

A Comparative Study of the Impacts of High Stakes Standardized Testing Between Teachers of
Core Content Areas and Teachers of Elective Courses.

Chris Shinski

SUNY Plattsburgh

A Comparative Study of the Impacts of High Stakes Standardized Testing Between Teachers of Core Content Areas and Teachers of Elective Courses.

Abstract: The rise of high stakes standardized testing is ubiquitous throughout secondary schools in the United States. Prior studies have revealed several unintended side effects of high stakes testing that have had significant influence on teaching strategies, teacher morale, student achievement, and curriculum design. The purpose of this qualitative study is to examine the existence of unintended side effects associated with New York's inclusion of standardized tests in secondary schools and compare the effects between classes with high stakes exams and classes that do not have a mandated high stakes exam. A survey of 22 questions was distributed to a total of 12 teachers at the secondary level with half teaching content area in the core areas of math, English, science and social studies. Findings suggest that stress levels are higher, more pressure exists, and instructors are more likely to teach directly to the test when a high stakes exam is involved. Teachers in classes without a high stakes exam repeatedly felt they had more freedom to teach the way they wanted to but were indirectly affected by high stakes exams. The research concludes with recommendations on how findings can be used for future purposes and what questions may be considered for further evaluation.

Keywords: High stakes, tests, curriculum, morale, stress, Regents

A COMPARATIVE STUDY OF HIGH STAKES TESTING

A Comparative Study of the Impacts of High Stakes Standardized Testing Between Teachers of Core Content Areas and Teachers of Elective Courses.

The iconic no. 2 pencil has survived decades in the classroom even amidst the modern technologies found in American school buildings of the 21st century. The painted yellow, pink-tipped hexagon filled with graphite is a simple tool over 85% of students in the United States will use on a state-mandated, high stakes test throughout the school year and serves as constant reminder of America's obsession with assessment (Fletcher, 2009).

In 1966, sociologist James Coleman led a report titled *Equality for Education* that raised awareness and controversy amongst Americans at all levels involved in education. Arguments over whether or not the amount of resources allocated by a school affected performance pushed policymakers to consider great changes. Coleman stressed the correlation between socioeconomic status and student achievement and argued standardization was the only effective means to judge one student against another to decide where funding was needed (Coleman, 1966).

Within the next two decades *A Nation at Risk* was published in 1983 by the National Commission of Excellence in Education (Barksdale-Ladd & Thomas, 2000). The report revealed American achievement was faltering against international competition and urged schools to improve by issuing higher standards and accepting more accountability. Pedulla et al. found that “standards-based reform efforts were designed to encourage schools, teachers and students to excel and meet tougher academic challenges as prescribed by state curricular standards and frameworks” (2003, p. 11).

Under the auspice of improving education, standardization had successfully infiltrated the arena of assessment by the 1990's and George Bush's claim of the Texas Miracle had set up the

A COMPARATIVE STUDY OF HIGH STAKES TESTING

country to fall into standardized assessments with claims of successfully curtailing the decline of student achievement (Haney, 2001). By the time the Texas Miracle was debunked (Haney, 2001; Stanford, 2013) different forms of standardized testing had already become a component of state-level assessments asked for by the 2001 act, No Child Left Behind (NCLB).

The main goal of NCLB reads, “An act to close the achievement gap with accountability, flexibility, and choice, so that no child is left behind” (“Public Law 107-110”, 2002, p. 1). Under this mask the act broadened the federal governments reach into state level education systems by demanding annual testing, proof of academic progress, detailed report cards, early reading programs, and higher teacher qualifications in order to receive federal support.

Other programs like Obama’s Race to the Top (RTTT) initiative, part of the 2009 American Recovery and Reinvestment Act, have also raised the stakes by incentivizing schools to use high stakes tests as a way to measure achievement in order to obtain millions of dollars in funding. Currently, all 50 states utilize some form of standardized testing amounting to over \$1.7 billion per year (Chingos, 2012). With the adoption of Common Core Standards in more than 45 states, the rigor and stakes associated to the mandated assessments will undoubtedly increase and have strong effects on school culture and student achievement in the forthcoming years.

PURPOSE STATEMENT

The American Educational Research Association uses the term “high stakes” when tests “carry serious consequences for students or educators” (Marchant, 2004, p. 2).

New York State’s adoption of the Common Core and subsequent participation in RTTT has resulted in more rigorous standards assessed through mandated high stakes testing. The New York State Regents’ exams directly impact RTTT and teachers’ Annual Professional Performance Review (APPR). Although Regents’ exams have been issued for over 100 years,

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Grant (2000, p. 7) notes that in the last decade students “[had] to pass five Regents examinations (English, mathematics, global studies, U.S. history, and science) in order to graduate.”

New York State’s exemption of state level testing in non-core content related classes such as agriculture, music, and art raises questions of how instructors of different content areas view and are affected by the requirements and growing pressures of mandated assessments. This study aims to reveal any differences in the effects of high stakes testing between core content areas (English, math, science, social studies) with state level, mandated assessments and elective classes with local assessments in New York State.

PRIOR STUDIES

Summary

Prior studies have used qualitative and quantitative methods to investigate various state mandated testing programs throughout the United States. While the idea of added rigor in assessment and holding education to higher standards is widely supported (Barksdale-Ladd & Thomas, 2000; Grant, 2000), numerous studies conducted in states with mandated, high stakes tests have found evidence of side effects including lower student achievement, lower teacher morale, curriculum narrowing, and higher stress levels amongst instructors.

Increased Pressure and Effects on Teacher Efficacy

In 1991, Shepard and Dougherty’s study of two high stakes testing districts revealed that more than half of the faculty felt “great pressure from the district administration or board of education to raise scores” (p. 5). Only 8% felt that pressure was slight or non-existent. In the same study, 66% of respondents felt a significant amount of pressure from the media and publishing of results to raise scores.

A COMPARATIVE STUDY OF HIGH STAKES TESTING

According to Powell, Higgins, Aram, and Freed, teachers in Maine are often “discouraged and want out of teaching” and one participant stated “low performing schools lose experienced teachers to early retirement or to private schools” (Powell et al., 2009, p. 25). The pressure to have students perform well on high stakes tests thanks to NCLB in combination with curriculum narrowing has also pushed teachers “to adopt teaching methods that lack innovation and creativity, often leaving teachers to feel that they are unable to use their professional judgment about what’s best for their students” (Powell et al., 2009, p. 27).

Descriptions from Barksdale-Ladd and Thomas’ (2000) study indicated that pressure and stress are constant when it comes to making sure students do well on high stakes testing. They continue to identify memos from administrators, teacher meetings, and media as sources of the “internalized pressure [that is] always with me” (Barksdale-Ladd & Thomas, 2000, p. 390). One middle school teacher in a study by Moon, Bright, Jarvis, and Hall (2007) said their principal, school board, and superintendent “want test results to look good in the local paper” and another high school teacher said “testing pressure has permeated almost every meeting and discussion, making them more stressful” (p. 81). In the same report, data from both the national survey and school visits conducted showed pressure to evolve from the top down.

A participant in Barksdale-Ladd and Thomas’ (2000) study felt high stakes testing had made them a worse teacher. “I am not the teacher I used to be”, they reported, “I used to be great... All of my powerful teaching tools I used to use every day are no good to me now because they don’t help children get ready for the test, and it makes me feel like a robot instead of a teacher” (Barksdale-Ladd & Thomas, 2000, p. 392). Pedulla et al. (2003) revealed roughly 70% of teachers are practicing methods that contradict their own views of sound teaching because of the pressures associated with high stakes tests. 85% of high school teachers and 93%

A COMPARATIVE STUDY OF HIGH STAKES TESTING

of elementary teachers reported feeling pressure from the superintendent to raise scores in the same study.

Research by Jones, Jones, Hardin, Chapman, Yarbrough, and Davis (1999) and Smith (1999) identify a negative relationship between high stakes testing and teacher morale. North Carolina's former state superintendent noted "With the help of a little public scrutiny, I think you'll be amazed at how much better those schools are going to get" (Jones et al., 1999, p. 200). With high stakes testing, it is becoming more commonplace for results to be published to the public with teachers often taking the brunt of criticism when achievement is not proficient or exemplary (Jones et al. 1999; Smith, 1991). The same study tells of teachers appreciating the increased accountability but when the test does not accurately reflect their work or students' abilities, teachers are citing feelings of anxiety, guilt and embarrassment. Barksdale-Ladd and Thomas (2000) found high stakes testing to result in more citations of lower self-esteem, shame, and alienation amongst teachers in content areas with high stakes tests. In addition, Jones et al. (1999) found 77% of teachers reported lower morale and 76% said their jobs were more stressful due to high stakes testing.

Hamilton, L., Stecher, B., Marsh, J., McCombs, J., Robyn, A., Russell, J, Naftel, S. and Barney, H. (2007) found teachers in three states to believe the rigors of state assessments were too difficult for their students. More than 70% of all teachers in California, Georgia, and Pennsylvania reportedly felt more pressure to raise student achievement to meet the requirements of the Annual Yearly Progress outlined in No Child Left Behind (Hamilton et al., 2007).

Teachers that have students with higher achievement are not exempt from anxiety and stress (Smith, 1999). Smith's study continued by stating administrators in high achieving districts continue to apply pressure on teachers to keep scores higher or exceed the previous year's

A COMPARATIVE STUDY OF HIGH STAKES TESTING

growth leaving many teachers feeling frustrated by goals they viewed as “technically impossible to meet” (Smith, 1999, p. 9).

Effects on Curriculum

High stakes testing has had impacts on both curriculum content and modes of instruction. In 1967 a study by Goslin (Shepard & Dougherty, 1991) showed that teachers rarely altered their instruction after reviewing the test results of their students. In the 1970's, a renewed focus on assessment promoted many changes and surveys as early as the 1980's showed that instruction was beginning to reflect content being tested on high stakes exams. In 1989, Romberg, Zarinnia, and Williams (Shepard & Dougherty, 1991) found that more than one in four mathematics teachers were changing lessons plans and instructional methods to cover topics emphasized on tests. Early studies reviewed by Herman and Golan (1991) differentiated between the effects of mandated assessments and high stakes testing. Their study found that when promotions, job security, or performance reviews were directly related to student achievement on mandated tests, teachers were more likely to adapt their curriculum and dedicate up to four weeks to prepare for one test (Herman & Golan, 1991).

Urduan and Paris (1994, p. 145) revealed 81% of surveyed teachers “believed that other teachers teach to the test.” Congruent with those teachers' beliefs, Barksdale-Ladd and Thomas (2000, p. 391) found that “75% of the teacher surveyed in their study had changed their instructional practices due to testing.” Teachers felt that they had discontinued more effective means in lieu of test preparation (Barksdale-Ladd & Thomas, 2000).

Pedulla et al. (2003) surveyed over 4,100 teachers from every state except Iowa, Oregon, and Idaho, and found that teachers in schools with high stakes exams were found to align their curriculum more to the content in mandated assessments than those in lower stakes groups where

A COMPARATIVE STUDY OF HIGH STAKES TESTING

test results had fewer consequences for students and teachers less. Their study showed more than half of teachers align their curriculum to state frameworks that are similar to previous tests.

A study of literacy related instructors conducted by White, Sturtevant, and Dunlap (2003) found that Virginia's Standards of Learning and state exams influenced 60% of interviewees decisions related to their instruction. One teacher said she "felt the need to push push push the kids all the time to their maximum limit" (White, Sturtevant, & Dunlap, 2003, p. 55).

Mesler's (2008) study found that some teachers focus on materials covered in the test. "If it's not tested, it's not going to be taught," was the response of one interviewee (Mesler, 2008, p. 24). Abrams, Pedulla, and Madaus (2003, p. 23) found teachers in high stakes areas had "significant decreases in time spend on instruction in fine arts, industrial/vocational education, field trips, class trips, enrichment assemblies, and class enrichment activities." The study also researched the content focus of teachers in high stakes areas and results were similar to Mesler's (2008) revealing that 43% of teachers in high stakes areas greatly increased time spent on tested areas (Abrams, Pedulla, & Madaus, 2003). Furthermore, 51% of teachers created classroom tests that reflected state exams and more than 30 hours were dedicated to these assessments by 44% of respondents.

In one study of two teachers, each of whom taught classes with state mandated high stakes testing and classes deemed electives, it was revealed that there was a large discrepancy between their curriculum planning based on the existence of the assessments. Gerwin (2006) found the teachers were unable to use creative methods or discuss current events through teachable moments because of the pressure to get students to pass tests when they existed. In the classes identified as electives, the same teachers were able to "discuss a handout in order to think about higher level concepts, evaluate historical events, and relate past events to current events"

A COMPARATIVE STUDY OF HIGH STAKES TESTING

(Gerwin, 2006, p. 270). Without state guidelines and a high-stakes test, one teacher's elective class left him in a position to "decide what content would be taught, how he would teach it, and how student learning would be assessed" (Gerwin, 2006, p. 266).

In the classes with Regents' exams, a focus on breadth over depth occurred and rote learning was part of daily instruction (Gerwin, 2006). The same study found the teachers struggling to find time for discussion and analysis in Regents' classes; the main goal was "to get them through Global History by equipping them with enough content and skills to pass the state exam" (Gerwin, 2006, p. 268).

Effects on Elective Courses

Little research has been done to study the effects of high stakes tests on elective courses. Few studies have examined teachers that teach both classes with high stakes exams and classes without them or teachers that have course loads of entirely elective classes without state mandated, high stakes exams.

The term "elective class" will be used in this study to define classes that earn credit hours towards graduation but are options amongst other classes and require no state mandated high stakes exam. In lieu of standardized exams, these classes have local assessments created by the instructor.

Gerwin and Visone's (2006) work showed teachers felt more freedom and flexibility within their elective classes to teach with methods they felt were stronger and promoted higher order thinking. In the elective courses, the participants were more concerned with student understanding as well as with more complex and even affective goals. In one teacher's lesson log, it was determined the instructor "was more likely to write that he had a good class in the elective than the 8th grade [class with a high stakes test]" (Gerwin & Visone, 2006, p. 270).

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Another teacher felt the same way and never reported having a great lesson or fielding interesting questions. The participant would later conclude pressure from administration to cover material on the Regents' exam led him to simply rule out any possibility of incorporating larger discussions. "The Regents, no matter what anyone wants to say, hangs over everyone's head. A teacher would be a fool not to recognize them" (Gerwin & Visone, 2006, p. 273).

Sabol's (2010) study found several educators felt "marginalized and devalued by colleagues, students, and school administration" after passage of NCLB and a renewed focus on math and science (p.141). More than half of study respondents reported spending more time developing assessments in art, explaining results to parents, and using various assessments more frequently. One in ten teachers also reported the additional focus on assessment took "the joy out of art" (Sabol, 2010, p. 133).

RESEARCH STUDY

Research Problem

Recent evidence indicates high stakes testing changes pedagogical strategies putting teacher efficacy at risk while raising stress levels in core content areas. By studying the effects of high stakes tests on teachers in elective classes and comparing their pedagogical strategies, levels of efficacy, and levels of stress to teachers within high stakes content areas, results may point towards common causalities that need to be addressed. Teachers that have experience in sections with and without a high stakes may exemplify different approaches to their curriculum based on the existence or absence of the test. Through data analysis in qualitative interviews this study aims to reveal any indications of differences in pedagogy between the two areas.

Target Population

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Ethical approval was granted by the SUNY Plattsburgh Committee on the Protection of Human Subjects prior to the beginning of the study. Written informed consent was obtained from each district's building administrator and each study participant before interviews were conducted.

A total of 12 secondary teachers in three districts were interviewed. Within each district, two teachers were interviewed from a core content area (math, science, ELA, or social studies) and two teachers were from an elective content (art, music, foreign language, health, etc.) area. Participants' experience ranged from 10 years of teaching to 35 with a mean number of 18 years in the classroom. Of the participants, 6 male teachers and 6 female teachers were selected to be interviewed.

Each district participating in the study was located in upstate New York and had a graduating class size of more than 50 and less than 150. Table 1 shows the schools were not in areas of extreme poverty as each was well below the national average of free and reduced lunch programs. Due to the size of each district, four participants taught sections in their content area that had both a high stakes test and lacked a high stakes test. The comparative analysis of their classroom experiences added validity to the study.

Instrumentation

A 22 question survey was developed and categorized into four main areas of study; test effects on stress level, test effects on morale, test effects on curriculum instruction, and test effects on teacher pressure as shown in the Appendix. The questions were finalized after a series of pilots were conducted to increase the validity of the survey and eliminate unnecessary or overlapping questions.

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Each question in the study was asked in an open ended manner and provided the participant the opportunity to express themselves freely and to a depth of their personal preference.

Methods

The principal investigator interviewed each participant in a series of one on one interviews over a two month period in a space within the interviewees home district. Each participant was made aware that the interview would be recorded, kept anonymous, transcribed, and that they had the option to skip any question they were uncomfortable or unable to answer. All transcriptions were electronically communicated back to the participant for review and confirmation of the results were received before publishing any data in the study.

Throughout the interviews the investigator could probe in depth for examples to general statements or heavily laden judgments to allow for richer data to be collected while notating any significant mannerisms or change in the tone of the subject's voice.

Limitations

The small sample size may present a limitation to the research of this study. In the process of selecting participants it was unknown what class sections each teacher taught in their respective school district and whether or not their sections included high stakes tests. The only consideration given prior to their selection was their content area (core or non-core). Participants taught either classes with high stakes tests exclusively, classes without high stakes test exclusively, or at least one section with a high stakes test and one without.

The types of high stakes tests included in this study are New York State Regents' exams, Advanced Placement (AP) exams, or other standardized exams that are mandatory for the class section in question. Content area classes absent of a standardized test that have a local

A COMPARATIVE STUDY OF HIGH STAKES TESTING

assessment created by the teacher are referred to as “elective” classes for the purpose of this study.

RESULTS

Overview

The data collected through qualitative analysis presented several themes that are concurrent with the findings of previously published literature. Teachers in the study with and without high stakes tests revealed they were negatively impacted by testing and cited few, if any, positive changes to their pedagogy or school culture due to high stakes tests.

Teachers with high stakes tests expressed deeper concern over lower morale amongst colleagues, the lack of creative instruction in classrooms, teaching to a test, and additional stress and pressure. Their colleagues without high stakes tests believed testing has contributed to marginalization, neglect, and the establishment of a hierarchy in schools where electives are placed at the bottom. One teacher in a high stakes area said they were “marching to the beat of a different drummer” and compared their pressure to teachers that were able to create their own classes and assessments.

Teachers that had sections in each category compared their pedagogical strategies, stress levels, and morale. The data suggests that there is a discrepancy between the way instructors feel they are able to present content to students, assess students, and take time to build understanding and ultimately indicated they preferred sections without high stakes tests.

Throughout the interview process information and quotations were collected for analysis and then populated into common themes as seen in Table 2. The data suggests teachers are supportive of measures to increase accountability in schools but are negatively impacted by the connection between their professional review and student performance on high stakes tests.

Effects on Pressure and Stress Levels

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Teachers in this study with sections including high stakes exams received regular pressure from their building leaders in regards to achievement levels on standardized tests. Nearly all teachers with a high stakes test said they had frequent and regular conversations in faculty meetings where scores were emphasized and most stated that high stakes test scores were mentioned at every staff meeting or very often.

While high stakes tests are often single handedly held accountable for the rise in stress levels, responses in this study suggests that stress is often associated to the way data from student performance impacts teachers' APPR ratings. Teachers in the high stakes area felt that absenteeism, a lack of intrinsic student motivation, and a lack of support at students' homes were all variables outside their control that increased stress because they negatively impact student scores that are now tied directly to their own evaluations. Multiple teachers believed that the quality of their instruction may not be accurately represented by unreliable data collected from student performance on high stakes tests.

Participants in this study with high stakes tests all revealed they had achieved or surpassed their building administrator's expectations for standardized test results and noted that they perceived less pressure than colleagues that had failed to meet goals established by building leaders. Figure 2 depicts the positive relationship between student achievement and the amount of pressure perceived by participants in their content areas from building administrators. One teacher summed up their experience with the publication of student data from high stakes tests by stating "it's not a direct pressure... it's an indirect understanding that If you're not performing, your number will be out there for everyone to see." Another teacher who had previously underperformed said they received less pressure from building leaders once their

A COMPARATIVE STUDY OF HIGH STAKES TESTING

students' scores met expectations but remembered being spoken to on a frequent bases when they had failed to meet expectations.

One participant with more than 20 years of classroom experience began to experience sleepwalking episodes that they directly attributed to stress caused by upcoming high stakes exams. Another teacher reported hearing of colleagues becoming nauseous and physically impacted by stress levels on the day tests were administered.

Several teachers with high stakes tests believed their students were negatively impacted by high stakes exams and at times this caused them to bring stress home while no teachers that created their own assessments indicated they were affected negatively outside of school.

In class sections without high stakes tests teachers revealed they felt little if any pressure from building administrators. Many of these teachers stated they were rarely, if ever, spoken to about test scores and had fewer meetings with principals on an individual or department level. One teacher with more than 30 years of experience and no high stakes tests said they never had a building leader mention anything about assessments in a faculty meeting to them throughout their career.

A participant that had previously taught sections with high stakes tests before they were removed by the state said they rarely received any pressure because "it's not a high priority because we don't have high stakes tests." Teachers of elective classes believed their locally created assessments led to lower stress levels because their test was designed towards the end of the year after their selected curriculum content was presented with considerations for their students' interests and personal teaching styles. In addition, these teachers believed their content was adequate and more closely aligned to what they deemed important as individual educators to their respective content areas.

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Effects on Morale

Within each district participating in the study, teachers with and without high stakes exams indicated their morale had been both positively and negatively impacted by the existence of high stakes tests. Experience in teachers' own classrooms or conversations and interactions with colleagues relating to high stakes testing had directly and indirectly affected teacher morale.

During the interviews teachers expressed a lower morale due to the impacts tests had on their students, teaching strategies, curriculum instruction, stress levels, APPR evaluations, and the overall lack of reliability of tests that they recognized. When asked the question "If at all, how many times in the last year has another teacher held a conversation with you that revealed they were negatively impacted by a high stakes test?" the responses showed a common theme of lowered teacher morale was pervasive in schools throughout all departments and even reaching paraprofessionals as shown in Table 2.

Participant responses indicated that high stakes testing alone was not always the antagonist in lowering morale. Additional effects related to the implementation of the Common Core and the link between test scores and APPR evaluation were more frequently mentioned when morale was discussed with both high stakes and non-high stakes teachers. Through analysis, the data indicated that teachers in core content areas had more significant impacts on their morale especially when a high stakes test was mandated.

The overall impact in one district was described by one participant as having "completely changed the culture" in a negative way and the overall use of test scores for funding or their personal performance review factored into their prospective retirement date.

The APPR rating was also mentioned as a negative influence by several teachers because of the lack of standardized resources and practices teachers have available to them from district

A COMPARATIVE STUDY OF HIGH STAKES TESTING

to district. More than four teachers in separate districts cited lower morale in their schools because colleagues in separate departments were rated as “developing” or “ineffective” teachers after their APPR scores were determined in part by high stakes testing. The research participants did not feel that their colleagues’ evaluations were representative of their abilities and were dismayed that student performance on high stakes tests negatively impacted their colleagues’ scores.

Additionally, more than half of the teachers with high stakes testing believed that student scores were not reflective of the quality of teaching practices being rated because results were unreliable and affected by problems outside of their control. Chronic absenteeism that affects scores was cited multiple times. Several teachers gave examples of students who performed well in class but were unable to produce high achieving results on exams because of test anxiety. In one instance a students’ stress was so severe that it caused “medical emergencies” where the classroom had to be cleared out for medical professionals to come in on multiple occasions. Participants across multiple districts also stated that some students may choose to “not care” about the test or said they believed inclusion classes ultimately leads to their scores being negatively impacted. Multiple teachers in a core content area that had experience with both types of classes said it was difficult for instructors to “give up that kind of control” over assessments.

Several other teachers expressed negative effects on their morale when administrators used high stakes test results to over comparing teachers to personnel in other districts. They disagreed with the practice of comparing results from multiple districts on the same test when the schools’ resources and philosophical approaches greatly differed.

In one district the building leadership had criticized a department for not producing the same results or “not doing as good of a job” on high stakes exams as a neighboring district.

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Within a couple of years the higher performing school was investigated for tampering with scores and no apology had been given to the staff that was ridiculed.

In the same district another teacher commented on the culture of the school by saying the teaching staff in the school was “monumentally frustrated” by changes associated to high stakes tests.

Interviewed teachers that were subject to high stakes tests and had a history of high level achievement scores felt that the publication of student data had helped them feel better about being a teacher. All teachers in the study that were rated effective or highly effective on their APPR scores reported feeling good about the experience and the data suggests that morale can be directly tied to student performance that affects achievement ratings.

One teacher with a Regents’ exam felt that published scores “reinforced” their performance and another said they see their students’ performance as a way to increase their “marketability” as a professional and said they have had experiences where other districts offered them a position based on the results at their present school.

Two teachers, one in a core content area and one in an elective, expressed a concern over the educational hierarchy they perceived in their schools due to high stakes testing. Each participant believed that their classes or content area were being marginalized by either administrators or department colleagues because of the absence of high stakes testing. One respondent stated “there’s a resentment to those of us who don’t have the high stakes testing” and continued to say that they believe their colleagues feel they have less pressure on them and less stress.

The same interviewee shared their perspective on the perceived decline of camaraderie in their district amongst faculty members. “For me to say it hasn’t affected me would be ridiculous

A COMPARATIVE STUDY OF HIGH STAKES TESTING

- because it's impacted the school and its impacted parent relationships and its impacted colleague relationships." According to the teacher, faculty members were more likely to spend time together outside of the classroom and communicate during school hours. He continued to say, "Now it's almost all the teachers are in their room 'til 3:05 if we're lucky that's when the students get to leave but, well that's when the late period ends for the students. But we're in our classrooms 'til, you know, 4:00, we come in at 6:00, 6:30 and we don't talk much and if we do talk, it's about this."

Effects on Curriculum Design and Instruction

The interview process revealed an abundance of impacts on pedagogical strategies, the choice of material to present in class, and the pace teachers felt they were able to teach at. Findings were consistent with previous studies and most teachers with high stakes tests in the study felt the impacts were negative and several said they were not conducive to student understanding while those with the ability to create a local assessment felt they had more freedom in the classroom to teach in ways they believed in.

Teachers that only taught class sections with high stakes tests believed they were essentially teaching to a test. One teacher described the experience as a "constant juggling act" where they had to decide what to teach and how long to spend on the content because of the probable point values each area would represent on a high stakes exam. Another teacher specifically wondered how much time they should spend in a topic if one question out of 50 would address a topic. The inability to explore topics of interests or slow down the pace to build understanding in these classes was viewed negatively by teachers.

Seven of the participants in the study had experience with sections involving high stakes tests and sections where no test was mandated. These teachers reported several differences

A COMPARATIVE STUDY OF HIGH STAKES TESTING

between their pedagogical approaches to each type of class whether they taught both types in the same school year or one in the past and one in the present.

When teaching sections with a high stakes test teachers were more likely to teach to the test and move on before student understanding was developed. One teacher mentioned tests driving “a specific pace that must be met” and compared it to their other elective sections by stating they could stop and “go deeper” in topics students are interested in. Several other teachers in the same scenario indicated they felt the same way. One wondered, “How am I going to get them to pass the test? Versus teaching them a concept, getting to know the concept?” and followed up by stating, “I’m more concerned about them performing on the test.”

In several other content areas teachers of various experience agreed that flexibility is more likely to exist in their classes without high stakes tests and they can take advantage of teachable moments opposed to feeling like they have to move on in order to cover content that would appear on the test due to standards.

Teachers generally stated they enjoyed and approved their pedagogical strategies more in classes without high stakes tests and related this to their perceived ability to do “fun” projects that are hands on, project based, allotted time for deeper understanding, and student centered. Less pressure to teach to a test allowed instructors to focus on content that is “important to teach” and one teacher mentioned a stronger sense of teaching authentic information that would be useful in the field now that the high stakes test was not in existence anymore. However, this teacher and several other participants believed they were still responsible for instruction and assessments modeled after high stakes tests in the event students take classes where they are required and made instructional choices based on this belief.

A COMPARATIVE STUDY OF HIGH STAKES TESTING

A teacher that had never had a high stakes test in their content area believed they were able to “teach the classes exactly how they wanted to”; a statement only echoed by teachers when they referred to their sections without high stakes exams.

All participants in the study would prefer to teach their content without high stakes tests. Teachers that had them believed they would be able to implement more creative lesson plans, use their personal skills, develop depth in areas of interest, use project based assessment, and effectively create assessments that align to their self-created curriculum - practices reported in elective classes. Teachers without high stakes tests felt that the possible implementation of a test in their content areas would be a “severely negative” change and they would also wind up “teaching to the test” as many of their colleagues have already done.

The changes to curriculum and pedagogical strategies appear to be directly related to pressure from building administrators and the existence of high stakes tests. The research suggests the same changes have impacted the level of stress teachers feel and the level of morale felt towards their own classrooms.

DISCUSSION

High stakes testing remains a considerable topic for debate amongst educators. Previous research and the data presented in this study suggest high stakes testing negatively impacts school culture throughout several facets of education. Disconnect between policymakers, administrators, parents, teachers, and students leads educators to discuss high stakes tests impacts on increased stress, lower morale, and the altering of pedagogical strategies in classrooms on a frequent basis.

The data compiled in this study provides evidence that the effects of high stakes tests are pervasive in public school districts and affects both teachers with and without these types of

A COMPARATIVE STUDY OF HIGH STAKES TESTING

assessments. The apparent hierarchical establishment of content areas based on the existence of high stakes testing in combination to other federal and state legislation appears to cause teachers with these tests to receive more pressure, feel more stress, and alter their curriculum in ways they are not supportive of for the purpose of producing high scores to demonstrate student achievement and understanding.

Many educators express support for accountability in the classroom but disagreed with the use of high stakes testing as a measurement tool and the overall application of student scores on standardized tests that affect their professional reviews.

While the intent of legislation at the federal and state level may not have been to cause teachers or students to experience more stress, alter their curriculum, or lead to lower morale in school buildings, the evidence suggests these are the realities many teachers face on a near daily basis due to the combination and misunderstandings of the Common Core, APPR, and Race to the Top policies adopted by New York State.

Teachers in multiple content areas empathize with one another and on multiple occasions participants spoke of a “pendulum swing” that they feel is on the horizon due to pushback related to high stakes testing. These teachers, both with high stakes tests and without them, believe that this can be attributed to mounting discontent from educators, parents, and students.

While the effects of high stakes testing are markedly more severe and rampant in classes that actually mandate standardized assessments they continue to appear in classes without them as well. Additional stress, pressure, and the establishment of an educational hierarchy will make it more difficult for teachers to collaborate and utilize teaching strategies that are considered to be of best practice.

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Implementing project based learning, differentiated instruction, collaborative techniques, inquiry based learning, executive function tasks, and data-driven instruction may be avoided by teachers as they continue to teach to a test and concentrate on rote skills in order to protect their livelihood and professional profile.

As one teacher explained, educational leaders must remain aware of what the true purposes of educational systems are and not become overwhelmed by data or misguided publications of the past. As American education pushes on with an emphasis on developing 21st century skills, leaders must examine the realities of their policies regardless of their original intents in order to correct the impacts high stakes testing has had on school culture, curriculum instruction, teacher morale, and the relationships between all parties involved.

IMPLICATIONS FOR FUTURE RESEARCH AND PRACTICE

Recent changes to educational policy in the wake of pushback in New York will require additional research to determine the effectiveness of decisions made moving forward. Qualitative and quantitative studies using elements from the results of these interviews in districts with more prevalent low income statuses and in schools with more cases of underperforming teachers may be conducted to further determine the effects of high stakes tests on different groups.

Given the findings in this study in combination with the literature presented, future research is needed to examine any impacts on student achievement due to alterations to teacher morale, stress, and curriculum instructed due to high stakes testing. Future studies may examine the same effects in larger schools while clearly differentiating between the types of high stakes standardized tests being discussed. Researchers may also be more selective in the selection of participants by carefully differentiating between teachers that have always or never had a high

A COMPARATIVE STUDY OF HIGH STAKES TESTING

stakes exam. This may help differences come to light as pedagogical strategies are less likely to overlap.

Educators of the future that look to retain autonomy in their classrooms must work with building leaders and policymakers to develop means of accountability while balancing a delicate scale that weighs the impact of assessment on instructional strategies, curriculum content, and their effects on teachers. Administrations must provide faculty members with professional development opportunities to learn more about Common Core, APPR, and Race to the Top so that they may receive additional knowledge on how to effectively teach their curriculum to standards while retaining personal values and skills in instructional strategies.

A COMPARATIVE STUDY OF HIGH STAKES TESTING

References

- Abrams, L., Pedulla, J., & Madaus, G. (2003). Views from the classroom: teachers' opinions of statewide testing programs. *Theory into Practice*, 42 (1), pp. 18-29. Retrieved from epx.sagepub.com/content/21/4/662.refs
- Barksdale-Ladd, M. & Thomas, K. (2000). What's at stake in high-stakes testing. *Journal of Teacher Education*, 51 (5), pp. 384-397. Retrieved from <http://www.soe.vt.edu/elementaryed/faculty/barksdale/docs/BarksdaleThomasHST.pdf>
- Chingos, M. (2012). "Strength in numbers: state spending on k-12 assessment systems." In *Brookings*. Retrieved from <http://www.brookings.edu/research/reports/2012/11/29-cost-of-ed-assessment-chingos>
- Coleman, James S. Equality of Educational Opportunity (COLEMAN) Study (EEOS), 1966. ICPSR06389-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-04-27. Doi:10.3886/ICPSR06389.v3
- Fletcher, D. (2009). Standardized testing. In *Time*. Retrieved from <http://content.time.com/time/nation/article/0,8599,1947019,00.html>
- Gerwin, D. & Visone, F. (2006). The freedom to teach: contrasting history teaching in elective and state-tested courses. *Theory and Research in Social Education*, 34 (20), pp. 259-282. www.tandfonline.com/doi/pdf/10.1080/00933104.2006.10473307

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Hamilton, L., Stecher, B., Marsh, J., McCombs, J., Robyn, A., Russell, J.,...Barney, H. (2007).

Standards-based accountability under No Child Left Behind. Santa Monica, CA: National Science Foundation. Retrieved from www.rand.org/content/dam/rand/people/h/27olorado_laura_s.pdf

Herman, J. & Golan, S. (1991). *Effects of standardized testing on teachers and learning -*

another look. (CSE Technical Report 334). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing. Retrieved from <https://www.cse.ucla.edu/products/reports/Tech334.pdf>

Jones, G., Jones, B., Hardin, B. & Chapmin, L. (1999). The impact of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan*, 81 (3), pp. 199-203.

<http://www.jstor.org/webdb.plattsburgh.edu:2048/stable/pdfplus/20439620.pdf>

New York State Agencies. (2014) KWIC Indicator: Children Receiving Free or Reduced-price

School Lunch - Public Schools. Retrieved from http://www.nyskwic.org/get_data/indicator_profile.cfm?subIndicatorID=52

Moon, T., Brighton, C., Jarvis, J., & Hall, C. (2007). *State standardized testing programs: their*

effects on teachers and students. (RM07228). Charlottesville, VA: The National Research Center on the Gifted and Talented. Retrieved from <http://files.eric.ed.gov/fulltext/ED505375.pdf>

A COMPARATIVE STUDY OF HIGH STAKES TESTING

No Child Left Behind. (2004). *Education Week*. Retrieved November 18, 2013, from

<http://www.edweek.org/ew/issues/no-child-left-behind/>

Pedulla, J., Abrams, L., Madaus, G., Russell, M., Ramos, M., & Miao, J. (2003). *Perceived effects of state-mandated testing programs on teaching and learning: findings from a national survey of teachers*. National Board on Educational Testing and Public Policy.

Retrieved from www.bc.edu/research/nbetpp/statements/nbr2.pdf

Powell, D., Higgins, H., Aram, R., & Freed, A. (2009). Impact of No Child Left Behind on curriculum and instruction in rural schools. *The Rural Educator*, 31 (1), pp. 19-28.

Retrieved from <http://files.eric.ed.gov/fulltext/EJ876130.pdf>

“Public Law 107-110.” (2002). In *United States Department of Education*. Retrieved from

<http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf>

“Public School Enrollment.” (2013). In *NYSED*. Retrieved from <http://www.p12.nysed.gov>

[/irs/statistics/enroll-n-staff/home.html](http://www.p12.nysed.gov/irs/statistics/enroll-n-staff/home.html)

Sabol, F.R., (2010). *No child left behind: A study of its impact on art education*. National Art

Education Foundation. Retrieved from http://www.arteducators.org/research/NCLB_Proj

[_ Report_2-10.pdf](http://www.arteducators.org/research/NCLB_Proj_Report_2-10.pdf)

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Smith, M. (1991). Put to the test: the effects of external testing on teachers. *Educational Researcher*, 20 (5). Retrieved from <http://www.jstor.org/stable/1176396> .

Smith, M., Noble, A., Heinecke, W., Seck, M., Parish, C., Cabay, M.,...Bradshaw, A. (1997). *Reforming schools by reforming assessment: consequences of the Arizona Student Assessment Program (ASAP): equity and teacher capacity building*. National Center for Research on Evaluation, Standards, and Student Teaching (CRESST). Retrieved from <http://www.cse.ucla.edu/products/reports/tech425.pdf>

Shepard, L. & Dougherty, K. (1991). *Effects of high-stakes testing on instruction*. Spencer Foundation and the Office of Educational Research and Improvement. Retrieved from <http://www.colorado.edu/education/sites/default/files/attached-files/Effects%20of%20High-Stakes%20Testing.pdf>

Stanford, J. (2013). "Bush's "Texas Miracle" debunked, lone star state sparks anti-testing revolution." In *MSNBC*. Retrieved from <http://www.msnbc.com/msnbc/bushs-texas-miracle-debunked-lone-star-st>

Urdan, T. & Paris, S. (1994). Teacher's perceptions of standardized achievement tests. *Educational Policy*, 8 (2), pp. 137-156. doi: 10.1177/0895904894008002003

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Appendix

GRAND TOUR QUESTIONS

- 1 - How do high stakes tests affect curriculum and instruction?
- 2 - How do high stakes tests affect teacher morale?
- 3 - Is high stakes testing associated with changes in teacher stress levels?
- 4 - Do high stakes tests affect teachers in different elective content areas differently?

GENERAL QUESTIONS

How many years have you been teaching?

How many years have you been teaching in your present district?

What content area/s do you teach and how long for each area?

HOW DO HIGH STAKES TESTS AFFECT CURRICULUM AND INSTRUCTION

Does the existence or absence of New York State assessments in your content area affect your curriculum planning? If so, how?

Does the existence or absence of New York State assessments in your content area affect the amount of time you are able to spend on a specific unit? If so, how?

What, if any, commercial products do you use (copies or originals) to teach your content area that relate to high stakes testing?

Do you feel that your teaching strategies are affected by the existence or absence of high stakes tests?

Are there areas of interest in your content area you'd like to teach but are not able to because of high stakes testing?

For the entire school year, approximately how much time in days, hours, or weeks do you spend practicing for high stakes tests?

How might your lessons change if New York State assessments were or were not required? Would the changes be positive or negative?

Overall, do you feel that you are able to teach your classes the way that you would like to? If not, what affects your approach?

HOW DO HIGH STAKES TESTS AFFECT TEACHER MORALE

Has high stakes testing had any effects on your sense of efficacy as a teacher and if so how?

How does the publication of student scores and district reports affect you in any way? If so, how?

If at all, how many times in the last year has another teacher held a conversation with you that revealed they were negatively impacted by high stakes tests?

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Do you think New York State assessments have had an effect on your level of satisfaction with the profession?

IS HIGH STAKES TESTING ASSOCIATED WITH CHANGES IN TEACHER STRESS LEVELS

Have you experienced any pressure from your building leadership when it comes to students performing well on tests? If so, what tests and how?

What examples of pressure from parents have you experienced in regards to testing?

How often are tests in your content area mentioned in staff meetings?

Has the existence of high stakes testing in your school affected you as a teacher and if so, how?

What is your impression of parent perspectives in regards to testing?

Has student stress related to high stakes exams related to high stakes exams carried over to you personally or in your classes?

FINAL THOUGHTS

Is there anything else you would like to add to our discussion?

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Table 1

Students Receiving Free and Reduced Lunches

	Number of Free Lunches	Percentage of Student Population	Number of Reduced Price Lunches	Percentage of Student Population
District A	69	14%	52	10%
District B	61	11%	23	4%
District C	92	11%	57	6%

Notes. 48.5% of students in New York qualify for free lunches and 52.4% qualify for reduced price lunches. Information supported by Kids' Well-being Indicators Clearinghouse (KWIC).

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Table 2

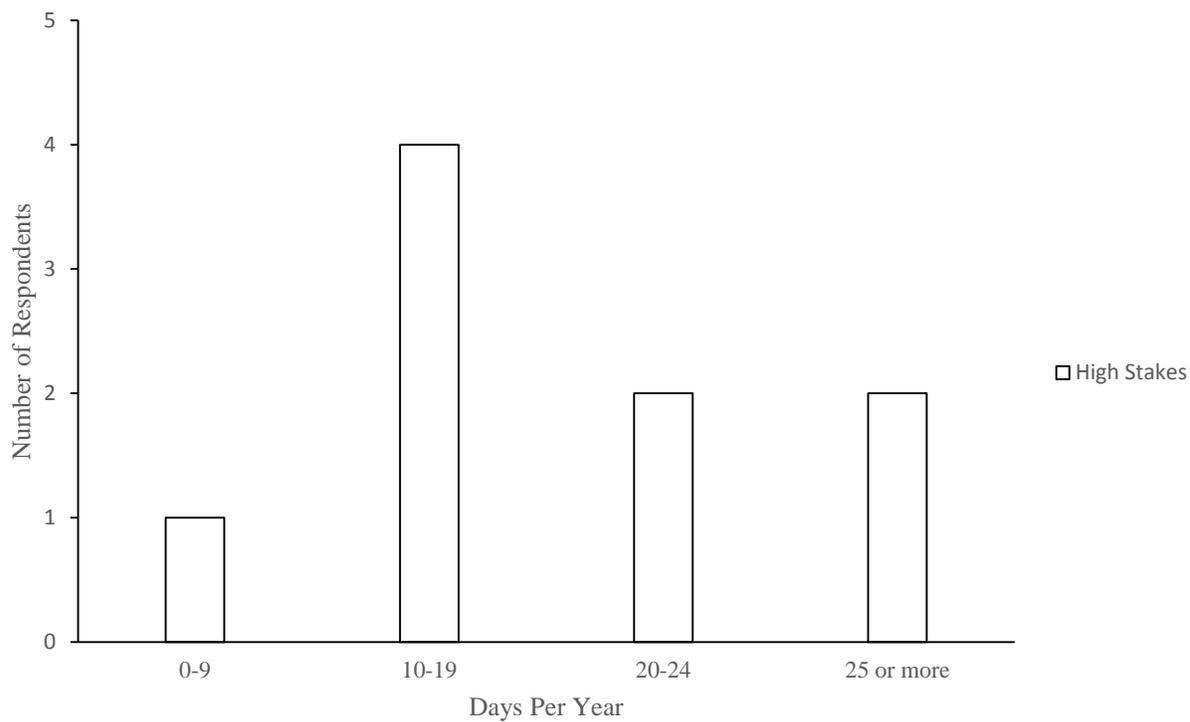
Example of Theme Development - Pedagogy

High Stakes Area	Quotation	Theme
	I believe that due to the fact that I don't have a high stakes test I have a lot more flexibility as far as my curriculum goes.	Curriculum Freedom
	I have more flexibility or more time to, you know, get more outside of the box – potentially be more creative	Curriculum Freedom
	With the 7 th grade I can do more fun things with them. Fun is more concrete too.	Curriculum Freedom
	When I had it I spent more time dissecting the exam and figuring out where the kids could get the most bang for their buck	Teach to Test
X	The test affects curriculum 100%; affects how much time is spent on a specific unit of study	Teach to Test
X	How am I going to get them to pass the test? Versus teaching them a concept, getting to know the concept, I'm more concerned about them performing on the test.”	Teach to Test
X	The test will shape what you are going to cover	Teach to Test
X	If students are interested in a topic in your class there are times when you have to cut them off because you need to move on and go to the next thing	Teach to Test

Notes. This process was repeated and included significant quotations to recognize major themes relating to the research questions and three main topics of study. Final categories were broken down into three areas: pressure and stress, morale, and curriculum instruction.

A COMPARATIVE STUDY OF HIGH STAKES TESTING

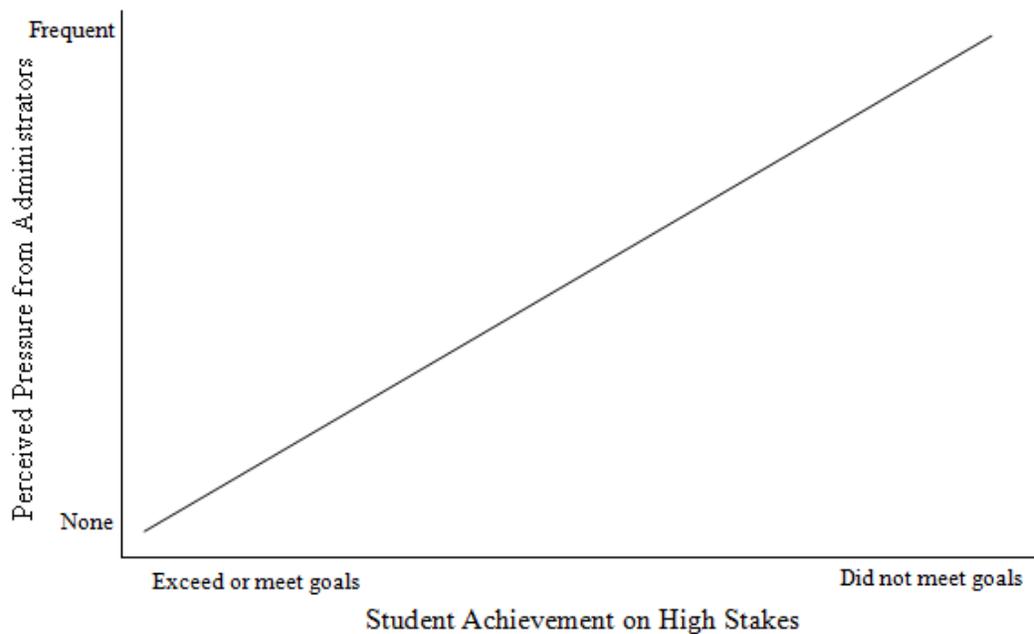
Figure 1



Note. Responses in minutes or hours per week dedicated to high stakes testing preparations are shown converted to days based on the duration of the respective district's class period.

A COMPARATIVE STUDY OF HIGH STAKES TESTING

Figure 2



Note. Teachers that met or exceeded district goals reported less conversations, emails, and communication from administrators regarding high stakes testing.