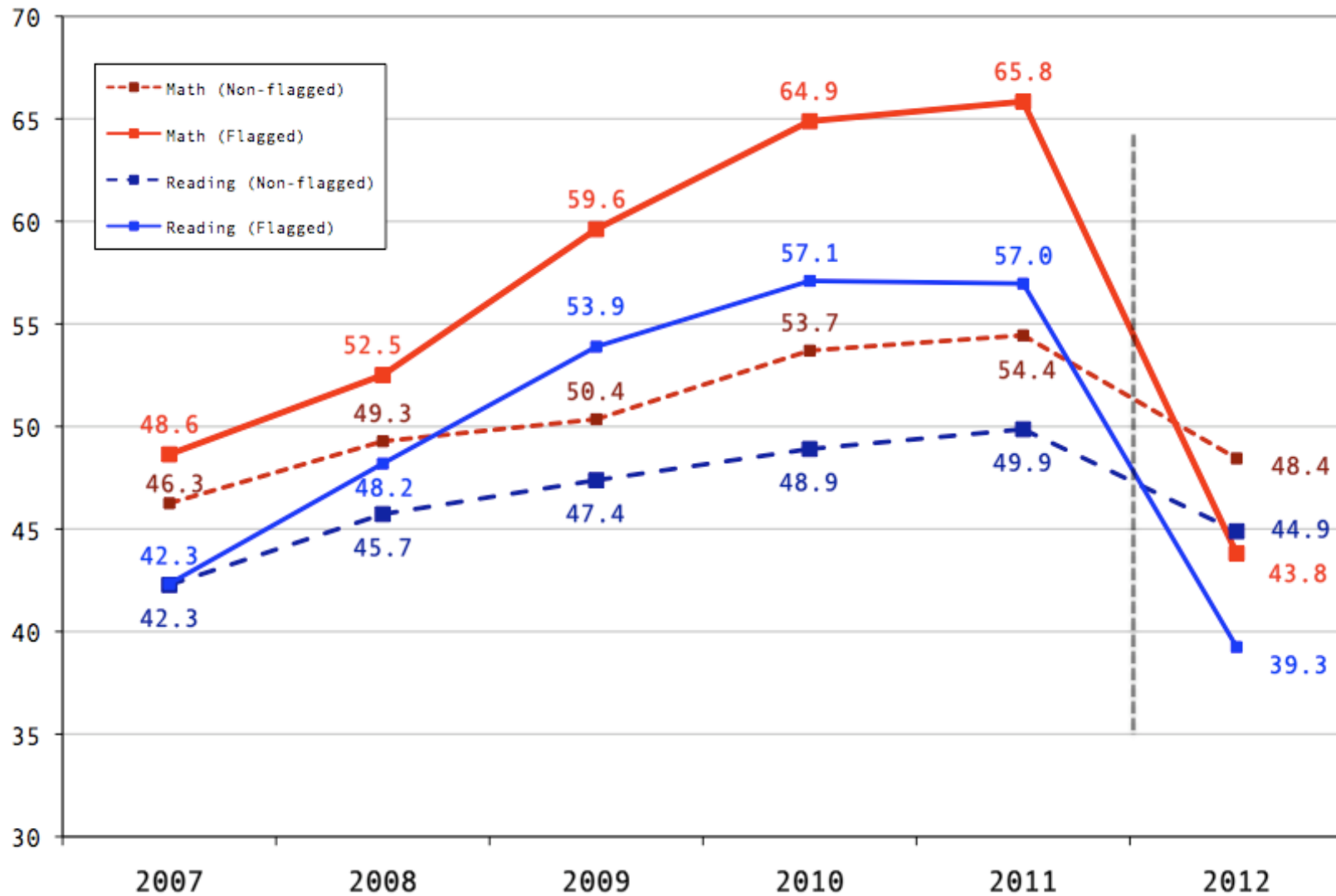
Pictorial Evidence

Graphs of the test score data—disaggregated by cheating status—can provide hints on what the statistical tests will find. As shown in Figure 2 at the end of Part IV, PSSA test scores drastically declined in Philadelphia in 2012. Math proficiency rates dropped 8.7 percentage points while reading proficiency rates dropped 7.1 percentage points. One way to predict if cheating was indeed the cause of this decline is to separate out the scores of the flagged schools and the non-flagged schools. Figure 3 below shows such a disaggregation.

Between 2007-2011, the non-flagged schools (dotted lines) made steady gains in proficiency rates in both reading and math. Meanwhile, flagged schools (solid lines) made significant gains in 2009 and moderate to no gains in 2010 and 2011. Thus, after the 2011 PSSA test, flagged schools' proficiency rates were 11.4 percentage points higher in math and 7.1 points higher in reading.

The 2012 test saw steep declines for both sets of schools, but the flagged schools experienced a much steeper decline. Flagged schools' proficiency rates dropped an astonishing 22.0 percentage points in math and 17.7 points in reading, compared to a 6 point drop in math and 5 point drop in reading for non-flagged schools. The drop was so severe in flagged schools that their proficiency rates were around 5 points lower than non-flagged schools in 2012 for both subjects.

Figure 3: Philadelphia PSSA Proficiency Rates (%) by Cheating Status (2007-2012)



Yet this school-level analysis does not tell the whole story of the elimination of cheating in flagged schools. This significant decrease in flagged schools could have been due to the elimination of cheating practices or other factors connected to the public label of being a “flagged” school. For example, perhaps the increased monitoring that occurred at flagged schools created a climate that was detrimental to a testing environment. These latter factors would theoretically affect the entire school—all grades, all subjects. However, several schools were only flagged for certain subjects and grades. Looking closer at these schools can shed light onto the causes of these decreases by isolating flagged grades/subjects at specific schools.

One way to analyze these school specific cases is to isolate the non-flagged and flagged grades within a flagged school and examine the decreases in 2012 for each category. For the purposes of this part in the study, a grade is considered “flagged” if 50% or more of its testing windows were flagged for irregularities between 2009-2011 (for math and reading).

For example, in Disston Elementary, the flagged grades are 3rd, 4th, and 6th grade while the non-flagged grades are 5th, 7th, and 8th grade (see table below).

Table 5: Disston Elementary School Flags, 2009-2011

	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
Subject	M	R	M	R	M	R	M	R	M	R	M	R
2009	Y		Y									
2010	Y	Y	Y	Y				Y				
2011	Y	Y	Y				Y	Y				
Flagged Grade?	YES		YES		NO		YES		NO		NO	

The schools used in this analysis include Disston Elementary, Kelly Elementary, and John Marshall Elementary, shown in the Figures below.

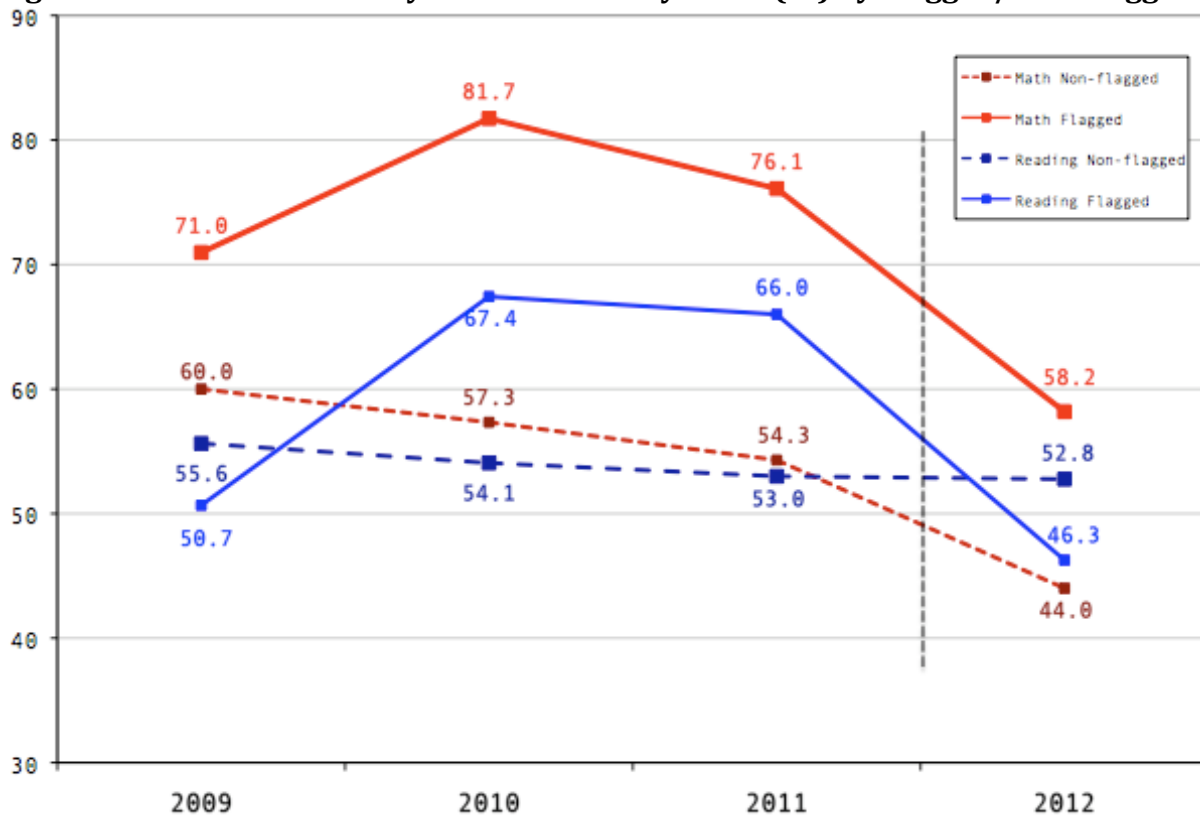
Because different grade levels have separate tests and standards of proficiency, the magnitude of declines *within* grade levels is more important than the declines *between* grade levels. Disston Elementary's flagged grades' (3rd, 4th, and 6th grade) math proficiency rates dropped 23.5% (17.9 percentage points) in 2012, compared to a 19.0% drop (10.3 percentage points) in non-flagged grades (5th, 7th, and 8th grades). The difference in declines is much more pronounced in Disston's reading scores, where flagged grades dropped 29.8% (19.7 percentage points) in 2012, compared to a tiny .4% drop (.2 percentage points) in non-flagged grades (See Figure 4).

Similar differences are found in Kelly Elementary (Figure 5), where flagged grades (3rd and 4th) dropped 29.3% in math and 34.9% in reading in 2012 while non-flagged grades (5th and 6th) only dropped 18.2% in math and 1.0% in reading.

At John Marshall Elementary (Figure 6), a K-5 school, proficiency rates in flagged grades (3rd and 4th) declined 36.5% in math and 39.5% in reading in 2012. The only non-flagged grade, 5th grade, still declined sharply in 2012, but not with the same magnitude as Marshall's flagged grades: 19.8% in math and 28.8% in reading.

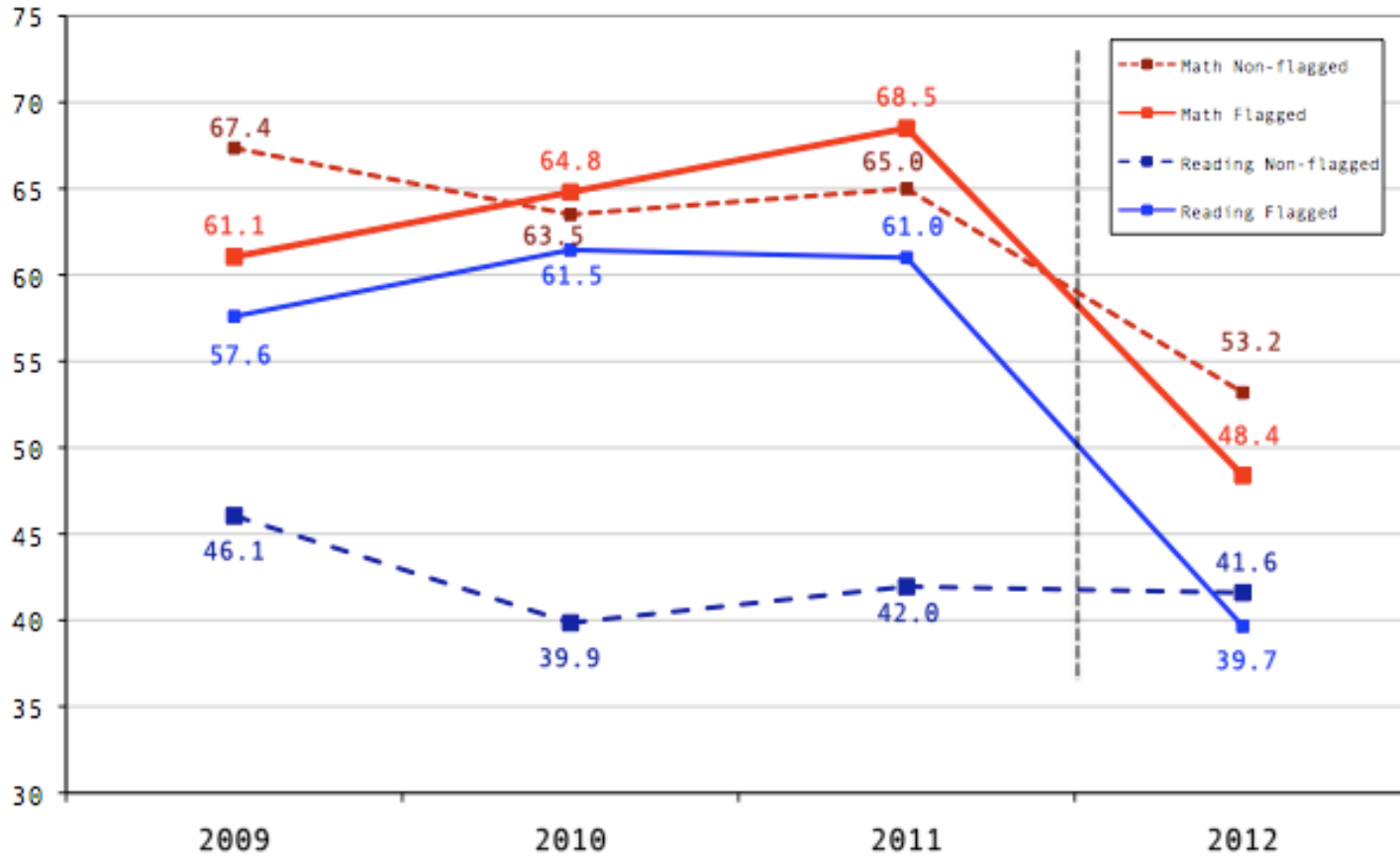
The results from three schools—ideal examples of institutions where alleged cheating only occurred in specific grades and classrooms—imply that while scores sometimes dropped significantly in all grades of a flagged school, the drops were steeper in those grades that contained previous evidence of testing improprieties.

Figure 4: Disston Elementary PSSA Proficiency Rates (%) by Flagged/Non-Flagged Grades (2009-2012)



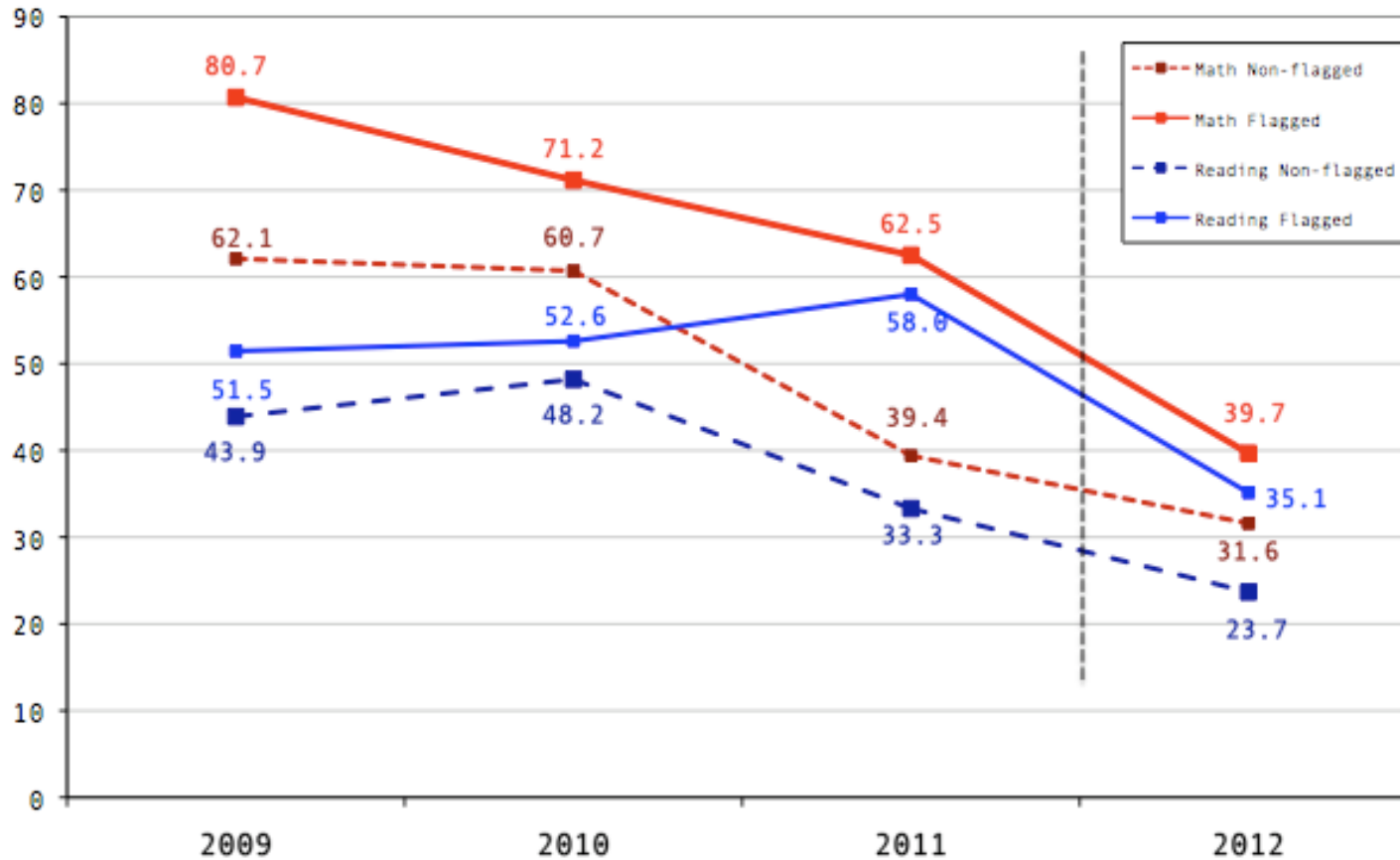
	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
Subject	M	R	M	R	M	R	M	R	M	R	M	R
2009	Y		Y									
2010	Y	Y	Y	Y				Y				
2011	Y	Y	Y				Y	Y				

Figure 5: Kelly Elementary PSSA Proficiency Rates (%) by Flagged/Non-Flagged Grades (2009-2012)



	Grade 3		Grade 4		Grade 5		Grade 6	
Subject	M	R	M	R	M	R	M	R
2009	Y	Y						
2010	Y	Y	Y	Y				
2011	Y	Y	Y	Y				

Figure 6: Marshall, J. Elementary PSSA Proficiency Rates (%) by Flagged/Non-Flagged Grades (2009-2012)



	Grade 3		Grade 4		Grade 5	
Subject	M	R	M	R	M	R
2009	Y		Y	Y	Y	
2010	Y	Y	Y	Y		
2011	Y	Y		Y		

School Level Regression Analyses (See Table 6)

Test #1 shows that a flagged school's proficiency rates were generally 15.3 percentage points higher in Math ($p < .01$) and 10.8 points higher in Reading ($p < .01$) in comparison to non-flagged schools. (The p-value for this test indicates that the probability of such a relationship occurring by chance is less than 1%. Generally, tests are considered statistically significant if the p-value is less than .05.)

Furthermore, between 2007-2011, the gap between flagged and non-flagged schools rose by 2.60 points in Math ($p < .01$) and 1.98 points in Reading ($p < .01$) every year. Thus, the overall flagged school effect for Math was an increase of 2.3 points in 2007 ($15.3 + (-5 * 2.60)$), 4.9 points in 2008 ($15.3 + (-4 * 2.60)$), and so on. In 2012, the combined effect for flagged schools is a 4.63 percentage point decline in Math ($p < .1$) and a 5.64 point decrease in Reading ($p < .05$) compared to non-flagged schools.

Test #2, which only analyzes proficiency rates in 2011-2012, shows similar results. In 2011, flagged schools' proficiency rates were 10.9 percentage points higher in Math ($p < .01$) and 6.76 points higher in Reading ($p < .05$) than non-flagged schools. In 2012, flagged school's proficiency rates were 5.11 points lower in Math ($p < .1$) and 5.97 points lower in Reading ($p < .05$) compared to non-flagged schools.

Test #3 is the same test as *Test #2* except it includes a change-in-budget effect. The change in budget coefficient shows that a 100% decrease in school budgets would lead to a 46.4 point drop in Math ($p < .01$) and a 55.1 point in Reading ($p < .01$). In more practical terms, for every 1% decrease in budget that a school experiences, their proficiency rates decrease by .464 points in Math and .551 points in Reading. Even when controlling for budget differences, the combined 2012

Table 6: School Level Regression Analyses on Philadelphia PSSA Proficiency Rates (Tests #1-#3)

Subject	Math (% Proficient or Advanced on PSSA)			Reading (% Proficient or Advanced on PSSA)		
	1	2	3	1	2	3
Test #						
Years Included	2007-2012	2011-2012	2011-2012	2007-2012	2011-2012	2011-2012
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Flagged School	15.282 (2.863)***	10.903 (2.928)***	11.398 (2.895)***	10.800 (2.638)***	6.761 (2.644)**	7.349 (2.629)***
Flagged School X Year == 2012	-19.915 (2.002)***	-16.009 (1.952)***	-15.944 (1.951)***	-16.439 (1.923)***	-12.726 (1.766)***	-12.647 (1.827)***
Flagged School X (Year - 2012)	2.597 (.475)***			1.976 (.478)***		
% Decrease in School Budget			-46.380 (10.524)***			-55.074 (10.173)***
Constant	46.276 (1.556)***	48.917 (1.491)***	55.332 (2.298)***	42.087 (1.486)***	45.216 (1.444)***	52.834 (2.208)***
Combined Flagged School Effect in 2012	-4.633*	-5.106*	-4.546*	-5.639**	-5.965**	-5.298**
Observations	1464	492	492	1464	492	492
R-squared	0.06	0.08	0.11	0.04	0.06	0.10

Robust standard errors in parentheses

* significant at 10% level; ** significant at 5% level; ***significant at 1% level

Flagged School effect saw decreases. Interestingly, even when controlling for budget decreases, this test shows that flagged schools in 2012 had proficiency rates 4.471 points lower in Math ($p>.1$) and 5.206 points lower in Reading ($p<.05$) compared to non-flagged schools.

School-Grade Level Regression Analyses (See Table 7)

Test #4—which controls for the fact that schools will demonstrate similar scores—shows that in the 2011 PSSA tests, a grade that was flagged in 2011 experienced proficiency rates that were 14.9 percentage points higher in Math ($p<.01$) and 13.2 points higher in Reading ($p<.01$) compared to grades not flagged in 2011. For the 2012 tests, proficiency rates for grades flagged in 2011 were 4.5 points lower in Math ($p<.05$) and 8.5 points lower in Reading ($p<.01$) compared to non-flagged grades.

Test #5 uniquely controls for similar scores experienced within grade levels of the same school. Using scores from 2007-2012, this test shows that for the 2012 tests, a grade flagged in 2011 performed 13.6 percentage points worse in Math ($p<.05$) and 14.4 points in Reading ($p<.01$).

Test #6 shows the difference in scores between flagged grades in flagged schools, non-flagged grades in flagged schools, and non-flagged grades in non-flagged schools. For the 2012 PSSA tests, grade levels that belonged to one of the 53 flagged schools but were not flagged themselves (in 2011) scored 11.5 percentage points lower in Math ($p<.01$) and 9.45 in Reading ($p<.01$) compared to non-flagged grades in non-flagged schools. On the same 2012 tests, grades that were flagged in

Table 7: School-Grade Level Regression Analyses on Philadelphia PSSA Proficiency Rates (Tests #4-#7)

Subject	Math (% Proficient or Advanced on PSSA)				Reading (% Proficient or Advanced on PSSA)			
	4	5	6	7	4	5	6	7
Test #								
Years Included	2011-2012	2007-2012	2007-2012	2011-2012	2011-2012	2007-2012	2007-2012	2011-2012
Year Fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed	Yes	No	No	No	Yes	No	No	No
School-Grade Fixed	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School-2012 Fixed	No	No	No	Yes	No	No	No	Yes
Flagged School Tested X Year == 2012			-11.460 (1.666)***				-9.447 (1.619)***	
2011 Flagged Grade	14.865 (1.637)***				13.208 (2.281)***			
2011 Flagged Grade X Year == 2012	-19.346 (2.409)***	-13.578 (2.493)**	-3.295 (2.631)	-4.423 (2.445)*	-21.626 (2.354)***	-14.413 (2.437)***	-6.266 (2.435)**	-8.496 (2.401)***
Constant	29.817 (0.299)***	36.843 (.576)***	36.668 (.586)***	31.101 (15.582)**	25.547 (0.260)***	26.676 (.492)***	26.486 (.501)***	33.703 (14.519)**
Combined Flagged Grade Effect in 2012	-4.481**		-14.755***		-8.418***		-15.713***	
Observations	1940	5708	5708	1940	1940	5708	5708	1940
R-squared	0.76	0.77	0.78	0.92	0.71	0.80	0.81	0.93

Robust standard errors in parentheses

* significant at 10% level; ** significant at 5% level; ***significant at 1% level

2011 performed 3.30 points lower in Math ($p > .1$) and 6.27 points lower in Reading ($p < .05$) than non-flagged grades in flagged schools. Compared to non-flagged grades in non-flagged schools, grades flagged in 2011 performed 14.8 points worse in Math ($p < .01$) and 15.7 points worse in Reading ($p < .01$).

Test #7, which has school fixed effects for 2012, shows that for the 2012 tests, grades flagged in 2011 scored 4.42 percentage points lower in proficiency for Math ($p < .1$) and 8.50 points lower in Reading ($p < .01$) than grades not flagged in 2011 (regardless of their school's flagged status).

Principal Analysis

For both 2010 and 2011, if a flagged school changed principals from the previous year, they experienced a 17.4% change in the percent of grades flagged (see Table 8). This could mean either a 17.4% increase or decrease as this analysis only calculated the absolute value of the change in percentage of grades flagged.

Table 8: Regression Analysis of Change in Flagged Grades on Change in Principal

	Change in % grades flagged
Change in Principal	17.378 (6.093)***
Constant	20.622 (2.367)***
Observations	106
R-squared	.07

Standard errors in parentheses

* significant at 10% level; ** significant at 5% level; ***significant at 1% level

IX. Analysis

Almost all of the results described in the above section support each of the five hypotheses listed in Part VI with statistical significance.

#1: Test scores in flagged schools will drop by a greater amount than test scores in non-flagged schools

Figure 3 clearly shows that in absolute terms, flagged schools dropped by a much wider margin than non-flagged schools in 2012. *Test #s 1-3* on Table 6 show that the decline in proficiency rates in flagged schools in 2012 (*Flagged School in 2012* coefficient) was anywhere from 12 to 20 percentage points higher than non-flagged schools (statistically significant with at least 99% confidence). *Test #3* shows these results even after accounting for drastic budget decreases. In absolute terms, flagged schools' proficiency rates dropped so drastically that reading scores were statistically significantly lower in flagged schools than in non-flagged schools in 2012. This can also be seen in Table 6 as flagged schools in both subjects performed worse than their non-flagged counterparts (though the gap was not statistically significant for Math for *Test #s 1-3*).

#2: Test scores will drop by a greater amount in those schools that received a greater reduction in funding between 2011 and 2012

Test #3 found a statistically significant link between a decrease in school budgets and a decrease in proficiency rates for both subjects. Specifically, a 1% decrease in budgets corresponds to a .464 proficiency rate drop in Math and .551

drop in Reading (both significant with over 99% confidence). In the 2012 school year, the average District school experienced a budget decrease of 14.09% (See Table 4 above). Thus, the data imply that budget decreases in 2012 led to a mean drop of 6.536 percentage points in the Math proficiency rate and 7.761 percentage points in the Reading proficiency rate (after controlling for schools' flagged status).

#3: Within a flagged school, test scores in flagged grades will drop more than in non-flagged grades

Test #s 4-7 all show that flagged grades dropped more (except Math results in *Test #7*) than non-flagged grades. Specifically, *Test #s 4* and *5* demonstrate that flagged grades dropped more than non-flagged grades regardless of whether or not the non-flagged grade belonged to a flagged school.

The hypothesis that directly compares flagged and non-flagged grades within flagged schools was tested by *Test #s 6-7*, with *2011 Flagged Grade in 2012* the key coefficient. While the Math results show that flagged grades in flagged schools decreased by 3-4.5 percentage points more than non-flagged grades in flagged schools, the results are not significant at the 95% confidence level (*Test #7* is significant at the 90% level). The Reading results are statistically significant and show that flagged grades dropped around 6.0 to 8.5 percentage points more than non-flagged grades in flagged schools. The Readings results strongly support Hypothesis #3 while the Math results are inconclusive.

#4: Across non-flagged grades, test scores will drop by more in flagged schools than in non-flagged schools

Test #6 specifically looks at this hypothesis as the *Flagged School Tested in 2012*, when isolated, shows how non-flagged grades in flagged schools performed in comparison to non-flagged grades in non-flagged school. On the 2012 math test, proficiency rates for non-flagged grades in flagged schools declined 11.460 percentage points more than in non-flagged grades in non-flagged schools ($p < .01$). On the reading test for the same year, proficiency rates declined 9.447 more points in non-flagged grades in flagged schools than in non-flagged grades in non-flagged schools ($p < .01$).

The results of the two previous hypotheses beg the question: why did a school's flagged status affect a grade-level's proficiency rates much more than a grade-level's own flagged status? One possible suggestion points to the effect of test monitors and public shaming. Increased test security measures were placed at all District-run schools; however, flagged schools received increased attention including testing monitors that patrolled the hallways and looked into tested classrooms. As mentioned in Part IV, the *Philadelphia Inquirer* released the names of the 53 flagged schools right before the 2012 PSSA tests were administered. Did either (or both) of these occurrences—through intimidation or by some other means—affect staff and students enough on test day that scores were lowered?

A more pessimistic explanation points at the method of flagging grades. The sole measure for flagging grades consisted of a suspiciously high number of wrong-to-right erasure markings. This method had no way of finding other possible

methods of cheating. It is possible that flagged schools employed multiple means of cheating, and only the flagged grades used erasure markings while the non-flagged grades used other means. These other methods would then be discouraged by monitoring in 2012 the same way as erasure markings were. The next section of this study will discuss other possible ways of cheating in more detail.

It is unclear why results are generally more significant in Reading than in Math. Amongst the 53 flagged schools, the 2012 declines in Reading and Math were nearly identical. On the school-grade level, out of the 189 total flagged grade-subject combinations in 2011, the majority of them (110 combinations or 58.2%) were flagged in Math. Therefore, a lack of cheating data is not to blame. The difference in significance implies that in the testing room, a school's cheating status affected each subject differently. If teachers and administrators were able to more easily cheat in ways other than erasing wrong answers in the math test, than perhaps a non-flagged grade in math decreased just as much as a flagged grade in math because cheating was eliminated in both grades, though only recorded in the one with suspicious erasure markings. Alternatively, perhaps the "stigma" of cheating that existed in the publicly named 53 flagged schools affected students more when taking the Math test than when taking the Reading test. Without further research, analysis, and concrete data, the difference between effects of a "flagged status" on a specific subject test is up for speculation.

#5: A change in principal will cause change in the percentage of flagged grades in a school

Test #8 shows strong support for this hypothesis, and the results suggest that top administrators may have had a role in the cheating that allegedly occurred at flagged schools. A principal may have brought in or out their cheating habits when a principal change occurred.

One extreme example comes from Edward Penn’s principal career (Table 9). Penn was principal at Thurgood Marshall Elementary during the 2008-2009 and 2009-2010 school year. Because of the impressive gains in proficiency rates that he oversaw at T. Marshall, Penn was promoted and then placed in charge of Clemente Middle School, a Promise Academy, in the 2010-2011 school year.⁸⁶ The State’s erasure analysis of PSSA tests from 2009-2011 revealed multiple flags within grade levels at Marhsall during Penn’s tenure. However, during the first year of testing after Penn left, Marshall received no flags. Before Penn arrived at Clemente, the school was innocent of testing improprieties in the eyes of the state. Upon Penn’s arrival, 75% of grade-subject combinations were flagged for suspiciously high erasure marks.

Table 9: Schools led by Edward Penn with Grades Flagged, 2009-2011

Grade Levels	Schools	3rd		4th		5th		6th		7th		8th		Percentage Flagged	Principal
		M	R	M	R	M	R	M	R	M	R	M	R		
Marshall, T.	09	Y	Y	Y	Y			Y	Y	Y				58.3%	Penn
	10	Y	Y	Y							Y			33.3%	Penn
	11													0.0%	<i>Other</i>
Clemente	09													0.0%	<i>Other</i>
	10													0.0%	<i>Other</i>
	11					Y	Y	Y	Y	Y	Y			75.0%	Penn

⁸⁶ Herold, Benjamin; Mezzacappa, Dale. “District promoted principals from schools under suspicion for cheating.” *The Philadelphia Public School Notebook*. Sep. 18, 2012.

Penn is certainly not the only case where a school's flags disappeared after a principal left. The data in no way proves that Penn himself orchestrated cheating, though it strongly suggests that some characteristic of his administration fostered testing improprieties. While the statistics from *Test #8* and these case studies prove no wrongdoing, they strongly suggest that the problems from alleged cheating emanate from multiple levels of school leadership.

X. Discussion: Limitations and the Blaming Game

Limitations

This study could be improved with higher quality data. PSSA classification rates (Below Basic, Basic, Proficient, and Advanced) are based on students' scaled scores. For example, a 3rd grade student earning a 1181 scaled score on the Math test would be considered Proficient while the same student with a 1179 scaled score would be Basic. Because these somewhat arbitrary cut scores are of extreme importance yet yield few similarities between students just below or above them, using scaled scores instead of proficiency rates would lead to more precise conclusions on the trends of scores from year to year. Although, because each subject level test has different cut off scores for every tested grade level, a school-level analysis would not be possible as scaled scores aren't comparable between grades.

Erasure data analysis is also incomplete. *The Philadelphia Public School Notebook* released a detailed school-grade level analysis on 2009 erasure analyses when the broke the cheating scandal in 2011. In 2009 alone, 88 schools had suspiciously high erasure marks. The State then conducted similar analyses for 2010 and 2011, and in the end 53 schools were flagged for further investigation encompassing 2009-2011. The State has refused to release the erasure analyses for 2010 and 2011. How did 88 schools in 2009 turn into 53 schools in 2009-2011? Did the state eliminate schools on the 2009 list because they didn't reappear on the 2010 or 2011 lists? Or did they pick and choose the 53 flagged schools by some

other means? Without this knowledge, the credibility of a schools' flagged status is slightly in doubt.

Furthermore, presumably the state (and District) also have more detailed erasure analysis that would reveal which specific classrooms had suspicious erasure markings and which students' tests were used. Tracking cheating data on this deeper level would yield stronger and more meaningful results. Unfortunately, only the school and school-grade level analyses are available to the public as the investigations are still ongoing.

The Blaming Game

The finger-pointing for who or what caused the 2012 test score declines can start from the highest levels of education policy and move down. Did high-stakes testing legislated by national politicians create an environment that led to cheating? Or did Governor Corbett and the Pennsylvania General Assembly inevitably cause the declines with massive education budget cuts that they orchestrated? What about the external effects such as monitoring and public scrutiny that affected the entirety of flagged schools, even grade levels not flagged for cheating? Or should the blame rest solely on those administrators who actually conducted the cheating? While there is plenty of blame to rightly pass around, cheating undoubtedly occurred at Philadelphia schools and those responsible have yet to face any kind of repercussions.

High Stakes Testing and No Child Left Behind

No Child Left Behind deserves credit for being the first national law of its kind to pay attention to the academic achievement of *all* student subgroups, including minority subgroups, economically disadvantaged students, English language learners, and special education students. However, the unrealistic expectation for meeting 100% proficiency by 2014 has caused a hostile environment in school reform where schools are publicly labeled “failures” if they fail to meet increasing academic targets. To try and avoid these labels, states and school districts alike implement quick-fix reform initiatives that often lead to drastic turnaround of teachers and staff. Teachers and principals are thus left with the prospect of losing their jobs if they don’t improve test proficiency rates by a significant margin year after year. The incentive to manipulate test scores may then overcome the desire to naturally improve students’ academic achievement.—as the case has been in Philadelphia, Atlanta, Washington DC, and other major urban school districts.

The successes touted by high-stakes testing and NCLB advocates are often misleading. As previously discussed, every single state has lower proficiency standards on their standardized tests than the national standards used in the National Assessment of Educational Progress tests (with the exception of Massachusetts). While 49 out of 50 states reported that over 50% of their 4th graders were proficient in Math in 2009, the NAEP tests found that *all* states were under this threshold.

Not only are the high proficiency rates a myth, but gains made between years are also fabricated by teaching to the test, testing improprieties, and other factors.

For example, the PSSA proficiency rate in Philadelphia for 8th Grade math increased by 6.1 percentage points between 2009 and 2011 (50.5% to 56.4%). According to NAEP results, during that time the Philadelphia 8th grade math proficiency rate only rose by 2 percentage points (16% in 2009, 18% in 2011).⁸⁷ Are Philadelphia's 8th graders truly increasing their achievement in math, or are they merely learning test-taking techniques more from teachers or improving due to soft and hard forms of administration cheating?

When it inevitably arrives, the next national education reform legislation should take note at the successes and failures of NCLB. High expectations should not be abandoned for the sake of practicality and hurting feelings; however, legislation should create incentives for teachers and administrators to work for their students by supporting troubled schools and taking into account the different out-of-school factors that students may bring in.

Unfortunately, the current rhetoric of education reform is going in the opposite direction. The Obama Administration has granted states No Child Left Behind waivers if and only if the states develop a teaching evaluation system that is directly connected to student performance (among other stipulations).⁸⁸ The notion has good intentions but may lead to further cheating if there is too much stress on student achievement on these evaluations and if professional development is not prioritized. Multiple studies show that out-of-school factors such as poverty account

⁸⁷ U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 and 2011 Mathematics Assessments.

⁸⁸ White, Jeremy. "Education Reform: As Obama Pushes for Teacher Evaluations, Little Agreement on How." *International Business Times*. Oct. 28, 2011.

for the majority of test score variations while differences in teacher quality account for around 10% of student achievement.⁸⁹ A comprehensive teacher evaluation system must keep these statistics in mind or else teachers will be overburdened with the task of overcoming poverty completely on their own. Only evaluations, professional development, strong leadership, and adequate financial support can level the playing field. One wonders how the outlook of public education would look today if Gore and his educational vision had won the 2000 election.

Budget Cuts from Governor Corbett and the PA General Assembly

The 2011-2012 budget cuts for basic education were a significant factor in the decline of proficiency rates. By Pennsylvania's own standards, schools in the state are drastically underfunded—by over \$4.5 billion according to the Costing-Out commissioned by the General Assembly and released in 2007. Pennsylvania slashed funding by almost \$1 billion for the 2011-2012 school year, and Philadelphia received the brunt of these cuts as special funding formulas that target areas with higher poverty concentrations were eliminated.

Governor Corbett called these cuts part of a “reality-based budget” to reign in excessive spending from previous governors.⁹⁰ Yet there were many possibilities to make room in his budget for a less hurtful cut to education. Corbett turned down the

⁸⁹ Hanushek, Eric A.; Kain, John F.; Rivkin, Steven G.. “Teachers, Schools, and Academic Achievement.” *National Bureau of Economic Research*. Aug. 1998

Berliner, David. “Poverty and Potential: Out-of-School Factors and School Success”. *The Great Lakes Center for Education Research & Practice*. March 2009.

Goldhaber, Dan. “The Mystery of Good Teaching”. *Education Next*. Spring 2002. Vol. 2, No. 1.

Ravitch, Diane. “The Myth of Charter Schools”. *The New York Review of Books*. Nov. 11, 2010.

⁹⁰ Olsen, Laura; Mauriello, Tracie. “Corbett’s state budget plan cuts \$866 million”. *Pittsburgh Post-Gazette*. Mar. 9, 2011.

opportunity to tax the gas derived from the recently discovered Marcellus Shale due to fears of scaring away business. He also refused to spend a penny of the \$785 million revenue surplus, instead saving it for a “rainy day”.⁹¹ Meanwhile, the state budget allocation for correctional facilities increased by 11%.

Clearly Governor Corbett did not prioritize education, and his budget for the 2012-2013 did little to staunch the financial bleeding from the previous year. While basic education funding remained the same from the year before, Accountability Block Grants—money that aided special programs in needy Districts such as Philadelphia—were eliminated.⁹²

The future of education funding may be clearer. With about a year and half left before the 2014 election, Governor Corbett’s polling numbers are foreboding. Corbett—a first-term Republican—has an overall approval rating of 33% in the state, with 58% disapproving of his job as Governor (as of March 12, 2013). According to Public Policy Polling, these rates make him the third most unpopular governor in the country.⁹³ Furthermore, he trails in head-to-head match ups against five potential democratic opponents, some of which are barely known in the state. Perhaps most worrying for him is that more Republicans would prefer another candidate to run in 2014 in his place. Unless he can save his already struggling campaign, a Democratic candidate who is likely to sympathize with education advocates will take the governorship, and the General Assembly will likely follow.

⁹¹ Markosek, Joe. “Revenue Surplus Exceeds \$785 million”. *PA House Appropriations Committee: Fast Facts*. July 1, 2011.

Butterfield, Sam. “Spend rainy-day fund, advocates tell governor”. *Pittsburgh Post-Gazette*. June 17, 2011.

⁹² Hardy, Dan. “Corbett’s Pennsylvania budget holds the line on school spending for the most part”. *The Philadelphia Inquirer*. Feb. 8, 2012.

⁹³ Jensen, Tom. “Corbett in dire shape for reelection”. *Public Policy Polling*. Mar. 12, 2013.

Poorer school districts such as Philadelphia have hope that they may experience funding relief in the near future.

As *Test #3* strongly implied, the massive decrease in funding likely directly affected the declines in proficiency rates in Philadelphia. When discussing the declines, the State Department of Education is quick to point towards the cheating scandal as the sole cause for Philadelphia's demise.⁹⁴ While cheating has a major role in what happened, the State should not hide the fact that their decision to cash-strap districts led to a severe lack of resources that affected students on test day.

Outside pressures from Testing Monitors and the General Public

While testing monitors and the media certainly created a negative atmosphere at flagged schools, their work can hardly be called blameworthy. As a news organization, *The Inquirer* had every right to publish the names of the 53 flagged schools when they received the list—even though the state asked them not to. Furthermore, the only way for the School District to ensure the credibility of the 2012 PSSA tests after the scathing cheating scandal broke out was to increase security throughout the District. Most probably, if these methods existed before 2012, administrators and teachers would not have had the opportunity to participate in cheating in the first place. From the stories told by whistle-blowers from Roosevelt Middle School and FitzSimons High School (now closed), cheating was obvious, open, and easily detectable. Yet the lax security in place was only improved after *The Public School Notebook* released the 2009 report.

⁹⁴ Graham, Kristen; Purcell, Dylan. "PSSA-cheating reforms yield lower scores across Pa.". *The Philadelphia Inquirer*. July 30, 2012.

Cheaters

Beyond a reasonable doubt, several schools in Philadelphia participated in cheating from 2009-2011 by erasing students' wrong answers and replacing them with correct ones. The insurmountable evidence shows that at many schools, test scores did not accurately represent student achievement. Those who participated in cheating may have been under immense pressure to raise proficiency levels. Yet their actions of misrepresenting students' knowledge ended up hurting those students who needed targeting instruction. Furthermore, their actions are an insult to the majority of teachers and administrators who tried to honestly raise student achievement and faced the consequences of the results of their efforts. The entire District, their employees, and their students have suffered because of the actions of the minority, and it will be a long time before the District can start to build up a more positive reputation.

What's perhaps most worrying about this scandal is that it only uncovered the most detectable forms of cheating. Schools may have got away with a multitude of other kinds of cheating before the District increased security in 2012. Easier and far less-detectable ways of cheating include: opening test booklets early in order to teach specific concepts and problems before the test, coaching students to answer a question correctly (whether by implication or simply telling them the correct response), and allowing students to take the test in groups or with outside help. These forms of "soft" cheating can only be detectable through observation. Higher scores than normal derived from these scores can be—and have been—explained away as implementations of reform initiatives (e.g. professional development, a new

curriculum, etc.). How much cheating actually occurred in Philadelphia schools before 2012? Since the analysis conducted by the state only includes erasure methods, the answer will likely remain unknown.

Although much remains a mystery, the State and District have damning evidence that specific teachers and administrators were involved in testing improprieties. In fact, the *Inquirer* has reported that some principals, assistant principals, and teachers from the 53 flagged schools “have confessed to investigators.”⁹⁵ Yet as of the time of this study, not a single District employee has faced repercussions for their involvement in the scandal.

In fact, the only possible firing of an employee that may have been connected to the cheating scandal occurred last summer under suspicious circumstances. According to *Notebook* reporter Benjamin Herold, Daniel Piotrowski—the former Executive Director for Accountability and Assessment—was assigned to personally monitor flagged schools in 2012 (along with several other central office staff). Piotrowski was placed at Wagner Middle School and found seventeen testing violations, including observing a teacher “coaching students on how to answer test questions.”⁹⁶ After reporting the violations, Piotrowski was removed from the school as a monitor and his allegations were quickly dismissed. Other schools that were reported for fewer violations had to undergo an investigation process that included interviews and a detailed report. The incidents at Wagner—a school where the principal in 2009 held the Chief Academic Officer position in the District at this

⁹⁵ Graham, Kristen & Purcell, Dylan. “PSSA-cheating reforms...”

⁹⁶ Herold, Benjamin. “New cheating concerns at Philly school met with ‘baffling’ response”. *The Philadelphia Public School Notebook*. Oct 23, 2012.

time—produced no such process. The District later fired Piotrowski in July 2012 for unknown reasons (neither Piotrowski nor the District commented in Herold’s story).

No matter the exact circumstances surrounding the events at Wagner, the District and State have failed to act swiftly in punishing those responsible for cheating. Both parties claim that the investigations are still ongoing. Over a year and a half has passed since the scandal first broke out, yet the only reshuffling of teachers and principals that has occurred has been for reasons disconnected to cheating. District principals who are highly implicated in the scandal still hold their jobs. Neither the District nor its students can move on and overcome the negative emotions brought until the State of District cleans house. Philadelphia students deserve to learn in peace without the politics of education policy hindering their progress in the classroom.

XI. Conclusion

On March 29, 2013, after a two and a half year investigation, the State of Georgia indicted 35 educators who participated in organized cheating in the Atlanta Public Schools system.⁹⁷ These educators include teachers, principals, and administrators—including the former Superintendent who may spend up to 45 years in prison for her role in the scandal. Will Philadelphia’s cheating scandal come to a similar resolution as Atlanta’s? Probably not. One reason is the scope of the investigations. While the then-Governor of Georgia set aside a budget “substantial enough to hire more than 50 state investigators,” the School District of Philadelphia, still under extreme financial stress, has hired a handful of lawyers pro bono to help with their load.⁹⁸ Furthermore, criminal charges cannot be sought against educators in Pennsylvania for cheating that occurred before 2012.

Once the State of Pennsylvania does release the findings of their investigation, however, one conclusion is inevitable and unsurprising: cheating occurred at an alarming rate in Philadelphia schools between 2009 and 2011. The erasure analysis data provides nearly irrefutable evidence that tests were tampered with. Schools that were flagged for cheating performed statistically significantly better in math and reading scores before 2012. But when the public caught wind of cheating and put an end to it in 2012, flagged schools and grade levels performed statistically significantly worse than their non-flagged counterparts—even when controlling for drastic budget cuts.

⁹⁷ Winerip, Michael. “Ex-Schools Chief in Atlanta Is Indicted in Testing Scandal”. *The New York Times*. Mar. 29, 2013.

⁹⁸ *Ibid.*

Without a doubt, other factors contributed to the massive test score declines experienced by Philadelphia schools in 2012. Governor Tom Corbett and his Republican-led Assembly crippled Philadelphia school resources by eliminating nearly 10% of their state funding. Even when controlling for a school's flagged status, school budget cuts had a statistically significant effect on school proficiency rates. Additionally, the evidence that non-flagged grades in flagged schools performed worse than non-flagged grades in non-flagged schools suggest that unexplained influences in flagged schools affected the proficiency rates across the entire school. Nevertheless, none of these findings should cloud the fact that Philadelphia has had a cheating problem, and that problem will not go away until those responsible leave.

Educators who participated in cheating may rightly claim that their actions were encouraged by high-stakes testing standards set by President Bush's No Child Left Behind act. The act, which has yet to be re-written by Congress during the Obama administration, labels schools as successful or failing based almost solely on test scores. In reaction to the strong influence of testing, states across the country have made their tests easier while districts and schools will stop at almost nothing to raise scores, leaving no school in the nation unaffected. Whatever the good intentions of the act were, the unintended consequences have caused a net negative effect on the nation's education system as a whole, especially in urban districts where schools with high poverty concentrations are asked to perform to high expectations without the necessary financial resources to do so.

This study shares a story about the neglect of low-performing students that starts on the national level and finishes in the classroom. Education is a business that is rife with politics that often provides incentives for educators to ignore the main purpose of education. Over the past decade, national policy has failed educators, and in turn educators have failed their students. As has been the case in urban education for decades, politics and the ambitions of self-minded educators has hindered instruction of the students with the highest need for a quality education.

Acknowledgements

This paper would not have been possible without the gracious help and support from advisors, friends, and family. Dr. Marc Meredith has been a research advisor of mine for two years. He assisted with the project from beginning to end, and his statistical expertise and guidance have been invaluable. I also wish to thank Penn's Political Science Department, including Dr. Doherty-Sil and Dr. DiIulio, for consistently inspiring, motivating, and supporting my academic efforts. Joe Tierney and the Fox Leadership staff have also helped fuel my passion for education since my freshman year at Penn. My internship supervisors at Public Citizens for Children and Youth—Shanee Garner—and the School District of Philadelphia—Dan Piotrowski—gave me wonderful opportunities to excel at their organizations and helped cement my interest in education policy. Finally, I want to thank my family and friends for consistently supporting me during the good times and especially during the more stressful ones.

Thank you.

Appendix

Table 10: Philadelphia Schools Flagged for Erasure Irregularities (Marked with Grade Levels and Subjects Flagged by School and Year)

Grade Levels	Schools	Year	3rd		4th		5th		6th		7th		8th		11th		Percentage Flagged
			M	R	M	R	M	R	M	R	M	R	M	R	M	R	
Bridesburg	09		Y	Y													33.3%
	10	Y	Y	Y													50.0%
	11			Y		Y											33.3%
Catharine	09			Y													16.7%
	10	Y		Y		Y											50.0%
	11	Y	Y	Y	Y												66.7%
Cayuga	09				Y	Y	Y										50.0%
	10	Y	Y		Y		Y										66.7%
	11	Y	Y	Y													50.0%
Childs	09					Y	Y	Y									25.0%
	10	Y		Y	Y	Y	Y							Y			50.0%
	11					Y		Y		Y							25.0%
Cramp	09	Y	Y	Y													37.5%
	10	Y	Y	Y		Y	Y	Y	Y								87.5%
	11	Y	Y	Y													37.5%
Day	09			Y	Y			Y				Y					33.3%
	10				Y			Y	Y	Y		Y					41.7%
	11	Y		Y	Y			Y	Y								41.7%
Decatur	09	Y								Y							16.7%
	10	Y	Y	Y		Y						Y					41.7%
	11	Y	Y	Y	Y			Y	Y								50.0%
Disston	09	Y		Y													16.7%
	10	Y	Y	Y	Y				Y								41.7%
	11	Y	Y	Y				Y	Y								41.7%
FS Edmonds	09	Y	Y	Y	Y	Y	Y	Y									87.5%
	10	Y	Y	Y	Y	Y	Y	Y	Y								100.0%
	11	Y	Y	Y	Y	Y	Y	Y	Y								100.0%
Emlen	09	Y	Y	Y	Y	Y	Y	Y	Y								100.0%
	10	Y	Y	Y	Y	Y	Y		Y								87.5%
	11	Y	Y	Y	Y	Y	Y	Y	Y	Y							100.0%
Forrest	09	Y	Y	Y	Y	Y	Y	Y	Y								100.0%
	10	Y	Y	Y	Y	Y	Y	Y	Y								100.0%
	11	Y	Y	Y	Y	Y	Y	Y	Y								100.0%
Fulton	09																0.0%
	10	Y	Y	Y				Y									50.0%
	11	Y				Y	Y	Y	Y								62.5%

Heston	09	Y	Y																	16.7%
	10							Y	Y	Y										25.0%
	11	Y	Y	Y	Y			Y		Y	Y									58.3%
Houston	09																			0.0%
	10					Y														8.3%
	11	Y	Y	Y		Y	Y				Y	Y								58.3%
Huey	09	Y	Y	Y	Y	Y	Y							Y						58.3%
	10	Y	Y	Y	Y	Y	Y	Y	Y					Y						75.0%
	11	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y						91.7%
Kearny	09	Y			Y															16.7%
	10	Y				Y	Y				Y									33.3%
	11					Y	Y	Y	Y	Y	Y	Y								50.0%
Kelly	09	Y	Y																	25.0%
	10	Y	Y	Y	Y															50.0%
	11	Y	Y	Y	Y															50.0%
Lamberton Elem.	09	Y		Y				Y												25.0%
	10			Y	Y									Y	Y					33.3%
	11	Y	Y	Y	Y	Y	Y	Y	Y					Y	Y					83.3%
Locke	09																			0.0%
	10	Y	Y					Y	Y											33.3%
	11			Y	Y	Y	Y	Y	Y	Y	Y	Y								66.7%
Loesche	09					Y														16.7%
	10																			0.0%
	11					Y	Y													33.3%
Marshall, J.	09	Y		Y	Y	Y														66.7%
	10	Y	Y	Y	Y															66.7%
	11	Y	Y		Y															50.0%
Marshall, T.	09	Y	Y	Y	Y			Y	Y	Y										58.3%
	10	Y	Y	Y									Y							33.3%
	11																			0.0%
McClure	09	Y	Y	Y	Y															100.0%
	10	Y	Y	Y	Y															100.0%
	11	Y	Y	Y	Y															100.0%
McDaniel	09	Y	Y			Y														25.0%
	10	Y	Y																	16.7%
	11	Y	Y	Y		Y		Y		Y	Y									58.3%
Mitchell	09	Y	Y																	33.3%
	10		Y				Y													33.3%
	11																			0.0%
Munoz-Marin	09	Y	Y	Y	Y	Y								Y						50.0%
	10			Y																8.3%
	11	Y	Y		Y									Y						33.3%
Olney El.	09	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100.0%

	11				Y	Y	Y	Y		100.0%
Tilden	09		Y	Y	Y	Y	Y	Y		100.0%
	10		Y	Y	Y	Y	Y	Y		100.0%
	11									0.0%
Wagner	09				Y	Y	Y	Y		66.7%
	10				Y	Y	Y	Y		66.7%
	11				Y			Y		33.3%
Bok Tech	09							Y	Y	100.0%
	10							Y	Y	100.0%
	11							Y		50.0%
Comm. Tech	09							Y	Y	100.0%
	10							Y	Y	100.0%
	11									0.0%
Frankford	09							Y		50.0%
	10							Y		50.0%
	11							Y		50.0%
Furness	09									0.0%
	10							Y		50.0%
	11							Y		50.0%
Kensington CAPA	09									0.0%
	10									0.0%
	11							Y		50.0%
King	09									0.0%
	10									0.0%
	11							Y		50.0%
Northeast	09							Y		50.0%
	10							Y	Y	100.0%
	11							Y	Y	100.0%
Phila. Military Academy @ Elverson	09							Y		50.0%
	10									0.0%
	11							Y		50.0%
Randolph	09									0.0%
	10									0.0%
	11								Y	50.0%
Strawberry Mansion	09							Y	Y	100.0%
	10							Y	Y	100.0%
	11							Y		50.0%

Bibliography

- Amrein-Beardsley, A.; Berliner, D.C.; Rideau, S.. "Cheating in the first, second, and third degree: Educators' responses to high-stakes testing". *Education Policy Analysis Archives*. Vol 18, No. 14. 2010.
- Anderman, Eric, & Murdock, Tamera. *Psychology of Academic Cheating*. Elsevier Academic Press. 2007. Pg. 292.
- Basham, Nick. "The Five-Year Legacy of Paul Vallas in Philadelphia". *The Philadelphia Inquirer*. June 2, 2007.
- Berliner, David. "Poverty and Potential: Out-of-School Factors and School Success". *The Great Lakes Center for Education Research & Practice*. March 2009.
- Butterfield, Sam. "Spend rainy-day fund, advocates tell governor". *Pittsburgh Post-Gazette*. June 17, 2011. <http://www.post-gazette.com/stories/local/breaking/spend-rainy-day-fund-advocates-tell-governor-302485/>.
- Chudowsky, Naomi; Chudowsky, Victor. "Many States Have Take a 'Backloaded' Approach to No Child Left Behind Goal of All Students Scoring 'Proficient'". *Center on Education Policy*. May 19, 2008.
- Cizek, Gregory J. "Cheating on Tests and Other Dumb Ideas". *Education Week*. July 25, 2011. <http://www.edweek.org/ew/articles/2011/07/25/37cizek.h30.html>.
- "Costing Out the Resources Needed to Meet Pennsylvania's Public Education Goals". *Augenblick, Palaich and Associates, Inc*. December, 2007. Pg. ii.
- Dee, Thomas S., and Brian A. Jacob. "The Impact of No Child Left Behind on Students, Teachers, and Schools." *Brookings Papers on Economic Activity* 2010.2 (2010): 149-207.
- Editorial: Applaud, but lightly". *The Philadelphia Inquirer*. June 19, 2010. http://articles.philly.com/2010-06-19/news/24962627_1_math-goals-school-aid-test-scores#axzz0rUrHFPX5.
- Fifield, Adam. "Power: Reformer R.I.P.". *Philadelphia Magazine*. September 2007. <http://www.phillymag.com/articles/power-reformer-rip/>.
- Gillum, Jack and Bello, Marisol. "When standardized test scores soared in D.C., were the gains?" *USA Today*. Mar. 30, 2011. http://usatoday30.usatoday.com/news/education/2011-03-28-1Aschooltesting28_CV_N.htm.

- Goldhaber, Dan. "The Mystery of Good Teaching". *Education Next*. Spring 2002. Vol. 2, No. 1.
- Goldstein, Andrew. "Bush and Gore: Who's The Education President?". *TIME Magazine*. Nov. 6, 2000.
<http://www.time.com/time/magazine/article/0,9171,998406-1,00.html>.
- Graham, Kristen. "26 Philadelphia schools losing full-service kitchens". *The Philadelphia Inquirer*. June 30, 2011. http://articles.philly.com/2011-06-30/news/29722033_1_school-meals-kitchens-elementary-schools.
- Graham, Kristen. "A test-score milestone for Phila. schools". *The Philadelphia Inquirer*. June 17, 2010. http://articles.philly.com/2010-06-17/news/24966078_1_empowerment-schools-math-on-grade-level-school-reform-commission#axzz0r7R5liGi.
- Graham, Kristen. "Education chief orders reviews of state exams since '09". *The Philadelphia Inquirer*. July 14, 2011. http://articles.philly.com/2011-07-14/news/29773644_1_pssa-cheating-on-state-tests-education-secretary-ronald-tomalis.
- Graham, Kristen. "Pa. education secretary orders reviews of all state exams since 2009, especially Phila.'s". *The Philadelphia Inquirer*. July 15, 2011. http://articles.philly.com/2011-07-15/news/29776859_1_pssa-education-secretary-ronald-tomalis-analysis-of-testing-data.
- Graham, Kristen. "Phila. District rejects claims of cheating at 2 schools". *The Philadelphia Inquirer*. July 7, 2011. http://articles.philly.com/2011-07-07/news/29747354_1_pssa-cheating-teachers.
- Graham, Kristen. "Philadelphia schools celebrate successful scores". *The Philadelphia Inquirer*. Aug. 21, 2010. http://articles.philly.com/2010-08-21/news/24974100_1_special-education-pupils-persistently-dangerous-school-empowerment-schools.
- Graham, Kristen. "Vallas on Ackerman". *The Philadelphia Inquirer: Philly School Files*. Sep. 6, 2011.
http://www.philly.com/philly/blogs/school_files/Vallas-on-Ackerman.html.
- Graham, Kristen; Purcell, Dylan. "City school's fast rising test scores questioned". *The Philadelphia Inquirer*. May 1, 2011.
http://articles.philly.com/2011-05-01/news/29493382_1_pssa-results-allegations-of-testing-irregularities-roosevelt-teachers.

- Graham, Kristen; Purcell, Dylan. "‘Compelling’ evidence of cheating in many Phila. Schools". *The Philadelphia Inquirer*. Mar. 11, 2012.
http://articles.philly.com/2012-03-11/news/31145463_1_city-charter-schools-vanguard-schools-education-secretary-ronald-tomalis.
- Graham, Kristen; Purcell, Dylan. "PSSA-cheating reforms yield lower scores across Pa.". *The Philadelphia Inquirer*. July 30, 2012.
http://articles.philly.com/2012-07-30/news/32924357_1_erasure-patterns-education-ronald-tomalis-pssa-results.
- Hanushek, Eric A.; Kain, John F.; Rivkin, Steven G.. "Teachers, Schools, and Academic Achievement". *National Bureau of Economic Research*. Aug. 1998
- Hardy, Dan. "Corbett’s Pennsylvania budget holds the line on school spending for the most part". *The Philadelphia Inquirer*. Feb. 8, 2012.
http://articles.philly.com/2012-02-08/news/31038123_1_school-districts-funding-million-gap.
- Hardy, Dan. "Pa. budget cuts hit poorer districts hardest, funding equity advocate says". *The Philadelphia Inquirer*. July 11, 2011.
http://articles.philly.com/2011-07-11/news/29761528_1_school-districts-funding-formula-education-law-center.
- Herold, Benjamin. "110 schools meet AYP targets" *The Philadelphia Public School Notebook*. Aug. 18, 2011. <http://thenotebook.org/blog/113967/110-schools-meet-ayp-targets>.
- Herold, Benjamin. "New cheating concerns at Philly school met with ‘baffling’ response". *The Philadelphia Public School Notebook*. Oct 23, 2012.
<http://thenotebook.org/blog/125270/wagner-cheating-follow>.
- Herold, Benjamin. "One-third of District schools had pattern of suspicious erasures". *The Philadelphia Public School Notebook*. July 11, 2011.
<http://thenotebook.org/blog/113873/one-third-district-schools-had-suspicious-erasures>.
- Herold, Benjamin. "Pa. wants 90 schools investigated for 2009 cheating". *The Philadelphia Public School Notebook*. July 13, 2011.
<http://thenotebook.org/blog/113881/pa-wants-90-schools-investigated-2009-cheating>.
- Herold, Benjamin. "State cheating probe was widened to involve more than 50 District schools". *The Philadelphia Public School Notebook*. Feb. 16,

2012. <http://thenotebook.org/blog/124532/state-cheating-probe-widened-involve-more-50-district-schools>.

Herold, Benjamin; Mezzacappa, Dale. "2009 report identified dozens of PA schools for possible cheating". *The Philadelphia Public School Notebook*. July 8, 2011. <http://thenotebook.org/blog/113871/2009-report-identified-pa-schools-possible-cheating#list>.

Herold, Benjamin; Mezzacappa, Dale. "District promoted principals from schools under suspicion for cheating." *The Philadelphia Public School Notebook*. Sep. 18, 2012. <http://thenotebook.org/blog/125145/district-promoted-principals-schools-under-suspicion-cheating>.

Herold, Benjamin; Mezzacappa, Dale. "New evidence: Philadelphia test cheating is likely far worse than previously revealed". *The Philadelphia Public School Notebook*. Aug. 16, 2012. <http://thenotebook.org/blog/125073/evidence-shows-state-test-cheating-philadelphia-likely-far-worse-previously-revealed>.

"How bad is Texas, really?" *The Economist*. Oct. 19th, 2000. <http://www.economist.com/node/397540>.

Jensen, Tom. "Corbett in dire shape for reelection". *Public Policy Polling*. Mar. 12, 2013. <http://www.publicpolicypolling.com/main/2013/03/corbett-in-dire-shape-for-reelection.html>.

Kerkstra, Patrick. "Nutter, Ackerman Split". *The Philly Post*. June 7, 2011. http://blogs.phillymag.com/the_philly_post/2011/06/07/nutter-ackerman-split/.

King, Larry. "Fairly or not, spikes in test scores haunt some schools". *The Philadelphia Inquirer*. July 16, 2011. http://articles.philly.com/2011-07-16/news/29780910_1_pssa-results-charter-schools-proficient-score.

Lecher, Colin. "What Are Your Odds Of Winning The Lottery? [Infographic]". *Popular Science*. Nov. 29, 2012. <http://www.popsci.com/science/article/2012-11/dismal-odds-winning-lottery-infographic>.

Markosek, Joe. "Revenue Surplus Exceeds \$785 million". *PA House Appropriations Committee: Fast Facts*. July 1, 2011. http://www.pahouse.com/HACD/series/16/FF_Revenue%20Est_070111.pdf

- Masch, Michael. "SDP Budget Update: Changes Since Initial Adoption on May 31, 2011". *School District of Philadelphia*. August 3, 2011. Slide 13.
- Mezzacappa, Dale. "Budget 'ominous,' further cuts required; District to hire recovery officer". *The Philadelphia Public School Notebook*. Jan 19, 2012. <http://thenotebook.org/blog/124427/budget-ominous-further-cuts-required-district-hires-recovery-officer>.
- Mezzacappa, Dale. "Hite: New plan will result in fewer students moving to lower-performing schools". *The Philadelphia Public School Notebook*. Feb. 19, 2013. <http://thenotebook.org/blog/135657/hite-new-plan-will-result-fewer-students-going-lower-performing-schools>.
- Mezzacappa, Dale. "Just in time for the holidays: more layoffs". *The Philadelphia Public School Notebook*. Dec. 22, 2011. <http://thenotebook.org/blog/114382/just-time-holidays-more-layoffs>.
- Mezzacappa, Dale. "State prohibits Philly teachers from administering PSSA to their own students". *The Philadelphia Public School Notebook*. Feb. 28, 2012. <http://thenotebook.org/blog/124560/new-rule-teachers-cannot-administer-pssa-tests-their-own-students>.
- Nichols, Sharon L. and Berliner, David C.. *Collateral Damage: How High-Stakes Testing Corrupts America's Schools*. Harvard Education Press. 2008. Pg. 34.
- Niederberger, Mary. "No Child Left Behind waiver to be sought by Pennsylvania". *Pittsburgh Post-Gazette*. Nov. 28, 2012. <http://www.post-gazette.com/stories/news/education/no-child-left-behind-waiver-to-be-sought-by-pennsylvania-663943/>.
- Olsen, Laura; Mauriello, Tracie. "Corbett's state budget plan cuts \$866 million". *Pittsburgh Post-Gazette*. Mar. 9, 2011. <http://www.post-gazette.com/stories/local/state/corbetts-state-budget-plan-cuts-866-million-211571/>.
- Owens, Erika. "Ackerman's tenure in Philadelphia: 2008-11 key events". *The Philadelphia Public School Notebook*. October 2011. Vol. 19. No. 2.
- "Philadelphia 2011: The State of the City". *The Pew Charitable Trusts*. Apr. 9, 2011.
- "Philadelphia Superintendent Wins Nation's Top Prize". *Council of the Great City Schools*. Oct. 21, 2010.

- Ravitch, Diane. "The Myth of Charter Schools". *The New York Review of Books*. Nov. 11, 2010.
- Resmovits, Joy. "Atlanta Public Schools Shaken By Cheating Report". *The Huffington Post*. Sep. 4, 2011.
http://www.huffingtonpost.com/2011/07/05/atlanta-public-schools-cheating_n_890526.html.
- Rosenbaum, David. "The 2000 Campaign: The Ad Campaign; Bush on Education". *The New York Times*. March, 18, 2000.
<http://www.nytimes.com/2000/03/18/us/the-2000-campaign-the-ad-campaign-bush-on-education.html>.
- Samuels, Christina. "Cheating Scandals Intensify Focus on Test Presures". *Education Week*. Aug. 4, 2011.
http://www.edweek.org/ew/articles/2011/08/04/37cheating_ep.h30.html.
- Scott, Katherine. "3,800 Philadelphia school personnel face layoffs". *6ABC.com*. June 6, 2011.
<http://abclocal.go.com/wpvi/story?section=news/local&id=8172513>.
- Snyder, Susan. "Vallas facing sharp criticism". *The Philadelphia Inquirer*. Nov. 11, 2006. http://articles.philly.com/2006-11-11/news/25407400_1_paul-vallas-deficit-cuts.
- Socular, Paul. "Cheating coverage honored". *The Philadelphia Public School Notebook*. April 2012. Vol 19. No. 5.
- Socular, Paul. "District statement: We never got 2009 report". *The Philadelphia Public School Notebook*. July 9, 2011.
<http://thenotebook.org/blog/113872/district-we-never-got-2009-report>.
- Socular, Paul. "Imagine 2014 gets a green light from the SRC". *The Philadelphia Public School Notebook*. Summer 2009. Vol. 16. No. 4.
- "Suspicious erasure patterns at 53 schools". *The Philadelphia Public School Notebook*. Aug. 16, 2012.
<http://thenotebook.org/blog/125074/suspicious-erasure-patterns>.
- "Technical Report for the 2012 Pennsylvania System of School Assessment". *Data Recognition Corporation*. 2012.

“Technical Report for the Pennsylvania System of School Assessment: 2009 Reading and Mathematics Grades 3, 4, 5, 6, 7, 8, and 11”. *Data Recognition Corporation*. July 2009.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1994, 1998, 2002, 2003, 2005, 2007, 2009 and 2011 Reading and Mathematics Assessments.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2009 Reading and Mathematics Assessments. SY 2008-09 Consolidated State Performance Report Part I

U.S. Department of Education: Mapping State Proficiency Standards Onto the NAEP Scales: Variation and Change in State Standards for Reading and Mathematics, 2005-2009. National Center for Education Statistics. 2011. pg. 14.
<http://nces.ed.gov/nationsreportcard/pdf/studies/2011458.pdf>.

Venkataramanan, Rajiv; Mezzacappa, Dale. “How state funding has changed since the costing-out study”. *The Philadelphia Public School Notebook*. October 2010. Vol. 18. No. 2.

White, Bill. “Pennsylvania Gov. Tom Corbett bombing public education into Stone Age”. *The Morning Call*. Sept. 26, 2012.
http://articles.mcall.com/2012-09-26/news/mc-bw-pennsylvania-gov-tom-corbett-bombing-public--20120926_1_tom-corbett-public-education-charter-school.

White, Jeremy. “Education Reform: As Obama Pushes for Teacher Evaluations, Little Agreement on How.” *International Business Times*. Oct. 28, 2011.
<http://www.ibtimes.com/education-reform-obama-pushes-teacher-evaluations-little-agreement-how-362690#>.

“Who runs the School District of Philadelphia?”. *Education Voters of Pennsylvania*. 2013.
<http://www.educationvoterspa.org/index.php/site/issues/who-runs-the-school-district-of-philadelphia/>.

Winerip, Michael. “Ex-Schools Chief in Atlanta Is Indicted in Testing Scandal”. *The New York Times*. Mar. 29, 2013.
<http://www.nytimes.com/2013/03/30/us/former-school-chief-in-atlanta-indicted-in-cheating-scandal.html?pagewanted=1&r=0>.

Worden, Amy. "PA budget highlights: On time, slimmed down". *The Philadelphia Inquirer: Commonwealth Confidential*. June 30, 2012.
http://www.philly.com/philly/blogs/harrisburg_politics/Budget-highlights.html.